**Direct and Indirect Searches for Axion Dark Matter**

by

Joshua W Foster

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Physics)
in the University of Michigan
2021

Doctoral Committee:

Professor Aaron Pierce, Co-Chair
Assistant Professor Benjamin Safdi, University of California Berkeley, Co-Chair
Professor Oleg Gnedin
Professor Dragan Huterer
Associate Professor Joshua Spitz

Joshua W Foster
fosterjw@umich.edu
ORCID iD: 0000-0002-7399-2608

# ACKNOWLEDGMENTS

For as long as I can remember, I've dreamed of a career in physics and aspired to contribute to the collective effort of the field to ask and answer fundamental questions about our universe. My past five years in Ann Arbor have been a real privilege, and I am grateful to the people in my life who have supported and enabled my growth as not just a scientist but also a person during this time.

The obvious first person for me to thank is my advisor, Ben Safdi. Ben took me on as his student before he had even arrived at Michigan, and from day one, he had me working on projects that had a real urgency and scientific impact associated with their outcomes. Ben has always expected excellence from me, and I wouldn't be the physicist I am today without him setting high bars that I hope to have occasionally managed to clear. But on top of that, Ben has been an incredibly supportive advisor, meeting with me for several hours every day for the past five years and teaching me to be careful and detail-oriented in my work while maintaining a big-picture perspective in choosing problems that have a real chance at advancing a field with a near-infinite number of outstanding questions. Beyond that, Ben has been uncommonly supportive of my engagement with physics beyond the Safdi group through participation in conferences, summer schools, workshops, and research with external collaborators, and any understanding I have of research done in the broader particle phenomenology community is to his credit. I feel enormously lucky to have been his student, and despite our usually sarcastic dynamic, this thanks is a very sincere one.

After Ben, there are several people who have had a particularly strong impact on my academic experience. In my first project in graduate school, I was fortunate enough to work with Nick Rodd, who was beginning his last year of graduate school at the time. Since then, we've worked on a number of projects together, and despite once promising to destroy me upon realizing we were on the same job application cycle, Nick was an invaluable source of support during that time. I've looked up to Nick since we met, and I am consistently reminded that I could not have had a better role model. Chris Dessert and Malte Buschmann have been my closest collaborators at Michigan. Many of the projects we've worked on together have only been possible thanks to Malte's skill in working with `C` and `Fortran` code frameworks and Chris' impressive expertise in stellar modeling and astrophysical data reduction. I'm grateful both for our friendship and for future opportunities to work together.

Beyond Ben, Nick, Chris, and Malte, there are a number of others who have shaped my time in graduate school. I have learned a great deal from Yoni Kahn and Aaron Pierce, both of whom took a real interest in my academic and personal well-being. Lindley Winslow and the rest of the ABRACADABRA collaboration allowed me to develop an analysis framework and apply it to a new experiment's first two data collections; it was a unique experience for me as a young theorist, and I appreciate the trust they placed in me. I'm also grateful to Josh Spitz and Henriette Elvang, who were burdened with the unenviable task of teaching me how to teach. Jim Liu, who I worked with as an undergraduate through the Michigan REU program, and Finn Larsen, who was the graduate chair at the time I began at Michigan, both did an enormous amount to make me feel at home in Ann Arbor. Karen Donovan, our theory group administrator, looked out for me (likely more than I'm aware of), was always good for some department gossip, and let me raid the group's espresso stash whenever I was undercaffeinated. Shruti, Brian, Callum, graduate school would not have been the same without your friendship. That statement goes double for Andy, Chris, Johnny, Tom, and Zach, who were consistent sources of support through graduate school and our time as undergraduates at Indiana, and Jason, who is both a great friend and a survivor of three years as my roommate.

Finally, this acknowledgment would be incomplete without mention of the most important people in my life. I struggle to imagine who I would be without my parents, John and Karin, who have encouraged my ambition, curiosity, and love of science for 27 years. Finally, to my partner Natalie, meeting you has been the most exceptional thing to happen to me in the past five years. I've treasured our life together in Ann Arbor, and I am indescribably excited for our future.

# TABLE OF CONTENTS

ACKNOWLEDGMENTS . . . . . ii

LIST OF FIGURES . . . . . vii

LIST OF TABLES . . . . . xxxi

LIST OF APPENDICES . . . . . xxxiii

ABSTRACT . . . . . xxxiv

CHAPTER

**1 Introduction . . . . . 1**

1.1 The Particle Dark Matter Paradigm . . . . . 2
1.1.1 Evidence for Particle Dark Matter . . . . . 2
1.1.2 Particle Models for Dark Matter . . . . . 6
1.2 Axion and Axion-like Particle Dark Matter . . . . . 7
1.2.1 A Classical Analogy for the Strong *CP* Problem . . . . . 7
1.2.2 The Strong *CP* Problem in QCD . . . . . 9
1.2.3 Solving the Strong *CP* Problem and Dark Matter with the Axion . . . . . 10
1.2.4 Additional Axion Motivations . . . . . 12
1.3 Searching for Axion Dark Matter . . . . . 12
1.3.1 Existing Constraints on Axion Dark Matter . . . . . 12
1.3.2 Searching for the Axion through its Photon Coupling . . . . . 14
1.3.3 Statistics for Dark Matter Detection . . . . . 18
1.4 Organization of this Thesis . . . . . 21

**2 Simulations of Axion Production in the Post-Inflationary Misalignment Scenario . . . 25**

2.1 Simulation Setup . . . . . 27
2.2 Analysis and Results . . . . . 30
2.3 Conclusion . . . . . 33

**3 The Statistics of Axion Direct Detection . . . . . 35**

3.1 Organization of this Chapter . . . . . 38
3.2 A Likelihood for Axion Direct Detection . . . . . 38
3.2.1 The Statistics of the Local Axion Field . . . . . 38
3.2.2 Coupling the Axion to a Broadband Experiment . . . . . 41

3.2.3 Coupling to a Resonant Experiment . . . 47
3.3 Experimental Sensitivity . . . 49
3.3.1 Asymptotic Distribution of the Test Statistics . . . 51
3.3.2 A Procedure for Stacking the Data . . . 54
3.3.3 Expected Upper Limit . . . 56
3.3.4 Expected Discovery Reach . . . 57
3.3.5 Comparison with $S/N = 1$ . . . 59
3.4 Application to the Bulk Halo . . . 61
3.4.1 Sensitivity . . . 62
3.4.2 Parameter Estimation . . . 64
3.5 Impact of a Realistic and Time-Varying DM Distribution . . . 67
3.5.1 Halo Annual Modulation . . . 68
3.5.2 Halo Gravitational Focusing . . . 70
3.5.3 Local DM Substructure . . . 71
3.6 Conclusion . . . 75

**4 The Statistics of Axion Direct Detection with Multiple Detectors . . . 82**

4.1 The Statistics of Multiple Detectors . . . 86
4.1.1 Construction of the Axion Field . . . 87
4.1.2 The Multi-Detector Covariance Matrix . . . 88
4.2 A Likelihood for Multi-Detector Axion Direct Detection . . . 91
4.2.1 The Multi-Detector Likelihood . . . 91
4.2.2 Asimov Test Statistic . . . 92
4.2.3 Limiting Cases of Zero and Infinite Separation . . . 94
4.3 Asimov Parameter Estimation . . . 96
4.3.1 The Minimal $\mathcal{N} = 2$ Example . . . 98
4.3.2 The Infinitely-Cold Stream . . . 99
4.3.3 The (boosted) Standard Halo Model . . . 102
4.3.4 The Sagittarius Stream . . . 104
4.4 Daily Modulation . . . 106
4.4.1 A Likelihood with Daily Modulation . . . 107
4.4.2 Asimov Examples with Daily Modulation . . . 108
4.4.3 Monte Carlo Example with Daily Modulation . . . 110
4.5 Conclusion . . . 111

**5 First Results from ABRACADABRA-10 cm . . . 113**

5.1 Magnet and Cryogenic Setup . . . 114
5.2 Data Collection . . . 116
5.3 Data Analysis Approach . . . 117
5.4 Results and Discussion . . . 120

**6 Second Results from ABRACADABRA-10 cm . . . 121**

6.1 ABRACADABRA-10 cm detector . . . 122
6.2 Data Collection . . . 125
6.3 Data Analysis and Results . . . 126

6.4 Conclusion . . . 129
**7 Radio Searches for Axion Dark Matter . . . 131**
7.1 Parametrics of Indirect Detection of Axion Conversion . . . 132
7.2 Data Acquisition . . . 133
7.2.1 GBT Observations . . . 133
7.2.2 Effelsberg Observations . . . 134
7.3 Analysis . . . 134
7.4 Results . . . 137
7.5 Conclusion . . . 139
**8 X-Ray Searches for Axions from Super Star Clusters . . . 142**
8.1 Organization of this Chapter . . . 144
8.2 Axion production in SSCs . . . 144
8.3 Axion conversion in Galactic fields . . . 146
8.4 Data analysis . . . 147
8.5 Conclusion . . . 151
**9 X-Ray Searches for Decaying Dark Matter in the Milky Way . . . 152**
9.1 Data Reduction and Processing . . . 153
9.2 Data Analysis . . . 153
9.3 Interpretation . . . 156
9.3.1 Axion Interpretation . . . 159
9.4 Conclusion . . . 160
**10 Conclusion . . . 161**

APPENDICES . . . 162

BIBLIOGRAPHY . . . 345

# LIST OF FIGURES

FIGURE

1.1 A comparison in the Milky Way of measured circular velocities as a function of radius (grey) with expected theoretical contributions from various mass components within the galaxy. The galactic bulge (blue) and the galactic disk (green) are baryonic matter distributions that dominate within the inner galaxy. The expected rotation curves associated with a Milky Way that consisted only of the observed baryonic matter are shown in dashed black and are visibly incompatible with the data at large radii. The contribution of the DM to the circular velocity (red) becomes appreciable at large radii. The rotational velocities expected for the combination of galactic bulge, galactic disk, and DM halo are shown in solid black, demonstrating good agreement with the observed data. Adapted with permission from [3] using data from [4]. . . . . 3

1.2 A depiction of inferred mass distributions within the Bullet Cluster. Bright blobs indicate the location of X-ray-luminous baryonic matter, while the contours indicate the central location of the inferred mass distributions of the two merging clusters. See [13] for further details. . . . . 4

1.3 In grey, the *Planck TT* spectrum, which measures the angular scales at which temperature inhomogeneities in CMB appear. In blue, we depict the predictions of the best-fit $\Lambda$CDM cosmology obtained in the *Planck* 2018 analysis. This cosmology describes a universe in which 84% of matter is comprised of DM, and its predictions demonstrate excellent agreement with the data [15]. If we hold the total matter abundance fixed but reduce the amount of DM by 10%, then we would expect the spectrum shown in dashed green, which is visibly a poor description of the data. Here, the predicted CMB spectra have been generated with the CLASS code package [16]. . . . . 5

1.4 The axion parameter space in terms of the axion mass, $m_a$, and the strength of its coupling to photons, $g_{a\gamma\gamma}$. Green, blue, red, and gray regions are excluded by various searches. The orange band indicates the mass-coupling relation for an axion which solves the Strong *CP* Problem. White space is unconstrained. Several of these constraints have been produced by works included in this thesis. See https://cajohare.github.io/AxionLimits/ and accompanying references for details. . . . . 13

2.1 Each panel illustrates the string network (yellow strings), domain walls (red mesh), and energy density of the axion field (blue-white intensity) before (left), during (middle), and after (right) the QCD phase transition (see animation). . . . . 27

2.2 (Left) A portion of a 2-D slice through the overdensity field $\delta(x)$ at the end of the QCD stage of our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$. Large overdensities and rings of relativistic radiation arise from oscillon decay. Slices through the clustered minihalos are outlined in red. (Right) As in the left panel, except the field is evolved to matter-radiation equality. The large overdensities largely disperse and the field is everywhere non-relativistic. . . . . 31
2.3 Differential mass distribution for axion minihalos for our most realistic simulation, as described in Fig. 2.2, computed by clustering the overdensity field at $\hat{\eta}_{\rm MR}$. The shaded "unresolved" region denotes the parameter space that is beyond our resolution limit. Small statistical uncertainties are displayed as grey error bands. . . . . 32
2.4 The DM density $\Omega_a$ as a function of the axion decay constant $f_a$, with statistical uncertainties (black) and correlated systematic uncertainties (red) indicated, for our top four simulations. We compare our results to those in [112] (Klaer and Moore), which agree relatively well with our own, and [118] (Kawasaki et al.), which predicts significantly higher $\Omega_a$ relative to what we find. . . . . 33

3.1 (Left) A comparison between the mean of 500 Monte Carlo simulations of a signal only PSD dataset (blue) and the analytic expectation given in (3.26) (black). The inset shows the distribution of the 500 simulated $S_{\Phi\Phi}$ versus the predicted exponential distribution, as in (3.24), at the frequency where the signal distribution is maximized, $\omega/m_a \approx 1.003$. This example was generated assuming the unphysical but illustrative parameters $A = 1$ Wb$^2$, $m_a = 2\pi$ Hz, and $v_0 = v_{\rm obs} = 220{,}000$ km/s. Importantly the simulations were generated by constructing the full axion field starting from (3.3), and so the agreement between theory and Monte Carlo is a non-trivial confirmation of the framework. (Right) As on the left, but with Gaussian distributed white noise added into the time-series data with variance $\lambda_B/\Delta t$, and taking $\lambda_B = 500\,{\rm Wb}^2\,{\rm Hz}^{-1}$. Again we see the theory prediction in good agreement with the average data, whilst at an individual frequency point the simulated data is exponentially distributed. See text for details. . . . . 44
3.2 A comparison between the look elsewhere effect improved survival function approximate result derived between (3.60) and (3.62), and the equivalent values derived directly from Monte Carlo simulations. The good agreement between the two, especially at large ${\rm TS}_{\rm thresh}$ demonstrates that our approximate result is useful for estimating how often the background can fluctuate to fake the signal at a given confidence level. Note the values plotted here correspond to signals varying from 0 to $4\sigma$, for derived values of $\lambda_B$ given in (3.28) and 2.5 million Monte Carlo simulations. We do not extend the plot up to the $5\sigma$ value relevant for discovery, as this would require roughly 100 times as many simulations. This statement in itself already demonstrated the utility of our approximate analytic result. . . . . 60

3.3 (Left) A comparison of the projected sensitivities for a hypothetical version of the ABRACADABRA (ABRA) experiment [71], with inner toroidal radius $R = 0.85$ m, an outer toroidal radius double this value, and a height $h = 4\,R$. A maximum magnetic field of 10 T is assumed, along with an interrogation time of 1 year. (Right) An equivalent comparison of projections for a future ADMX experiment. Here we take a total run time of 5 years, a volume of 500 L, quality factor of $10^5$, magnetic field of 7 T, and a system temperature of 148 mK. In both panels the exact sensitivities are contrasted with an estimate obtained from the signal-to-noise ratio, $S/N = 1$. . . 62
3.4 An actual limit obtained from a single Monte Carlo simulation, with the broadband readout, compared to the various expectations for the broadband ABRACADABRA framework used in Fig. 3.3. The data was simulated with an injected signal corresponding to $m_a = 10^{-8}$ eV and $g_{a\gamma\gamma} = 2.21 \times 10^{-16}$ GeV$^{-1}$, and indeed we can see that right near the frequency corresponding to the axion mass, we are unable to exclude the corresponding signal strength. . . . . 65
3.5 The posterior distribution for a model with annual modulation where the signal strength is at the threshold of annual modulation detection at $5\sigma$. The true parameter values are indicated in blue, with the $1\sigma$ confidence intervals on the parameter estimations indicated by the dashed black lines in the one parameter posteriors. The two parameter posteriors show the 1 and $2\sigma$ contours. The axion mass, $m_a$, was also scanned over, and is recovered accurately but not shown here. Note that this example uses the Asimov dataset. All times are measured in days and velocities in km/s, while the units of $A$ are arbitrary. . . . . 77
3.6 As in Fig. 3.5, but this time the data includes gravitational focusing and the model only includes the parameters $A$, $\alpha$ and $\bar{t}$. (Left) Gravitational focusing, while present in the Asimov data, is excluded from the model template. The estimations of $A$ and $\bar{t}$ are off at the $\sim 2\sigma$ and $\sim 1\sigma$ levels, respectively. (Right) As in the left panel but including gravitational focusing in the model template. As expected, the parameter estimation is quite accurate in this case. . . . . 78
3.7 The enhancement expected in the TS in the presence of a coherent DM stream. The TS is shown as a ratio with respect to the case where only the bulk halo is present and as a function of the fraction of the local DM within the substructure. . . . . 78
3.8 The axion contribution to the PSD as a function of frequency in the presence of DM substructure. (Left) We show the effect of a Sagittarius-like stream that makes up $\sim$5% of the local DM density at two different times of year, corresponding to the dates of maximum TS (June 5) and minimum TS (November 23), where all dates are for 2017. Annual modulation plays an important role for cold substructure because the Earth's orbital velocity may be larger than the substructure velocity dispersion. (Right) As in the left panel, but for a dark disk that makes up $\sim$20% of the local DM density. The dark-disk is co-rotating with the baryonic disk, with a lag speed $\sim$50 km/s, and so the contribution to the PSD is at lower speeds compared to the stream case. Gravitational focusing also plays an important role for the disk since the solar-frame velocities are relatively low. In this case the maximum and minimum TS occur on November 18 and June 5 respectively. For both of these panels, the signal is generated using $m_a = 1$ MHz, $A$ set to the value for the threshold for detection of the SHM, and $\lambda_B$ set to the minimum SQUID noise. . . . . 79

3.9 A Monte Carlo parameter estimation for the bulk halo parameters at the threshold of detection for annual modulation in the presence of a Sagittarius-like stream containing 5% of the DM and with a narrow velocity dispersion of 10 km/s. The accuracy of the parameter scan is worsened by the failure to account for the substructure in the analysis. . . . . 80
3.10 A simultaneous Monte Carlo parameter estimation for a signal containing a bulk halo and a Sagittarius-like stream with 5% of the DM using identical seed parameters as Fig. 3.9. Scanning for the bulk halo and substructure simultaneously allows us to accurately recover the signal parameters. Left, the bulk parameter scan results, right, the stream parameter scan results. . . . . 81

4.1 The imprint of DM interferometry. A single wave-like DM experiment is sensitive to the DM speed distribution $f(v)$. Two detectors separated by a vector $\mathbf{x}_{12}$, however, are sensitive to the speed distribution modulated by the $\mathbf{k}\cdot\mathbf{x}_{12}$ phase of the DM wave, replacing $f(v)$ with functions $\mathcal{F}_{12}^{c,s}(v)$ as defined in (4.3). As the figures demonstrate, the modified speed distributions exhibit daily modulation and carry additional information about the velocity distribution $f(\mathbf{v})$ that would be invisible to a single detector. For this example we take $m_{\rm DM} = 25.2\ \mu$eV [79], near the window where the HAYSTAC collaboration is searching for axion DM. Taking the Standard Halo Model ansatz for $f(\mathbf{v})$ in (4.50), we place one detector at a latitude and longitude of $(41^\circ\ \mathrm{N},\ 73^\circ\ \mathrm{W})$, and a second instrument $\sim 20$ m to the North, corresponding to $d \sim 2\lambda_c$. A curve is shown for every ten minutes starting from midnight on January 1st of 2020. Note that as $\mathcal{F}_{12}^{c,s}(v)$ are functions of $m_{\rm DM}d$, qualitatively similar effects exist for *e.g.* $m_{\rm DM} \sim 10^{-9}$ eV, in the mass range probed by ABRACADABRA and DM-Radio, for $d \sim 500$ km. . . . . 83
4.2 The modified speed distribution, $\mathcal{F}_{12}^{c}(v)$, that carries the imprint of DM interferometry. Here we show the particularly simple example of an isotropic SHM for $\mathcal{N} = 2$ detectors, in which case the expression is given in (4.40). The result is shown for various choices of the two detector separation $d$ as compared to the axion coherence length $\lambda_c = (m_a v_0)^{-1}$, with $v_0 = 220$ km/s. The limiting cases of $\mathcal{F}_{12}^{c}(v) \to f(v)$ for $d \ll \lambda_c$ and $\mathcal{F}_{12}^{c}(v) \to 0$ for $d \gg \lambda_c$ are apparent. For $d \sim \lambda_c$, however, the profile is modulated with the interference inherent in the cross-spectrum. In this simple case, there is no additional information about the velocity distribution that may be extracted by having multiple detectors. . . . . 98
4.3 The expected uncertainty on the angle between the detector axis and solar velocity, $\theta_\odot = \arccos(\hat{\mathbf{v}}_\odot \cdot \hat{\mathbf{x}}_{12})$, as a function of $d/\lambda_c = d \times m_a v_0$. In this example we have set the true orientation to $\theta_\odot^t = \pi/4$. With this configuration, we find that the maximum precision is obtained for $d \approx 2\lambda_c$. . . . . 104

4.4 (*Left*) A Mollweide projection of the Asimov test statistic $\widetilde{\Theta}(\theta, \phi)$ for the SHM divided by the co-located detection significance $\mathrm{TS}_0$. The detectors are configured so that the displacement vector between them is parallel to the SHM boost velocity, and the Mollweide plot is rotated so that it is centered on the maximum test statistic. (*Center*) As on the left, but for a detector configuration where the displacement vector is at a $45^\circ$ angle to the North ($\theta^t_\odot = \pi/4$) with respect to the SHM boost velocity. (*Right*) As on the left, but for a detector configuration where the displacement vector is perpendicular to the SHM boost velocity ($\theta^t_\odot = \pi/2$). In this configuration the location of the boost velocity can only be localized to a great circle on the celestial sphere. . . . 105

4.5 (*Left*) As in Fig. 4.3, but for the Sagittarius (SGR) stream rather than for the SHM. As before, the maximum precision for the inferred value of $\theta_{\mathrm{str}}$ is achieved at $m_a v_0 d \approx 2$, although the overall dependence is somewhat softened outside of the extremes at $m_a v_0 d = 0$ and $m_a v_0 d = 2\pi$. The values of $\sigma_{\theta_{\mathrm{str}}} \times \sqrt{\mathrm{TS}_0}$ are also considerably smaller than those found in the SHM example, indicating that the angle $\theta_{\mathrm{str}}$ can be reconstructed with much greater precision for the SGR stream as compared to the SHM. (*Right*) The Asimov TS $\widetilde{\Theta}(\theta_{\mathrm{str}})$ for the SGR stream rescaled by the co-located detection significance $\mathrm{TS}_0$ as a function of $\theta_{\mathrm{str}}$ for a detector configuration where the true stream direction is $\theta^t_{\mathrm{str}} = \pi/4$ (dashed vertical line). We have fixed $m_a v_0 d = 2$. The TS $\widetilde{\Theta}(\theta_{\mathrm{str}})$ is maximized at the true value of $\theta_{\mathrm{str}}$, but there is considerable nontrivial global structure with a large number of local minima and maxima in $\widetilde{\Theta}$. . . . 106

4.6 As in Fig. 4.4, we construct Mollweide projections of the Asimov test statistic $\widetilde{\Theta}(\theta, \phi)$ for the SHM rescaled by the co-located detection significance $\mathrm{TS}_0$. However, we now perform a joint likelihood over data collected over a 24-hour period so that the daily modulation of the detector displacement vector produces a time-varying signal, which helps break degeneracies in the reconstructed directional parameters. The Mollweide projection for a configuration in which the detectors are oriented along an East-West (North-South) orientation is shown on the left (right). While the results of obtained in an East-West configuration do not depend on the latitude of the detectors, the North-South configuration results do, so for definiteness, we have taken the detectors to be located in New Haven, CT, the site of the HAYSTAC detector. In both configurations, the SHM boost velocity direction can be localized effectively, although there remains a non-trivial degeneracy in the East-West map between two points on the sphere. . . . 108

4.7 As in Fig. 4.6, but for the Sagittarius stream example. For a fixed axion mass, the physical detector separation $d = 2\lambda_c$ is a factor of 20 larger than in Fig. 4.6 because of the larger coherence length of the stream. While there are many local maxima in both configurations, the North-South orientation produces only a single global maximum, at the true detector localization, while the East-West orientation leads to two degenerate global maxima (one at the true detector location and the other displaced). An animated version of these figures, showing how the localization improves throughout the day as more orientations of $\mathbf{x}_{12}$ are sampled, can be found at github.com/joshwfoster/DM_Interferometry. . . . . 109

4.8 The posterior distribution for a model with daily modulation where the signal strength is at the threshold of an expected $5\sigma$ detection for a 100 second observation with a single detector. Monte Carlo data is generated for 24 hours of data collection with two detectors separated along the North-South direction by a distance with $2 \times (m_a v_0)^{-1}$. The true parameters are indicated in blue, with the $1\sigma$ confidence intervals on the parameter estimations are indicated by the dashed black lines in the single-parameter posteriors. The two parameter posteriors show the $1\sigma$ and $2\sigma$ contours. On the left, we display the posterior distributions for the overall signal strength, the boost speed of the SHM, and the velocity dispersion of the SHM, all of which are parameters accessible in a single detector configuration. On the right, we display the posterior distributions for the angles $\Delta\theta_\odot = \theta_\odot - \theta^t_\odot$ and $\Delta\phi_\odot = \phi_\odot - \phi^t_\odot$ which specify the orientation of SHM boost velocity and are only accessible in a multiple-detector configuration. Both $\theta_\odot$ and $\phi_\odot$ are determined at degree precision in this scenario. . . . . 111

5.1 *Left:* Rendering of the ABRACADABRA-10 cm setup. The primary magnetic field is driven by 1,280 superconducting windings around a polyoxymethylene (POM) support frame (*green*). The axion-induced field is measured by a superconducting pickup loop mounted on a PTFE support (white). A second superconducting loop runs through the volume of the magnet to produce a calibration signal. All of this is mounted inside a superconducting shield. *Right:* Picture of the exposed toroid during assembly. . . . . 115

5.2 Flux spectrum averaged over the the data used in this analysis. (*a*) The spectrum over the frequency range $11\,\mathrm{kHz} < f < 3\,\mathrm{MHz}$, corrected for the pre-digitizer filters (*blue*). For comparison, we also show the digitizer noise floor, corrected for pre-digitizer filters (*gray*) and the characteristic SQUID flux floor (*green dashed*). The axion search range is between the dotted black lines. (*b*) A zoomed view of the 10 MS/s spectrum (*blue*) with $\Delta f = 100\,\mathrm{mHz}$ and and an example axion signal at the 95% upper limit (*red dashed*). (*c*) A zoomed view of the 1 MS/s spectrum with $\Delta f = 10\,\mathrm{mHz}$. Note that the digitizer data was collected at a different time from the SQUID data, and shows a few transient peaks that are not present in the SQUID data. 116

5.3 The limit on the axion-photon coupling $g_{a\gamma\gamma}$ constructed from ABRACADABRA-10cm data described in this work. We compare the observed limit, which has been down-sampled in the number of mass points by a factor of $10^4$ for clarity of presentation, to the expectation for the power-constrained limit under the null hypothesis. This down-sampling excludes the 87 isolated mass points vetoed in the discovery analysis; further details will be presented in [82]. Additionally, we show the astrophysical constraint on $g_{a\gamma\gamma}$ in this mass range from the CAST helioscope experiment [263]; the region above the grey line is excluded. . . . . 119

6.1 *Top*: Schematic of ABRACADABRA-10 cm showing the effective axion-induced current (blue), sourced by the toroidal magnetic field, generating a magnetic flux (magenta) through the pickup cylinder (green) in the toroid bore. *Bottom*: Simplified schematic of the ABRACADABRA-10 cm readout (full circuit diagram in Fig. E.1). The pickup cylinder $L_p$ is inductively coupled to the axion effective current $\mathbf{J}_{\rm eff}$. The power spectrum of the induced current is read out through a DC SQUID inductively coupled to the circuit through $L_{\rm in}$. An axion signal would appear as excess power above the noise floor at a frequency corresponding to the axion mass. . . . . 123
6.2 The gain shown here is defined as the change in amplifier output voltage over a corresponding change in input flux amplitude on the pickup cylinder ($\partial V_{\rm out}/\partial \Phi_a$). Both transfer functions roll off at high frequencies due to the amplifier bandwidth, which we estimate to have a cutoff frequency of approximately 1 MHz. We believe the difference in calculated and measured gain is due to inconsistency in the total inductance of the pickup circuit. . . . . 125
6.3 The survival function of TS values from the likelihood analysis of the Run 3 results. The $y$-axis indicates the fraction of mass points tested with a discovery TS at or above the value on the $x$-axis. Under the null hypothesis, the distribution should follow the survival function of the one-sided $\chi^2$ distribution with one degree of freedom ("Expected," dotted gray). This is indeed the case after data cleaning for *e.g.* single-channel excesses in time slices, magnet-off vetoes, and the inclusion of a systematic nuisance parameter, which is tuned in a sliding window at $4\sigma$ local significance to give the correct number of excesses at or above that significance, masking the signal of interest. No excesses are found beyond our indicated $5\sigma$ LEE-corrected discovery threshold. . . . . 128
6.4 (*Left*) The one-sided 95% upper limit (U.L.) on $g_{a\gamma\gamma}$ from this work excludes previously unexplored regions of ADM parameter space. The $1\sigma$ and $2\sigma$ containment regions are constructed by taking the appropriate percentiles of the distributions of the limits over narrow mass ranges; note that this means that $\sim$16% of the upper limits lie at the bottom of the green band. Around 11.1 million mass points are analyzed, so the plotted data is smoothed for clarity. Our limits surpass those from a number of indicated astrophysical and laboratory searches in this mass range, including CAST (solid grey region), super star cluster constraints (dashed grey line), and SHAFT (solid grey line). See text for details. (*Right*) The un-smoothed limit in a narrow mass range between 2.99790 and 2.99798 neV. This provides a detailed view of variations in the limit at each axion mass that arise from statistical fluctuations across the collected data that are not visible in the smoothed data shown in the left plot. This range also depicts the location where our maximum sensitivity is achieved, with our strongest limit at $g_{a\gamma\gamma} \lesssim 3.2 \times 10^{-11}\ \mathrm{GeV}^{-1}$. . . . . 129

7.1 The 95% upper limits on the signal flux for the indicated sources from the GBT and Effelsberg observations. These upper limits apply to monochromatic signals at the widths $\delta f_{\rm fid}$ given in Tab. 7.1. These curves have been down-sampled for visualization purposes. We compare these limits with the 95% upper limits expected from the ideal radiometer equation under the assumption that the only source of statistical uncertainty is thermal noise at the total system temperature. . . . . 137

7.2 The one-sided 95% upper limits on $g_{a\gamma\gamma}$ as a function of the axion mass $m_a$ from this work are shown as colored lines (GBT INS observations) and black lines (Effelsberg GC observations). Previous limits from the CAST helioscope and the UF and RBF haloscopes are shown in shaded grey. The range of couplings expected for the QCD axion is shaded in orange. Note that the fiducial GC limits assume an NFW DM profile and the conservative NS population model (Model II) from [104]. The green band depicts theoretical uncertainties on the $g_{a\gamma\gamma}$ limit associated with the GC analysis for the Effelsberg data. The top of the band assumes an NFW DM density profile with a 0.6 kpc core, while the bottom of the band uses the alternate NS population model in [104] (Model I). . . . . 140

8.1 The stacked and pixelated background-subtracted count data (10 - 80 keV) from the NuSTAR observations of the Quintuplet SSC. The locations of the stars are indicated in black, while the 90% energy containment region for emission associated with the SSC is indicated by the black circle, accounting for the NuSTAR point spread function (PSF). $RA_0$ and $DEC_0$ denote the locations of the cluster center. We find no evidence for axion-induced emission from this SSC, which would follow the spatial counts template illustrated in the inset panel. . . . . 143

8.2 The spectra associated with the axion-induced templates from the Quintuplet and Wd1 SSCs constructed from the NuSTAR data analyses, with best-fit points and $1\sigma$ uncertainties indicated. In red we show the predicted spectra from an axion with $m_a \ll 10^{-11}$ eV and indicated $g_{a\gamma\gamma}$. Note that for Wd1 we do not analyze the 10 - 15 keV energy bin because of ghost-ray contamination. . . . . 149

8.3 The 95% upper limits (black) on $g_{a\gamma\gamma}$ as a function of the axion mass from the Quintuplet and Wd1 data analyses. We compare the limits to the $1\sigma$ (green band) and $2\sigma$ (yellow band) expectations under the null hypothesis, along with the median expectations (dotted). The joint 95% upper limit, combining Quintuplet and Wd1, is also indicated (expected joint limit not shown). At low masses our limits may be surpassed by those from searches for $X$-ray spectral modulations from NGC 1275 [364], though we caution that those limits have been called into question recently, as discussed further in the text [365]. . . . . 150

9.1 Our fiducial $D$-factor, which is proportional to the expected DM signal flux. Values are given in all 30 annuli, which are $6^\circ$ wide in angular distance from the GC (with $|b| > 2^\circ$), and we define a signal and background ROI as shown. In each ring, we compute the $D$-factor of all MOS or PN exposures, weighted according to the observation time and field of view. The horizontal line indicates $D_{\rm bkg}$, the mean $D$-factor in the background ROI. . . . . 154

9.2 The background-subtracted MOS data for the innermost annulus, downbinned by a factor of four for presentation purposes. The indiciated best fit null and signal models, for a 3.5 keV UXL, are constructed using the GP modeling described in the text. . . . 157

9.3 (Upper) The power-constrained 95% upper limit on the DM lifetime from this work, presented in the context of the sterile-neutrino mixing angle $\sin^2(2\theta)$, as a function of the DM mass $m_\chi$. The dark grey regions correspond to theoretical bounds for DM underproduction in the $\nu$MSM or bounds from previous X-ray searches (1)–(5); see text for details. (Lower) The associated sign-weighted significance for the UXL. Vertical grey regions denote background lines and are at least partially masked. Green and gold regions indicate 1/2$\sigma$ expectations under the null hypothesis. These results are shown in the context of more general DM models as constraints on the DM lifetime in Fig. H.7. . . . . 158
9.4 Our decay search interpreted in the context of limits on the axion decay to two photons. Our limits, along with those from additional blank sky searches are indicated as in Fig. 9.3 . . . . . 160

A.1 Double differential mass fractions for axion minihalos as a function of the concentration parameter $\delta$ and mass $M$. In the top left we compute that mass function using the field immediately after the QCD phase transition, at $\hat{\eta} = 7$, while in the bottom left we use the more correct procedure of first evolving to $\hat{\eta} = \hat{\eta}_{\rm MR}$ before performing the clustering procedure. Evolving to matter-radiation equality gives the most overdense regions time to expand and results in less dense overdensities, as compared to the incorrect procedure shown in the top left. This is perhaps even more apparent in the single differential mass fractions as a function of the mass $M$ (top right) and concentration parameter $\delta$ (bottom right). These results are based on our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$. Error bars are statistical, and we do not extend the $df/d\log M$ curves to lower masses as we are unable to resolve those properly. . . . . 168
A.2 Illustration of an oscillon ($\log(\rho/\bar{\rho})$) at different times in a 2D simulation. Two scenarios are considered with different truncation points of the mass growth, $\hat{\eta}_c = 4.0$ and $\hat{\eta}_c = 6.0$. The three left panels are identical in both scenarios, while the two top right panels are for $\hat{\eta}_c = 6.0$, and the two bottom panels are for $\hat{\eta}_c = 4.0$. The radius of the oscillon is proportional to the oscillation frequency $\sim m_a(T)^{-1}$ (circles of that radius are shown in dashed blue) and as such is decreasing over time. The oscillon central density slowly dissipates after the mass growth ends, as seen in the bottom right panels for $\hat{\eta}_c = 4.0$. . . . . 170
A.3 Our results for the DM density today $\Omega_a$, inferred at $\hat{\eta}_{\rm MR}$, from simulations at different values of $\tilde{\lambda}$ for our most realistic $\hat{\eta}_c$: $\hat{\eta}_c = 3.6$. The uncertainties are the inferred statistical uncertainties arising from the spread in the DM density determinations as a function of $\tilde{\lambda}$. No trend is discernible for the dependence of $\Omega_a$ on $\tilde{\lambda}$, above the statistical noise. . . . . 171
A.4 A comparison of the predictions for the relic abundance of axions dark matter as a function of $f_a$ obtained in [118] (Kawasaki et al.) and [112] (Klaer and Moore) with the simulation results realized in this work. Error bars are combined statistical and correlated systematic errors, with the former dominating at $\hat{\eta} = 7$ due to large field gradients and the latter at $\hat{\eta} = \hat{\eta}_{MR}$. . . . . 172

A.5 A comparison of the power spectra realized in simulations for $\tilde{\lambda} = 5504$ for different choices of $\hat{\eta}_c$. New features in the power spectrum emerge as we push to larger values of $\hat{\eta}_c$, and we cannot exclude the possibility that further features would emerge were we to simulate with a greater value for the cutoff. On the other hand, the power-spectrum is highly non-Gaussian at small scales, so the distribution $\Delta_k^2$ alone is not adequate for understanding the small-scale nature of the overdensity field. . . . . 174
A.6 A comparison of the distribution of the squared magnitudes of Fourier components for four different fixed reference momentum $k$. The expected exponential distribution for a Gaussian field is also indicated. While the distributions are Gaussian at large scales, they become increasingly non-Gaussian at small scales. The momentum mode $|\mathbf{k}| = 500$ corresponds to approximately 6.5 grid sites. These distributions were constructed from our most realistic simulation with $\tilde{\lambda} = 5504$ and $\eta_c = 3.6$. . . . . 175
A.7 Comparison between differential mass fractions as a function of the minihalo mass $M$ from our simulations at different $\hat{\eta}_c$. In this plot we have rescaled the minihalo masses such that we achieve the correct DM density $\bar{\rho}$ observed in the Universe, but for the solid curves we have not applied the Hubble volume rescaling factor to reach our target $f_a$. However, the dashed curves do have the Hubble volume rescaling factor included, but here we take our target $f_a$ to be that corresponding to our most realistic simulation with $\hat{\eta}_c = 3.6$. The difference between the dashed mass functions and the solid black mass functions gives a sense of the systematic uncertainty introduced by applying the naive mass rescaling factors instead of simulating with the correct value of $\hat{\eta}_c$ ($f_a$). . . . . 176
A.8 Comparison between differential mass fractions as a function of the concentration parameter $\delta$ and minihalo mass $M$ for different $\hat{\eta}_c$ and $\tilde{\lambda}$ at $\hat{\eta} = \hat{\eta}_{\rm MR}$. Error bars are statistical. Shown as dotted lines is a fit to the $df/d\delta$ curves as described in the text. We do not extend the $df/d\log M$ curves to lower masses as we are unable to resolve those properly. . . . . 177
A.9 Comparison between cumulative mass fractions, defined in the text, for our simulation at $\hat{\eta} = 7$ (solid blue) and $\hat{\eta}_{\rm MR}$ (solid black). We use our fit to the differential mass fraction $df/d\delta$ to extrapolate to high $\delta_0$ for our $\hat{\eta}_{\rm MR}$ data (dotted black). Error bars are statistical. We compare our results to those from Kolb and Tkachev [430] obtained at $\hat{\eta} = 4$ by using the fit to their data presented in [99] (red curve). . . . . 178
A.10 We depict the variation of $f_1$, $f_2$, and $f_3$ as a function of $\hat{\eta}$ over the relevant range of $\hat{\eta}$ for our simulations accounting for a varying $g_*$. For fixed $g_*$, we would expect $f_1$ and $f_2$ to be constant at value 1. We additionally show the behavior of $f_3$, which describes the evolution of the quantity $m_a(\eta)^2/H_1^2$, normalized to $\tilde{f}_3$, wherein we compute $m_a(\eta)^2/H_1^2$ assuming a fixed $g_*$. Assuming a fixed $g_*$ causes the axion to reach its zero-temperature value earlier in $\hat{\eta}$, but the ratio ultimately reaches unity as the same zero-temperature mass is reached. . . . . 183

A.11 Numerical evolution of the idealized circular string-domain wall collapse. (*Left*) A comparison of the radius of the circular domain wall as a function of conformal time $\hat{\eta}$ for the simulation parameter $\tilde{\lambda} = 5504$ (solid black) and for a physically-motivated parameter value $\tilde{\lambda} \approx 10^{30}$. The collapse of the domain wall occurs at around $\hat{\eta} = 2$, *i.e.*, after a Hubble time. (*Right*) The ratio of the domain wall radius as a function of $\hat{\eta}$ for the two values of $\tilde{\lambda}$. We see that the collapse rate is largely insensitive to the value of $\tilde{\lambda}$. . . . . 185
A.12 The string length parameter $\xi$ shown as a function of the ratio of simulation temperature $T$ to the temperature $T_{PQ}$ at which the PQ phase transition occurred, including the results of our fit to the functional form of (A.58). We observe significant evidence for logarithmic deviation from the scaling regime. Extrapolating this result to the QCD phase transition (vertical dashed line) gives the prediction that $\xi$ should be around a factor of 15 higher at the beginning of the QCD phase transition than in the final state of our most realistic simulation. . . . . 187
A.13 The fraction of the string length in super-horizon length strings. Like [124], we find roughly $80\%$ of the string length resides within long strings. . . . . 188
A.14 The present day axion abundance as a function of the string density parameter $\xi$ at the beginning of the QCD simulation at $\hat{\eta}_i = 0.4$. Individual data points are labeled by their box length $L_{\rm QCD}$. The error bars are estimates of the statistical uncertainties, and no clear trend is visible in the data. . . . . 188
A.15 Differential mass spectrum as a function of the minihalo mass $M$ for different box sizes. Error bands include statistical errors and the uncertainty on the overall normalisation. . . . . 188

B.1 A comparison between the variation in the 95% upper limit found in Monte Carlo (MC) simulations to that derived analytically with the Asimov dataset. As shown the two are in good agreement. . . . . 197
B.2 Schematic depiction of the approximation made to the model used to derive $\mathrm{TS}_{\rm thresh}$. Specifically we assume that the signal model is non-zero only within a finite frequency range, and equal to the background outside this, and within this range the combined signal and background is flat. . . . . 198

C.1 Monte Carlo validation that the statistics of DM interferometry are as claimed in App. C.2. In the left figure we confirm that the variances of the real and imaginary signal-only data sets, collected for the $\mathcal{N} = 2$ experiments, is as claimed in (4.16). This was proven directly in the text, but in the plot we show that the average of 4,000 Monte Carlo simulations provides a consistent prediction for the variances as a function of frequency in the different cases. On the right figure, for the frequency where $\langle R^{(1)} R^{(1)} \rangle$ achieves its maximum, we show the distribution of values across the simulations. In detail, we see that the real and imaginary components are normally distributed, and consistent with a mean-zero normal distribution, where the variance is given as on the left, here $\sigma^2 \approx 25\ \mathrm{Wb}^2/\mathrm{Hz}$. We found that the distributions were consistent with the Gaussian expectation at the level of $p > .05$ using the D'Agostino and Pearson omnibus normality test [432, 433]. In both cases, each Monte Carlo simulation involves a direct construction of the axion field starting from (C.16) with $N_a = 100,000$, taking $m_a = 2\pi$ Hz, and $A = 1\ \mathrm{Wb}^2$. Further, we take the velocity distribution to follow a variant of the SHM in (4.50), but with $v_0 = 0.07$ and $\mathbf{v}_\odot = (0, 0.08, 0)$, both in natural units. The (unphysically) large velocity helps simplify the computation of the Fourier transform. The detector separation is $\mathbf{x}_{12} = d(0, 1, 0)$, with $d \approx 4.4\lambda_c$. . . . . 206

D.1 (a) Three of the 80 Delrin wedges that form the toroid structure stacked together. The black bar indicates a $\approx$ 1 cm scale. (b) Cutaway rendering of the toroid with the 1 mm diameter wire pickup loop in the center. A 0.5 mm diameter wire runs through the center of field region to form the calibration loop. Toroid height is $\approx$ 12 cm. (c) Rendering of the ABRACADABRA-10 cm support structure. The pickup loop is supported by a PTFE (white) tube through the center. The magnet is supported by an outer G10 support structure and thermalized with two copper bands. (d) Photo of the assembled ABRACADABRA-10 cm, with the top of the superconducting shield and support structure removed. . . . . 214

D.2 Gain of the combined high-pass and anti-aliasing filters. All spectra are corrected for this response function – unless otherwise noted. Measured in-situ, using injected signals at different frequencies. This also defines the usable range of data. For our search we use the range 75 kHz – 2 MHz. . . . . 216

D.3 A conceptual diagram of the ABRACADABRA-10 cm calibration circuit. The calibration loop, $L_C \approx 300$ nH, is concentric with the pickup loop, $L_P = 95.5$ nH. The circuit is plugged into the SQUID with input inductance $L_{\mathrm{in}} \approx 150$ nH. The parasitic resistance in the circuit is measured as $R_P \approx 13\,\mu\Omega$. . . . . 218

D.4 Low frequency SQUID spectra from ABRACADABRA-10 cm taken with an accelerometer attached the 300 K plate. The spectrum is that of the SQUID output, with the degree of correlation with the accelerometer indicated by color (i.e. the correlation coefficient). The accelerometer begins to lose sensitivity above a few kHz, so it is not clear from this measurement how far up the correlation continues. These data were taken with a larger dynamic range on the digitizer, so have a relatively high ADC noise floor of $\sim 5 \times 10^{-6}$ mV$^2$/Hz. (Data taken without signal shaping filters.) . 219

D.5 Example $\bar{\mathcal{F}}_{10\mathrm{M}}$ SQUID spectra with magnet on (*blue*) and off (*orange*), along with the digitizer noise floor (*gray*). SQUID spectra are averaged over ≈9 h, digitizer data averaged over ≈16 h. The typical SQUID noise floor is shown in green dashed line. Note: The spectra were collected at different times and some of the transient noise peaks are not seen in all spectra. . . . . 222
D.6 (a) Example calibration peak at 850 kHz with 10 mVpp excitation and 90 dB of attenuation. Bin width is 1 Hz wide and all power is contained within a single bin. Output voltage is measured at the output of the amplifier electronics. (b) Measured detector response for four different input amplitudes taken with the magnet on. The measured gain is a factor of $\approx 6.5$ below the expected response (dashed line at the top). The outlier in the 20 mVpp spectrum is the result of a background line contributing power to the measured peak. . . . . 224
D.7 (a) The number of $3\sigma$ excesses accounting for the look-elsewhere effect in each spectrum after vetoing the excesses that are present in the corresponding Magnet Off data. (b) The distribution of local TS values in the full month of analyzed data after removing periods of transient noise. In blue is the observed distribution of local TS values prior to vetoing the Magnet Off excesses. In green, the observed distribution of TS values after the Magnet Off veto. In red, the expected distribution under the null hypothesis. We see that after applying vetoes, there is excellent agreement down to very low survival counts, with no remaining $5\sigma$ excesses. . . . . 228
D.8 (*Top row*) The recovered signal parameters as a function of the injected signal parameters in four Monte Carlo realizations with identical mean background levels. Green and yellow bands indicate the expected 1 and $2\sigma$ containment for the upper $95\%$ limit on the axion coupling under the hypothesis of no axion signal. (*Bottom row*) The observed and expected test statistic for discovery as a function of the injected signal strength. The dashed red line indicates the threshold for a discovery at $5\sigma$ significance accounting for the LEE, while the dashed black line indicates the upper $95\%$ limit on the observed test statistic under the null hypothesis. . . . . 230

E.1 ABRACADABRA-10 cm Run 3 calibration circuit diagram. A fake axion signal generated in the signal generator is attenuated by 93 dB (including 3dB loss in the combiner) before being coupled into the pickup cylinder analogously to an axion signal. The resulting power excess is readout on the SQUID and measured in the ADC digitizer. In Run 3, calibration is performed with the magnet turned on and the active feedback circuit running. During data taking, the signal generator is replaced with a 50 Ω terminator. The FLL feedback resistor and inductor are omitted for clarity. . . . . 234
E.2 The SQUID flux for Run 3 over the 70 kHz to 2 MHz frequency range at which we collect data. The magnet on noise level (magenta) is elevated compared to data taken with the magnet off (gold) primarily due to vibrating fringe magnetic fields. For comparison, the noise level from a similar SQUID without anything plugged into its input is plotted in teal. . . . . 235
E.3 As in Fig. 6.3, but evaluated on the 10% of unblinded Run 2 data against which we calibrated our analysis procedure. . . . . 240

E.4 (*Left*) The histogrammed data acceptance fraction under the data filtering over all masses analyzed in Run 3 data. (*Right*) The fraction of masses removed by magnet off vetoes as a function of frequency in Run 3 data. The acceptance fraction is determined within 100 log-spaced bins between the minimum and maximum axion masses within our analysis range. Note that while we display the Run 2 results, those were used only to develop our analysis protocols and not in the physics analysis. . . . . 240
E.5 The hyperparameter, $\sigma_A$, converted to the units of $g_{a\gamma\gamma}$, for the systematic nuisance parameter, $g^{\rm nuis}_{a\gamma\gamma}$, as a function of axion mass (labeled Systematic Nuisance Param.). We compare the systematic nuisance hyperparameter to the statistical uncertainties (labeled Hessian Statistical Error), which are computed from the Hessian for the log-likelihood without systematic uncertainties about the best-fit axion coupling, $\hat{g}_{a\gamma\gamma}$. . . 242
E.6 The time evolution of the broad excess that is associated with the putative signal candidate in the Run 2 data that survived all analysis cuts. The excess persists after the magnet is turned off and evolves in frequency, indicative of a background source. The magnet off veto did not anticipate this level of time evolution and so did not remove these excesses. Since this feature was found after unblinding, we report it here but do not consider it to be a credible axion detection. . . . . 244
E.7 A comparison of our fiducial limits that include a nuisance hyperparameter correction (black) and those without any correction (blue). Limits set with the nuisance hyperparameter are slightly weaker, but the features and limit-setting power are broadly similar. The figure is smoothed for clarity. . . . . 245
E.8 (*Top row*) The best fit and 95% upper limit on the recovered signal strength as a function of the injected signal strength at five mass points evaluated on the real Run 3 data. The results are compared to the $1\sigma$ and $2\sigma$ expectations for the $95^{\rm th}$ percentile upper limit under the hypothesis of no axion signal as determined by 2560 Monte Carlo (MC) realizations of the null model fits to the real data at each injected signal strength. (*Bottom row*) In black, the recovered detection test statistic for the signal injected in the real data as a function of injected signal strength. The dashed red line indicates the threshold for a $5\sigma$ detection significance account for the look-elsewhere effect while the green and yellow bands indicate the $1\sigma$ and $2\sigma$ expectations for the detection significance determined from 2560 MC realizations of the null model combined with appropriately varied injected signal strength. . . . . 246

F.1 (*Left*) A comparison of the 95% upper limits of the flux density spectra measured with our windowed analysis for the GBT and Effelsberg observations of the Galactic Center and radiometer expectations. Data are analyzed at an approximately 1.831 MHz frequency resolution corresponding to the fiducial resolution for the GBT analysis. Although the Effelsberg data is consistent with the radiometer expectations at its original resolution, when down-binned to the GBT resolution, it demonstrates similar incompatibility with the radiometer expectations. (*Right*) The 95% upper limits on the signal flux for the indicated sources from the GBT observations. These signal flux limits are compared to the expected flux density limit appropriately computed from the radiometer equation. The analysis is performed at the fiducial analysis bandwidth, see Tab. F.1. . . . . 248

F.2 (*Upper Left*) The raw, uncalibrated Effelsberg data collected in the L-Band at the GC and in the Off Position at frequencies between 1.3-1.45 GHz. Detector features much larger than the expected width of an axion signal and coincident RFI lines in On and Off data are clearly visible. (*Upper Right*) The raw, uncalibrated GBT data collected from the INSs RX J0720 and RX J0806. For visual clarity, we do not show the corresponding Off Position data. Thickets of RFI are especially visible around 1.575 GHz and 1.62 GHz. (*Lower Left*) The flux density limits as a function of frequency obtained from our analysis of the Effelsberg GC data. With the exception of locations of narrow RFI, the limits are flat and characterized by the expected statistical variations from channel to channel. (*Lower Right*) The flux density limits as a function of frequency obtained from our analysis of the GBT INS data. . . . . 250
F.3 The interval-by-interval acceptance for two adjacent frequency channels for data taken from the GC observation by the GBT. Channels I and II are located at 1.61908569 GHz and 1.61899414 GHz, respectively. Data for each channel are shown in black and blue, respectively, with correspondingly colored highlighted regions identifying the reference interval for each channel. The antenna response is shown in arbitrary units. Time intervals accepted in both channels are highlighted in green, with those rejected in both channels highlighted in red. Intervals which are accepted in only one channel are not highlighted. . . . . 253
F.4 The channel-by-channel acceptance fraction for each ON-position measurements of the observation target in each observing session. Acceptances are averaged over the two polarizations and downsampled by a factor of 50 for presentation purposes. In the top row, we display the acceptances in the ON observing position; in the bottom row, the acceptances in the OFF observing position. . . . . 254
F.5 A comparison of the calibrated flux density for the GBT observation of the GC (*blue*) to the Effelsberg observations of the GC in the L-Band and S-Band (*green*). Note that the Effelsberg data is calibrated to follow the black curve, averaged over large frequency scales. The calibrated L-band Effelsberg data is around 20% different than the calibrated GBT data, suggesting that errors from the calibration procedure impacting sensitivity to $g_{a\gamma\gamma}$ are only on the order of 10% and subdominant compared to other sources of uncertainty. . . . . 257
F.6 (*Top Left*) A noise-free example flux density spectra for an axion of mass $m_a = 3.46\pi$ GHz with a coupling strength of $g_{a\gamma\gamma} = 10^{-11}\,\mathrm{GeV}^{-1}$ generated for the GBT observation of the GC at $\delta f_{\rm fid} = 1.831$ MHz. We assume Model I for the NS population (the model with more NSs participating in the conversion process) and take the DM to follow an NFW density profile. (*Top Right*) As in the top left, but using a cored DM density profile with a core radius of 600 pc. (*Bottom Left*) As in the top left, but assuming the conservative Model II for the NS population. (*Bottom Right*) As in the top right, but assuming the conservative Model II for the NS population. . . . . 258

F.7 As in Fig. F.6 but for the Effelsberg observations of the GC in the S-band, for an axion with $m_a = 2\pi \times 2.5$ GHz and $g_{a\gamma\gamma} = 10^{-11}\,\mathrm{GeV}^{-1}$. The panels indicate the assumed DM density profile for the Milky Way (NFW or cored NFW with a core radius of 0.6 kpc) and also the NS population model (Model I or Model II, as described in the text). Note that in this case we search for the brightest converting NS. We have shifted each of the MC realizations around in frequency space such that the brightest converting NS appears at $f = 2.5$ GHz. Note that in the scenario with Model 1 and a cored DM profile, the brightest converting NS is not always that much brighter than the signal flux in the sidebands, from other NSs within the Effelsberg beam, which makes it harder to discover an axion signal in this case. . . . . 260

F.8 (*Above*) The ON-position antenna temperature and raw antenna data for ON- and OFF-position measurements for the M31 observation. A narrow feature appears at the indicated central frequency channel in each of the datasets, although with larger relative magnitude in the antenna temperature. The fact that the feature appears in all datasets suggests it is not an axion signal. (*Below*) The test statistics for the central channel excess as a function of the central channel for the analysis of the ON-position antenna temperature and the raw antenna data for ON- and OFF-position measurements for the M31 observation. At the location of the narrow feature, the test statistic is quite large for all analyzed datasets, and the excess in the antenna temperature is vetoed as the test statistic in the OFF-position data exceeds the veto threshold. . . . . 264

F.9 (*Left*) The one-sided 95% upper limit on the axion-photon coupling as a function of the injected signal strength. The limit lies above the injected signal strength, indicating that we are not excluding an axion signal when present. (*Right*) The test statistic (TS) for discovery as a function of the injected signal strength. For sufficiently large signal strengths the TS exceeds our $\mathrm{TS} = 100$ threshold for an axion signal to be discovered. 266

F.10 Maser lines as detected in the GBT data. For each maser line, we show the antenna temperature (black) and the raw OFF data (blue), with each independently rescaled so as to fit within the same figure. The expected frequency location and width of the maser line, which is set by the line-of-sight velocity of W3OH, is indicated by the light red band. The width of the central frequency channel in which the maser line is detected is indicated by the light blue band. We additionally provide the TS associated with the maser line detection in the antenna temperature and the maximum percentile of the variable-width OFF position TS for each line. None of the detections are vetoed as none the maximum OFF position TS percentiles exceed the $97.5^{\mathrm{th}}$ percentile value that triggers vetoing. . . . . 267

F.11 (*Left*) The discovery TS survival function for all of the observations considered in this Letter. Note that the survival function is defined as the fraction of TSs observed at or above the indicated value. This figure excludes frequencies that are vetoed from the OFF position observation analyses. The "MC Expected" curve shows the expectation under the null hypothesis, as determined by MC simulations. We note that all observations are from GBT except those labeled "Eff", which are from the Effelsberg telescope. (*Right*) As in the left panel, but including frequencies that would be vetoed by the OFF data. Without the OFF vetoes there would be a significant number of frequencies with TSs exceeding the TS detection threshold, which emphasizes the importance of the OFF position vetoing procedure. . . . . 269

F.12 (*Top Left*) The Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 2.51 GHz in the S-band observation of the GC. Frequencies are plotted relative to the frequency corresponding to the excess channel frequency. This excess has $\mathrm{TS} \approx 41$. While this excess is not vetoed by the OFF data analysis, the OFF data does should a feature at the central frequency. (*Top Right*) Similarly, the Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 2.69 GHz in the S-band observation of the GC. This excess is also not vetoed, but like the previous excess there does appear to be a corresponding feature in the OFF data. (*Bottom Left*) The Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 1.34 GHz in the L-band observation of the GC. It also appears that there is a similar, though not so significant, feature in the OFF data. (*Bottom Right*) The GBT data shown in the analysis window around the excess candidate located at a central frequency of approximately 1.59 GHz in the observation of RX J0720.4−3125. As before, frequencies are plotted relative to the frequency corresponding to the excess channel frequency. The excess only appears at high significance in the antenna temperature; similar coincident features are observed in both ON and OFF data, coincident features appear in the raw ON and OFF data, although not at high enough significance in the OFF data to result in a veto of the excess. As before, this excess does not exceed our detection threshold, although it does come closer, with $\mathrm{TS} \approx 90$. . . . . 271

F.13 A comparison of survival functions for various polynomial background models for the analysis of Effelsberg S-Band data. The flat background model is unable to accurately model the null hypothesis and a significant improvement in the quality of the fits is seen by going to the linear background model. On the other hand, there is little improvement to the quality of the fits when going from the linear to quadratic background models, except at very high TS values. Note that we use the quadratic background model in our fiducial analyses. Cubic background models produce similar results but are most computationally intensive to implement. . . . . 272

F.14 The discovery TS survival function for the INSs analyzed with and without time-series data filtering applied. Applying the time-series filtering eliminates a number of high significance excesses that appear due to transient noise that appears in the data. . . . . 273

F.15 Limits on the axion-photon coupling for different combinations of assumptions about the DM density profiles in the observed galaxies and the properties of the NSs within those galaxies (see [104] and text for details). . . . . 274

F.16 Comparison between the profile likelihood and percentile upper limits methods for M54 observations with GBT. The black line (green area) shows the 95% C.L upper limits ($1\sigma$ containment band) obtained with the percentile method. The red line shows the upper limits obtained with the profile likelihood method and calibration used as default in our main pipeline; these flux limit curves are used in the main text. . . . . 276

F.17 Same as Fig. F.16, for Effelsberg GC observations. . . . . 276

F.18 An illustration of how outgoing electromagnetic waves are refracted towards the normal vector to the conversion surface, labeled here by $\hat{\mathbf{z}}$. . . . . 277

G.1 (Left) The HR diagram for the Quintuplet template star of mass 85 $M_\odot$ and initial surface rotation of 300 km/s. The coloring indicates the year before the run was stopped, approximately a few years from supernova. We mark with black squares, in order of occurrence, when the star enters the WNh phase, when it is 3 Myr old, when its core undergoes helium ignition, when it enters the WN, WC, and WO phases, and finally when the run ends at 3.85 Myr. (Right) A $\log$T-$\log \rho$ diagram for the template star with the same points of interest marked. We also show the relevant degeneracy zones, showing that the star is entirely in the nonrelativistic nondegenerate regime. . . 282
G.2 (Left) The abundances of hydrogen (black), helium (red), carbon (yellow), and oxygen (green) in the center of the star as a function of time, for the simulation described in Fig. G.1. With dashed-black vertical lines, we mark several points of interest: "WNh" indicates the time the star enters the WNh phase, "He ignition" when its core undergoes helium ignition, and "WN","WC", and "WO" indicate the beginning of the WN, WC, and WO phases, respectively. (Right) The same as in the left panel, but for surface abundances. . . . . 283
G.3 (Left) The stellar core temperature as a function of time for the simulation described in Fig. G.1. (Right) The hydrogen and helium luminosities in the core through the CNO cycle and the triple-alpha process, respectively. The dashed-black vertical lines retain their meanings from Fig. G.2. . . . . 284
G.4 The stellar mass (black) and radius (red) as a function of time from the simulation described in Fig. G.1. The dashed-black vertical lines retain their meanings from Fig. G.2. . . . . 284
G.5 (Left) Axion volume emissivity over the interior of the star. In this figure we have taken the stellar model to be the one at the start of the WC stage and fixed $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$. For comparison purposes, we also show the temperature profile. (Right) Axion luminosity spectrum for those same stages marked in Fig. G.2. . . . . 285
G.6 We denote the projections of the Galactic magnetic field onto the plane normal to the propagation direction by $B_1$, $B_2$. (Left) The transverse magnetic field components in our fiducial model (the JF12 model, black) and alternate model (PTKN11, orange) towards the Quintuplet and Arches clusters. Note that in our fiducial $B$-field model we extend the JF12 model to distances less than 1 kpc from the GC using the field values at 1 kpc. The true magnetic field values in the inner kpc almost certainly surpass those from this conservative model (see text for details). (Right) The two field components towards the Wd1 cluster, which is taken to be at a distance of 2.6 kpc from the Sun. The conversion probabilities towards Wd1 are much larger in the alternate model (PTKN11) than in our fiducial model (JF12), though we stress that random fields are not included and could play an important role in the conversion probabilities towards Wd1. . . . . 286
G.7 (Left) The free electron density $n_e$ towards the GC in our fiducial model (YMW16) and the alternate model (`ne2001`). (Right) As in the left panel but towards the Wd1 cluster. The free-electron density gives the photon an effective mass and thus affects the axion-photon conversion probability. . . . . 288

G.8 (Left Column) The axion-photon conversion probabilities $p_{a\to\gamma}$, assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$, computed as a function of the axion energy $E$ (and assuming $m_a \ll 10^{-10}$ eV) using the formula given in (G.5). (Top Left) The conversion probabilities for axions produced in the Quintuplet or Arches clusters for different modeling assumptions for the Galactic magnetic field and free-electron density. Our fiducial result is shown in solid black. Note that the plasma mass, induced by the free-electron density, becomes more important at lower axion energies and induces the lower-energy features. The dashed black curve shows the effect of changing from the YMW16 free-electron model to the `ne2001` model. Removing the $B$-field within the inner kpc leads to the results in red, while only modeling a 50 $\mu$G field in the inner 400 pc leads to the results in blue. Changing to the PTKN11 model (and masking the inner kpc) gives the results in orange. We estimate that if the axions traverse the GC radio arc, located near the Quintuplet and Arches clusters, the conversion probabilities could be enhanced to the values shown in grey. (Bottom Left) As in the top left panel but for axions emitted from the Wd1 cluster. (Right Column) The effects of the different conversion probability models on the 95% upper limits on $g_{a\gamma\gamma}$ for Quintuplet (top right) and Wd1 (bottom right). Note that Arches is similar to Quintuplet, since they are both assumed to have the same conversion-probability models. . . . . 297

G.9 (Left) As in Fig. 8.1, but for the total observed counts between 10 - 80 keV instead of the background-subtracted counts. (Center) The best-fit background model, summed from 10 - 80 keV, for the Quintuplet data set shown in the left panel. (Right) The predicted axion-induced signal template from Quintuplet, in counts, normalized for an axion with $g_{a\gamma\gamma} = 7 \times 10^{-12}$ GeV$^{-1}$ and $m_a \ll 10^{-11}$ eV. . . . . 298

G.10 (Upper Left) The Quintuplet axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 10 - 80 keV, is shaded in red. (Upper Right) The individual contributions of each stellar classification to the Quintuplet axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification (per star) in Quintuplet. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band. . . . . 298

G.11 (Left) We inject a synthetic axion signal into the Quintuplet NuSTAR data with axion coupling $g^{\rm inj}_{a\gamma\gamma}$. We then pass the hybrid synthetic plus real data through our analysis pipeline and show the best-fit coupling $g^{\rm rec}_{a\gamma\gamma}$, along with the recovered $1\sigma$ and $2\sigma$ uncertainties. (Middle) The discovery TS for the axion signal for the test illustrated in the left panel. The square root of the TS is approximately the discovery significance. (Right) The 95% upper limit recovered for the injected signal test. Importantly, the 95% upper limit is above the injected signal value, for all injected signal strengths, and the upper limit is consistent with the 68% and 95% expectations for the upper limit under the null hypothesis, which are indicated in green and gold, respectively. . 299

G.12 As in Fig. 8.2, except for different ROI sizes, as indicated. . . . . 299

G.13 As in Fig. G.9, but for the Wd1 cluster NuSTAR analysis. The red star indicates the location of the magnetar CXOU J164710.2–45521, which is masked at 0.5'. Also shown is the background-subtracted count data, as in Fig. 8.1. . . . . 300

G.14 (Upper Left) The Wd1 axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 15 - 80 keV, is shaded in gray. (Upper Right) The individual contributions of each stellar classification to the Wd1 axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification in Wd1. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band. . . . . 301
G.15 As in Fig. G.12 but for the Wd1 analysis. Note that we only include energies above 15 keV in our analysis because of ghost-ray contamination. . . . . 302
G.16 (Upper Left) The Arches axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 20 - 80 keV, is shaded in gray. (Upper Right) The individual contributions of each stellar classification to the Arches axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification in Arches. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band. . . . . 302
G.17 (Top Panel) As in Fig. G.9, but for the Arches cluster. (Bottom left) We show the best-fit emission associated with the halo template that describes emission from the nearby molecular cloud. (Bottom right) As in in Fig. 8.1, but for Arches. . . . . 303
G.18 (Left) The Arches spectrum measured with and without the halo template. Note that we use the spectrum with the halo template in our fiducial analysis, though the difference between the two results is relatively minor above $\sim$20 keV. (Right) As in Fig. G.12 but for the Quintuplet analysis. Note that these spectra are computed while profiling over halo emission. Above $\sim$20 keV the different ROIs produce consistent results. . . . . 303
G.19 As in Fig. 8.3 but from the analysis towards the Arches SSC. No evidence for axions is found from this search. . . . . 304
G.20 (Left) The evolution of the nitrogen abundance Z(N) over time from MESA simulations of a non-rotating 85 $M_\odot$ star with initial metallicity $Z = 0.01$ to $Z = 0.04$. The bolded sections of the lines correspond to the WNh phase. The gray shaded region indicates the measurements of nitrogren abundances of the Arches WNh stars from [329]. . . . . 304
G.21 (Left) The variation to the 95% upper limit found by varying the initial metallicity and rotation in the range $Z \in (0.018, 0.035)$ and $\mu_{\rm rot} \in (50, 150)$ km/s for the Quintuplet analysis. The blue region indicates the maximum and minimum limit found, while the black curve shows our fiducial limit. (Right) As in the left panel but for Wd1. . . . . 305

H.1 Examples of the signal region spectra for MOS (top panels) and PN (bottom panels) in Ring 1 (left panels) and Ring 8 (right panels) with and without background subtraction in red and black, respectively. The background-region spectra are shown in grey. Many of the large instrumental features that are removed when looking at the background-subtracted data. Note that for visual clarity these spectra have been down-binned by a factor of 4. . . . . 311

H.2 The same background-subtracted data sets illustrated in Fig. H.1 (also down-binned), but now shown along with their best-fits under the null hypothesis. The best-fit model prediction is shown in black, which may be decomposed into the contribution from the GP model (dark red) and the contributions from the individual background lines (colored curves). Note that the background lines to include in the analysis are determined independently in each annulus, as described in the text. . . . . 314
H.3 The spurious-signal hyperparameter $\sigma^2_{\rm spur,m}$ (labeled MOS Sys. and PN Sys.), as computed in (H.9), as a function of the DM mass. For both MOS and PN the nuisance parameter $A_{\rm spur}$ is predominantly active at low energies, and it plays a more significant role in PN than in MOS. We compare the hyperparameter to the statistical uncertainties (labeled MOS Stat. and PN Stat.), which are computed from the Hessian of the log-likelihood (without the spurious-signal) about the best-fit mixing angle at a fixed energy. We note that several of the sharp variations of the expected sensitivity shown in Fig. 9.3 arise as a result of the variations of the spurious signal hyperparameter shown here. . . . . 317
H.4 (Left) The survival function of the test statistic for discovery in the analysis of the MOS data. Under the null hypothesis, and for a large number of samples, the survival fractions are expected to follow the $\chi^2$ distribution, as verified by MC (as labeled). At a finite number of samples the expected chi-square distributions are found from MC to be expected to be contained within the green and gold shaded regions at 68% and 95% confidence, respectively. The negligible effect of the systematic nuisance parameter can be seen by comparing the survival function without the nuisance parameter (red, labelled "Data") and with the nuisance parameter (blue, labeled "Data w/ Nuisance Parameter"). (Center) As in the left panel, but for the PN analysis. (Right) The survival function for the joint analysis of MOS and PN data. In blue, the survival function for the joined PN and MOS analysis without systematic nuisance parameters; in red, the survival function for the joint analysis when the PN and MOS results are corrected by their independently-tuned systematic nuisance parameters prior to joining. . . . . 318
H.5 As in Fig. 9.3, but for the MOS (left panel) and PN (right panel) analyses individually and with and without the spurious-signal nuisance parameter. The $1\sigma$ and $2\sigma$ expectations are shown only for the case with the spurious-signal nuisance parameter. The limits without the nuisance parameter are slightly stronger at low masses. The sharp variations in the expected sensitivity, especially visible in the PN results, arise from how the spurious-signal hyperparameter is determined through the sliding window procedure. . . . . 319
H.6 A comparison of all results obtained in the joint analysis of PN data with and without the inclusion of Ring 3, which may be subject to systematic mismodeling. Note that for this comparison we do not profile over the spurious-signal nuisance parameter. . . 319
H.7 As in Fig. 9.3, but interpreted as limits on the DM lifetime. This figure applies for DM whose decays produce a single mono-energetic photon at energy $m_\chi/2$. If the DM decay produces two photons (as in an axion model), then the lifetime limits are twice as strong. . . . . 320

H.8 (Upper Left) The best-fit signal flux, and 1 and $2\sigma$ uncertainties, as a function of the central UXL energy across our full energy range for the innermost MOS ring. (Lower Left) The corresponding significance in favor of the signal model, multiplied by the sign of the best fit UXL normalization at that energy, along with the 1/$2\sigma$ expectations under the null hypothesis. (Right Panel) As in the left panel but for the innermost PN annulus. . . . . 321
H.9 As in Fig. H.8 but for annulus 2. . . . . 322
H.10 As in Fig. H.8 but for annulus 3. . . . . 323
H.11 As in Fig. H.8 but for annulus 4. . . . . 324
H.12 As in Fig. H.8 but for annulus 5. . . . . 325
H.13 As in Fig. H.8 but for annulus 6. . . . . 326
H.14 As in Fig. H.8 but for annulus 7. . . . . 327
H.15 As in Fig. H.8 but for annulus 8. . . . . 328
H.16 As in Fig. H.4 but for the individual MOS annuli. Note that the systematic nuisance parameter has not been applied since that is only incorporated in the joint likelihood that combines the results from the individual annuli. . . . . 329
H.17 As in Fig. H.16 but for the PN data sets. . . . . 330
H.18 The results of the analysis of the hybrid data that consists of the real MOS and PN data plus a synthetic DM signal. The DM signal is generated with mass $m_\chi = 7.0$ keV and mixing angle $\sin^2(2\theta) = 2.5 \times 10^{-11}$ as described in the text. The top, middle, and third rows are analogous to Figs. H.4 and H.5, but for the hybrid data set. The last row shows the 1, 2, and 3 $\sigma$ recovered parameter space for the signal in the mass and mixing angle plane. The best-fit recovered signal is indicated in dark blue, while the red star denotes the true value injected. The synthetic signal is appropriately recovered, adding confidence that our analysis procedure has the ability to detect real DM signals if present in the data. . . . . 332
H.19 (*Top Row*) In red, the median $95^{\rm th}$ percentile upper limit on the recovered signal as a function of the injection signal strength at two neutrino masses evaluated on synthetic data. We additionally indicate the 1 and $2\sigma$ containment intervals for the ensemble of upper limits realized at each injected signal strength. Note that these upper limits are not power constrained. These results demonstrate that our analysis framework places robust upper limits that do not rule out an injected signal. (*Bottom Row*) In black, the median recovered detection test statistic for a signal injected in the synthetic data as a function of the injected signal strength, with the 1 and $2\sigma$ containment intervals also indicated. Under the null hypothesis, the detection test statistic should follow a $\chi^2$-distribution; the median and $1\sigma$ and $2\sigma$ percentile values of the $\chi^2$-distribution are indicated by dashed grey lines. These results demonstrate that our detection test statistic follows its theoretically expected distribution under the null hypothesis ($\sin^2(2\theta_{\rm inj}) = 0$) and that our analysis framework can robustly identify a signal which is present in the data. The results are smoothed with a Savitzky–Golay filter for clarity. . . . . 333

H.20 As in Fig. 9.3, but for three different DM density profiles, all based upon Ref. [416]. In solid curve we show our fiducial results, corresponding to the uncontracted NFW profile with a conservative density. The dashed curve then shows our results using the best fit NFW profile, whereas in dashed we show the stronger limits that would be obtained with a contracted DM distribution. Details of the distributions are provided in the text. . . . 334
H.21 The analogues of Figs. H.4 and H.5, but changing the kernel correlation length to $\sigma_E = 0.2$ (c.f. our fiducial value of $\sigma_E = 0.3$). The differences between the $\sigma_E = 0.2$ and $0.3$ results are minor. . . . 335
H.22 As in Fig. H.21 but with $\sigma_E = 0.4$. The limit is slightly strengthened, although again the differences are not significant. . . . 336
H.23 As in Fig. H.21 but with $\sigma_E$ treated as a profiled nuisance parameter. The results demonstrate that even providing our background model this additional freedom does not have a significant impact on the limit. . . . 337
H.24 As in Fig. H.21 but with the alternate GP kernel, in (H.10), with $\sigma^2 = 0.5$ keV$^2$. . . . 338
H.25 As in Fig. H.24 but with $\sigma^2 = 1.0$ keV$^2$. Adopting a large scale length again slightly strengthens the limits, although again the systematic variation of our results with the kernel is relatively small. . . . 339
H.26 A comparison of the limits obtained across the full mass range for each variation of the GP correlation-length hyperparameter considered. In particular we show results for variations of the relative-scale and fixed-scale kernels (denoted $\sigma_E$ and $\sigma^2$ respectively), as well as the relative-scale kernel where the scale profiled independently in each annulus. . . . 339
H.27 As in Fig. H.21, but with the fiducial GP kernel at $\sigma_E = 0.3$ and the inclusion of 3.32 and 3.68 keV lines in all analyzed annuli. The newly masked region associated with these two lines is highlighted in light red. . . . 340
H.28 A close inspection of the limits set in our fiducial analysis and the modified analysis that includes a 3.32 and 3.68 keV line in each annulus. We compare the limits set in these two analyses both with (solid lines) and without (dashed lines) the inclusion of our systematic nuisance parameter designed to test for and correct possible mismodeling. 341
H.29 The same results as presented in Figs. H.4 and H.5, however on a modified data set where instead of analyzing the signal ROI divided into eight individual rings, we stack the inner three rings into a single annulus. As in our primary approach, we subtract the background ROI flux from the signal-region data. The results are comparable to, although slightly weaker than, those from our fiducial approach, consistent with the reduced information available. . . . 342
H.30 As in Fig. H.29, however considering the stacked signal ROI without subtracting the background. The limit is noticeably worse, and several excesses appear, highlighting the importance of the background subtraction. . . . 343

H.31 Here we compare our fiducial results using a GP model, shown in black, to the result of an approach where the continuum background contribution is modeled with a second order polynomial, shown in red, as described in the text. Both results are shown without imposing a systematic nuisance parameter. While our fiducial approach uses the background-subtracted signal-ROI data, the alternate approach uses the un-subtracted data. We see that in both cases the expected and resulting limits are in qualitative agreement, demonstrating that our choice of GP modeling in our fiducial analysis does not drive the sensitivity of our results. . . . . 344

# LIST OF TABLES

TABLE


7.1 The targets observed by the GBT and Effelsberg for evidence of axion DM. "Pop." refers to populations of NSs, while "INS" refers to a single isolated NS. The bin widths $\delta f_{\rm obs}$ correspond to those of the original observation, but we down-bin the data before performing the axion line search to the resolution given by $\delta f_{\rm fid}$ to account for the finite width of the signal. The INS (GC) observations were performed with the GBT (Effelsberg radio telescope). The GBT INS observations cover the frequency range 1.15 to 1.73 GHz, with a gap from 1.2 to 1.35 GHz, and the L-band (S-band) Effelsberg observation covers 1.27 to 1.45 GHz (2.4 to 2.7 GHz). Note that the $t_{\rm exp}$ are the ON exposure times. . . . . 135

D.1 Summary of the ABRACADABRA-10 cm detector design parameters. . . . . 216

F.1 As in Tab. 7.1 but for the GBT observations of the GC, M31, and M54. . . . . 249

G.1 The number of stars $N_{\rm star}$ for each stellar class in the Quintuplet cluster, along with the predicted axion luminosities (all in erg/s). Note that Quintuplet is $\sim$30 pc away from the GC. Except in the last column, the axion luminosities are summed over all energies. All entries assume $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ and are summed over all stars for the given stellar class. . . . . 291

G.2 As in Tab. G.1 but for Wd1. . . . . 293

G.3 As in Tab. G.1 but for Arches. . . . . 294

H.1 The list of spectral lines of instrumental and astrophysical origins which are included in our background model for the MOS camera. For the line in each ring, we provide the value of $\Delta\chi^2$ associated with the addition/removal of the line from the best-fit background model which is obtained after our line-dropping procedure. Bolded values indicate the inclusion of a line in a ring's background model. . . . . 312

H.2 The list of spectral lines of instrumental and astrophysical origins which are included in our background model for the PN camera. For the line in each ring, we provide the value of $\Delta\chi^2$ associated with the addition/removal of the line from the best-fit background model which is obtained after our line-dropping procedure. Bolded values indicate the inclusion of a line in a ring's background model. . . . . 313

H.3 The goodness-of-fit of the null model fit in each annulus for PN and MOS data sets as measured by the $\chi^2$ divided by the number of degrees of freedom (dof). With the exception of Ring 3 of the PN data set, this measure indicates an acceptable goodness-of-fit to the data under the null, as quantified through the $p$-value (see text for details). 320

H.4 The best fit normalization of the GP kernel for each ring in PN and MOS. We present $\sqrt{A_{\rm GP}}$ in units of photons/cm$^2$/s/keV for $A_{\rm GP}$ defined in (H.3). . . . . 320
H.5 The best-fit scale $\sigma_E$, determined under the null model, when this scale is treated as profiled nuisance parameter. In all cases except Ring 6 of PN data, the best-fit scale is larger than the scale of the kernel used in our fiducial analysis, indicating that our fiducial choice of $\sigma_E$ = 0.3 was conservative and endowed the GP model with sufficient flexibility. . . . . 326

# LIST OF APPENDICES

**A Simulations of Axion Production in the Post-Inflationary Misalignment Scenario . . . 162**

**B The Statistics of Axion Direct Detection . . . 189**

**C The Statistics of Axion Direct Detection with Multiple Detectors . . . 202**

**D First Results from ABRACADABRA-10 cm . . . 212**

**E Second Results from ABRACADABRA-10 cm . . . 232**

**F Radio Searches for Axion Dark Matter . . . 247**

**G X-Ray Searches for Axions from Super Star Clusters . . . 280**

**H X-Ray Searches for Decaying Dark Matter in the Milky Way . . . 306**

# ABSTRACT

The majority of the matter in the Universe is non-luminous and unaccounted for by any known particle, making the unknown nature of dark matter one of the most urgent problems in fundamental physics. Amidst a broad landscape of particles proposed to explain the dark matter, axions have emerged as a particularly well-motivated candidate as they naturally arise in extensions of the Standard Model and can simultaneously reproduce the observed dark matter abundance while solving other outstanding mysteries in particle physics. Despite this, axions have remained largely unprobed, and new insights and innovative approaches are required to carefully test the axion dark matter hypothesis. This thesis aims to advance prospects for axion detection by identifying how axion signals may appear, developing optimized searches for these signals, and implementing robust analysis strategies. I will begin by showing how simulations of axion production in the early universe can direct search efforts toward the best-motivated mass range for axions that solve the Strong *CP* Problem related to the absence of a neutron electric dipole moment in quantum chromodynamics. I will then discuss the development of rigorous analysis frameworks for axion direct detection and their application to the search for axion dark matter with the ABRACADABRA detector. Lastly, I will show how astrophysical observations with *X*-ray and radio telescopes can be used in novel searches for axion dark matter. This thesis contributes to an increasingly comprehensive search program that will either discover or exclude axion dark matter in the coming years.

# CHAPTER 1

# Introduction

The unknown nature of dark matter (DM) is one of the most significant unsolved problems in physics. An overwhelming accumulation of evidence in the form of observations of gravitational interactions on astrophysical and cosmological scales suggests that a form of cold, nonluminous matter exists in nearly five times greater abundance than the so-called ordinary matter described by the well-established Standard Model (SM) of particle physics. Gravitational interactions of DM are unlikely to provide a complete characterization of this unidentified form of matter since all forms of matter interact gravitationally. Instead, the discovery of the microphysical nature of DM will require identifying its presently unknown interactions through the development of well-motivated hypotheses with detailed phenomenologies, the design of precise experimental and observational probes, and the execution of sensitive analyses that can untangle the signals of new physics from enormous datasets in the presence of potentially confounding backgrounds.

Despite the associated challenges, work toward identifying DM represents arguably the best prospect to advance understanding of particle physics. Characterizing DM and developing a more fundamental theory that describes ordinary matter, DM, and the interactions within and between the two sectors could lead to insights on outstanding theoretical questions. Moreover, as the dominant fraction of matter in the universe, DM plays a central role in cosmological and astrophysical histories by seeding large-scale structure and driving the dynamics of galaxies and galaxy clusters. Therefore, the study of DM is highly compelling as an understanding of its precise details is intrinsic to understanding the evolution of the universe.

This thesis will attempt to ask and answer the question of how we can work toward the detection of arguably the best-motivated DM candidate, the axion. This chapter will serve as an introduction to the key concepts that underlie these efforts and is structured as follows. In Sec. 1.1, we will consider the strong evidence for the existence of DM and its nature as an unidentified particle and provide a general discussion on particle dark matter candidates. In Sec. 1.2, we will discuss in detail theoretical aspects of the Strong *CP* Problem and its solution through the introduction of an axion. In Sec. 1.3, I will review the two strategies I have pursued in the search for DM: direct detection with precision laboratory experiments and indirect detection through observation

of astrophysical systems. Finally, in Sec. 1.4, I will describe the organization of the remainder of this thesis.

## 1.1 The Particle Dark Matter Paradigm

In this section, I will discuss several measurements that provide strong evidence for the existence of DM, largely inspired by the excellent historical review provided in [1]. This discussion will be far from comprehensive and instead is intended to illustrate the concordance of the many complementary probes that suggest DM is the dominant form of matter in the universe. I will then discuss the general aspects of DM model building an the choices made in performing model-dependent searches for DM candidates.

### 1.1.1 Evidence for Particle Dark Matter

The most straightforward evidence for DM comes from measuring the circular velocity of stars in galaxies. The key concept is that the velocity of a star in a bound orbit of a galaxy must be related to the gravitational force it experiences determined by the mass of the galaxy interior to the orbit. If a star moves too fast, it will escape the galaxy; if it moves too slow, it will fall into the galaxy's potential well. In Newtonian theory, the requisite circular velocity $v_c$ for a star located a distance $r$ from the center of the galaxy is given by

$$v_c(r) = \sqrt{\frac{GM(r)}{r}}, \tag{1.1}$$

where $G$ is the universal gravitational constant and $M(r)$ is the mass of the galaxy contained within radius $r$. By measuring the circular velocities of many stars at many different radii, one can piece together the continuous mass distribution of a galaxy. An example of measured rotation curves with comparison to theory expectations for the Milky Way galaxy is provided in Fig. 1.1. Similar features are observed across a vast catalog of galaxies. Interpreted in the absence of DM, the discrepancy between theory and observation is striking. At radii beyond the galactic disk, beyond which there is little ordinary matter, circular velocities would be expected to scale like $v_c(r) \propto r^{-1/2}$. However, the observed circular velocities are roughly constant out to much larger radii. This suggests the presence of "dark" matter that results in an enclosed mass scaling like $M(r) \propto r$ to radii much larger than that of the galactic disk [2].

However, while certainly suggestive, galaxy rotation curves are only gravitational anomalies and do not uniquely prescribe the presence of unaccounted-for matter; galaxy rotation curves might instead be evidence for undiscovered aspects of gravitational interactions. There is his-

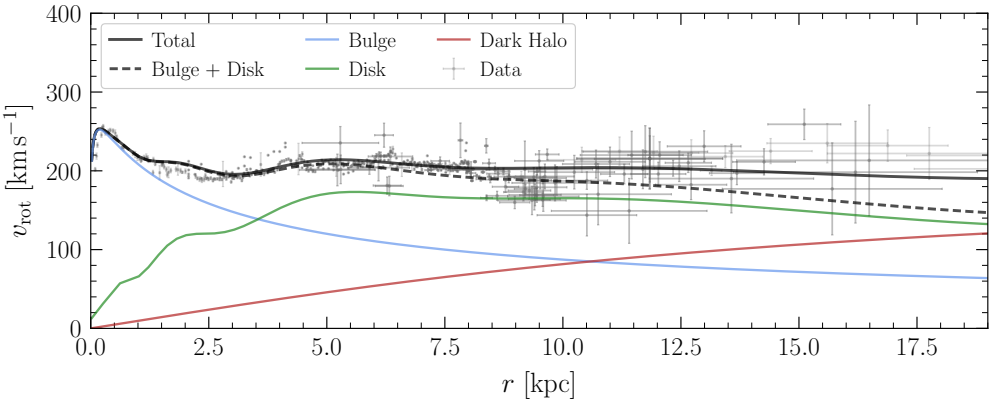



Figure 1.1: A comparison in the Milky Way of measured circular velocities as a function of radius (grey) with expected theoretical contributions from various mass components within the galaxy. The galactic bulge (blue) and the galactic disk (green) are baryonic matter distributions that dominate within the inner galaxy. The expected rotation curves associated with a Milky Way that consisted only of the observed baryonic matter are shown in dashed black and are visibly incompatible with the data at large radii. The contribution of the DM to the circular velocity (red) becomes appreciable at large radii. The rotational velocities expected for the combination of galactic bulge, galactic disk, and DM halo are shown in solid black, demonstrating good agreement with the observed data. Adapted with permission from [3] using data from [4].

torical precedent for both possibilities, as is noted in [5]. In the mid-1800s, measured deviations of Uranus from its expected orbit did lead to the discovery of missing matter (the planet Neptune). Later, in the early 1900s, the precession of the perihelion of Mercury, which was anomalous as interpreted in Newtonian gravity, lead to the discovery of General Relativity. Similarly, an alternative hypothesis to DM could be made for galaxy rotation curves. They could instead be explained by a class of theories known as the MOdified Newtonian Dynamics (MOND) [6–8]. We now consider several additional probes that provide strong evidence for DM and arguably lesser consistency with the MOND hypothesis.[1] While we will not explicitly discuss the tensions of MOND with these observations, reviews can be found in several references [5, 10–12].

Unaccounted-for matter can also be inferred on super-galactic scales from the Bullet Cluster, a pair of merging galaxy clusters. X-ray-emitting gas in the Bullet Cluster can be used to trace its baryonic matter distribution, while its total matter distribution can be measured from the gravitational lensing of photons propagating through the potential of the massive objects. A depiction of these inferred mass distributions is provided in Fig. 1.2. Strikingly, the baryonic matter distribution does not coincide with the total matter distribution, suggesting that not only does the Bullet Cluster primarily consist of non-baryonic matter, but also that that non-baryonic matter must have

[1] We also note with that the advent of large-scale astrometric surveys like *GAIA* and its planned successors, MOND predictions may be sensitively probed on the galactic distances scales at which it is engineered to address the rotation curve problem [9].

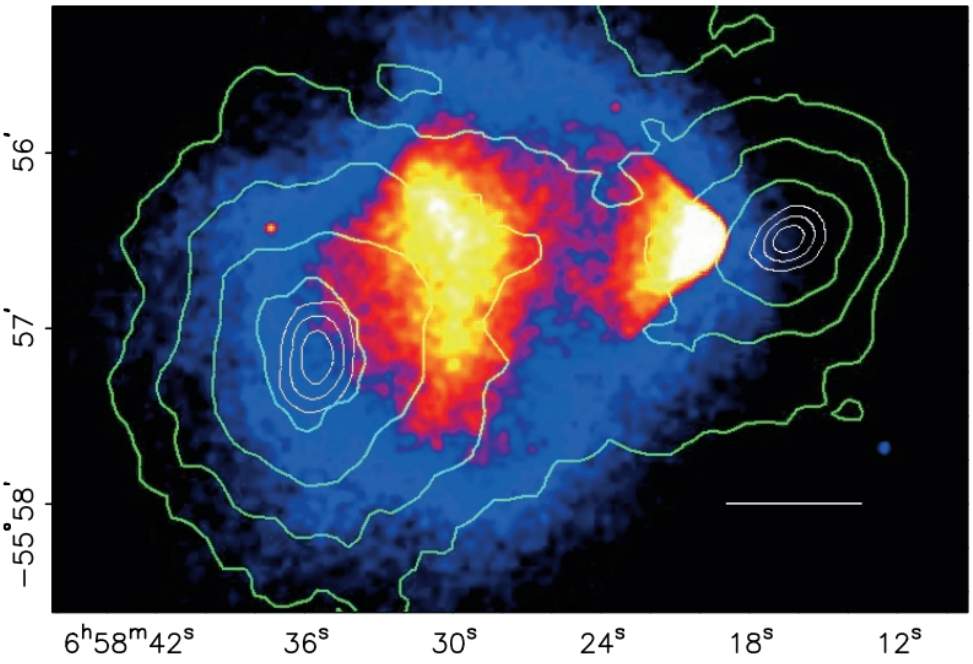



Figure 1.2: A depiction of inferred mass distributions within the Bullet Cluster. Bright blobs indicate the location of X-ray-luminous baryonic matter, while the contours indicate the central location of the inferred mass distributions of the two merging clusters. See [13] for further details.

experienced qualitatively different dynamics during the collision event [13]. An identical conclusion is reached by a similar analysis of the galaxy cluster merger MACSJ0025.4-1222 [14]. From these observations, we infer that while the baryonic matter was slowed by friction induced by its particle physics interactions during the collision, the DM was essentially noninteracting and went unperturbed on its gravitational trajectory.

Finally, we consider cosmological probes of DM enabled by ΛCDM, a phenomenological model describing a universe with an energy budget allowing for dark energy (Λ) and cold, noninteracting DM (CDM) in addition to the baryonic matter.[2] ΛCDM has achieved remarkable success in describing the CMB, large- and small-scale structure, the late-time accelerated expansion of the universe, and relic elemental abundances [17, 18]. For simplicity, we will limit ourselves to a discussion of the CMB, where the relevant observable is the angular power spectrum of temperature fluctuations on an otherwise uniform background. This angular power spectrum, as measured by *Planck*, is presented in Fig. 1.3 alongside theory predictions in ΛCDM cosmologies. The power spectrum contains three notable peaks which describe the angular scales at which the temperature inhomogeneities appear on the sky, with a variety of physical processes are imprinted within

[2]The description of DM as cold is in contrast with ordinary matter. Although the DM may initially be characterized by relativistic velocities, observational constraints require that it become cold at times when ordinary matter remains hot.

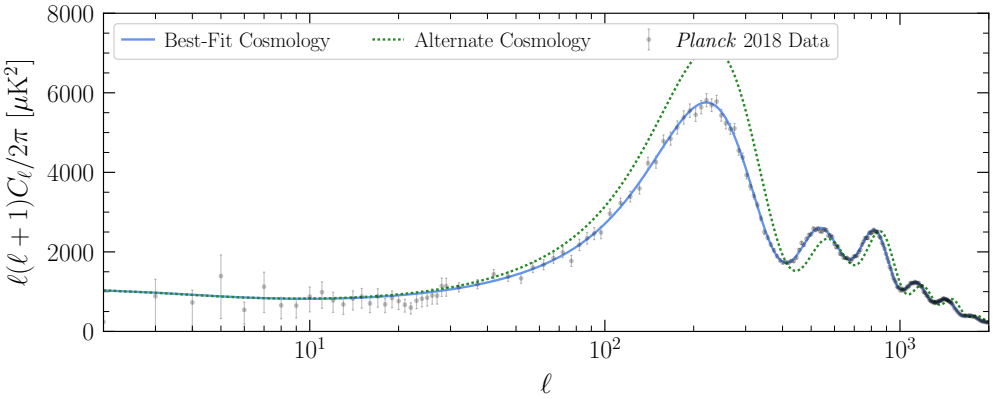



Figure 1.3: In grey, the *Planck TT* spectrum, which measures the angular scales at which temperature inhomogeneities in CMB appear. In blue, we depict the predictions of the best-fit ΛCDM cosmology obtained in the *Planck* 2018 analysis. This cosmology describes a universe in which 84% of matter is comprised of DM, and its predictions demonstrate excellent agreement with the data [15]. If we hold the total matter abundance fixed but reduce the amount of DM by 10%, then we would expect the spectrum shown in dashed green, which is visibly a poor description of the data. Here, the predicted CMB spectra have been generated with the CLASS code package [16].

the location and relative height of each peak. A more detailed review of CMB phenomenology is provided by [19], but in highly simplified terms, the first peak informs us of the total energy density content of the universe, the second tells us about the total baryonic content of the universe, and the third is sensitive to the relative abundance of DM compared to baryonic matter. The most recent CMB analysis from the *Planck* satellite provides measurements of a ΛCDM cosmology in which only 16% of the matter in the universe is baryonic, with the remaining 84% accounted for by DM [15].[3]

We conclude by again emphasizing that this discussion of the evidence for DM is far from a complete one. Any attempt to construct such a list is doomed to failure, but we attempt to provide a slightly more comprehensive catalog by noting that, among others, weak lensing [20], strong lensing [21, 22], large-scale structure [23, 24], small-scale structure [25–27], and local measurements [28] all strongly support a consistent interpretation of DM that is abundant in the Milky Way and throughout the universe.

[3] In fact, ordinary matter and DM only account for 32% of the total energy budget. The remaining 68% is described by the mysterious dark energy. There is relatively less effort toward identifying dark energy within the particle physics community due to the limited scope of well-motivated models testable with present-day technology. Nonetheless, a successful characterization of dark energy would be equally significant, if not more so, as the identification of DM in advancing fundamental physics.

### 1.1.2 Particle Models for Dark Matter

Since all forms of matter interact gravitationally, gravitational probes tell us little about its particle physics nature. Moreover, the universe as we observe it today tells us that the non-gravitational interactions of DM with itself or with ordinary matter must be either very weak or very infrequent. If this were not the case, DM would not admit a good effective description in $\Lambda$CDM cosmology as noninteracting. As a result, the design of effective search strategies requires detailed models that enable us to identify how weak DM signals may manifest and which of those signals can be sensitively probed. This section will provide a brief overview of several DM models and their motivations before proceeding to a more detailed discussion of axions and axion-like particles, which are the primary focus of this thesis.

For my model-building colleagues, constructing Beyond the SM theories with DM is not too difficult a task. Because DM must have weak interactions, it is fairly straightforward to add new sectors to the theory that include DM candidates without disrupting agreement with the SM. The much greater challenge is to identify DM candidates which might be considered natural or well-motivated in order to make use of limited resources in the search for new physics.[4]

One widely adopted guideline is to search for new physics candidates that can resolve an existing shortcoming of the SM while also accounting for the observed abundance of DM. This has motivated a considerable focus on weakly interacting massive particles (WIMPS) and axions. WIMPS might be related to the resolution of the hierarchy problem related to the mass of the higgs boson in the SM and have been the target of decades of direct, indirect, and accelerator-based searches [29]. Axions, which will be discussed at greater length later in this section, are motivated as a DM candidate that may also solve the Strong *CP* problem in quantum chromodynamics related to the absence of the neutron electric dipole moment [30]. An alternative principle is to consider DM candidates that may not solve other problems but instead are sufficiently qualitatively similar to the particles already known to us such that their addition to existing fundamental theories might be considered a "minimal" extension. For example, sterile neutrinos and dark photons, in analogy to the active neutrinos and photons of the SM, might be considered DM candidates within this category [31, 32]. Finally, we must acknowledge nature may not be simple or well-aligned with the assumptions that underlie these perspectives. In that sense, it is valuable to perform searches underneath the lamppost of existing experiments and observations so that our theory biases do not preclude the discovery of an "exotic" DM candidate within our reach.

This question of how we choose which DM models to search for might seem academic, but it is a necessary first step in performing any search. In addition to being challenging to detect by construction, different DM candidates produce qualitatively distinct signals that require unique

[4]Note that "natural" has a technical meaning, but we restrict ourselves to its nontechnical meaning here.

search strategies. Although there may be synergies between searches for various DM candidates, DM will likely not be identified in a model-agnostic search. As a result, we must decide what we are searching for before searching for it, and the principles and perspectives we have discussed are merely ways of navigating a model landscape that may be too broad to ever be fully explored.

## 1.2 Axion and Axion-like Particle Dark Matter

In light of the challenging decisions required in searches for DM, this section aims to introduce the axion as a DM candidate and explain its status as a particularly well-motivated one. Here, we are adopting the perspective that one should look for DM candidates that relate to existing problems within the SM, and in the case of the axion, we are motivated by the Strong *CP* Problem in QCD. Following along the lines of [33], we will begin with a nontechnical discussion that provides an intuitive picture of the Strong *CP* Problem before we proceed to its more technical statement and resolution with the introduction of the axion.

### 1.2.1 A Classical Analogy for the Strong *CP* Problem

We start with a classical analogy for the Strong *CP* problem by considering a neutron, a familiar SM particle comprised of one up quark and two down quarks. All these quarks have an electric charge, but because the up quark has charge $+2e/3$ and each down quark $-e/3$, the neutron has no total electric charge, *i.e.*, it has no electric monopole moment. This means that if we expose a neutron at rest to an electric field, the neutron will stay exactly where we placed it. Similarly, we can think about a compass and its needle. The needle has zero total magnetic charge, *i.e.* no magnetic monopole moment,[5] which is why it stays on the compass rather than flying off in the direction of the Earth's magnetic field. However, the needle has "positive" and "negative" ends, making it a magnetic dipole and causing point north to aligns with the Earth's magnetic field. As we have now seen in this example, having zero monopole moment while having nonzero dipole moment is perfectly acceptable and readily realized in nature.

This line of reasoning leads us to the natural question of whether the neutron has an electric dipole moment, which we can imagine testing by observing whether or not a neutron aligns itself in a particular direction in the presence of an electric field. But before we perform such an experiment, we would want to equip ourselves with some theory expectations to understand our result. If we treat the quarks inside the neutron as point charges, we can compute the expected electric dipole moment by integrating the charge density weighted by the displacement vector over the neutron

[5] Magnetic monopoles do not appear to be realized in nature.

volume to obtain

$$\mathbf{d} = \int dV \mathbf{r}\rho(r) = \frac{e}{3}(2\mathbf{r}_u - \mathbf{r}_{d,1} - \mathbf{r}_{d,2}) \tag{1.2}$$

where $\mathbf{r}_u$ is the location of the up quark and $\mathbf{r}_{d,1}$ and $\mathbf{r}_{d,2}$ are the locations of the two down quarks within the neutron. In this problem, we have nine unknowns (three spatial coordinates for each of the three quarks) and three constraints imposed by the vanishing of each component of the dipole vector. This tells us that generic configurations of the quarks will typically have some nonvanishing dipole moment, and only special configurations of the quarks, such as when the three are collinear, and each of the down quarks is an equal distance from the up quark, will produce a vanishing neutron electric dipole moment (NEDM). Moreover, the distribution of quarks inside the neutron is governed by the strong force of QCD. Because the QCD interactions are independent of electric charge, it would be surprising if QCD conspired to arrange the quarks so that the electric dipole would vanish. Without an informed guess for the precise charge distribution of a neutron, we can make a rough dimensional estimate of the size of the dipole moment by

$$d \approx e r_{NS} \approx 10^{-13}\, e \times \mathrm{cm} \tag{1.3}$$

where we have used the neutron radius of approximately 1 fm.

Our estimate for the NEDM is small but not so small that it cannot be measured in the lab. The most sensitive existing measurements would be able to detect the NEDM so long as $d \gtrsim 10^{-26}\, e \times \mathrm{cm}$. So it is somewhat surprising that all attempts at measuring the NEDM have resulted in a measurement perfectly consistent with zero [34].

The absence of an NEDM has interesting implications for QCD. To see this, we consider the action of a charge-parity (CP) transformation on a neutron. For the NEDM, the *CP* transformation roughly acts as $q \to -q$ and $\mathbf{r} \to -\mathbf{r}$. We can then see that the dipole is invariant under the *CP* transformation since

$$\mathbf{d} = \frac{e}{3}(2\mathbf{r}_u - \mathbf{r}_{d,1} - \mathbf{r}_{d,2}) \to -\frac{e}{3}(-2\mathbf{r}_u + \mathbf{r}_{d,1} + \mathbf{r}_{d,2}) = \mathbf{d}. \tag{1.4}$$

We must also recall that neutrons are characterized by an intrinsic spin, which we will take to point in the direction $\hat{s}$, that transforms under a *CP* transformation $\hat{s} \to -\hat{s}$. Hence, the quantity $\hat{s} \cdot \mathbf{d}$, transforms to $-\hat{s} \cdot \mathbf{d}$, showing us that the neutron is not *CP* invariant if it is characterized by an NEDM. However, since the NEDM appears to be zero, $\hat{s} \cdot \mathbf{d} = 0$ is preserved under a *CP* transformation, suggesting that the QCD interactions which govern the structure of a neutron respect *CP* symmetry. As we will now see, given the structure of the theory of QCD, this is unexpected.

### 1.2.2 The Strong *CP* Problem in QCD

We can advance beyond our crude estimate through a calculation of the NEDM in QCD using quantum field theory, the framework used to describe particle physics interactions. For simplicity, we consider QCD restricted to the neutron's consistuents, the gluons, up quarks, and down quarks, but the generalization to the full theory of QCD is straightforward.

For our two-flavor theory of QCD, we have the lagrangian density

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \sum_{q\in\{u,d\}} \left[ i\bar{q} \not{D} q - m_q \bar{q} e^{i\theta_q \gamma^5} q \right] \tag{1.5}$$

where $G$ is the gluon field strength tensor, $u$ and $d$ are the four-component quark spinors, $\not{D}$ is the covariant derivative contracted with the gamma matrices, and $m_u$ and $m_d$ are the magnitudes of the quark masses. The quark masses are generally complex, which generally produces *CP* violation, as they arise through electroweak dynamics which also mix the quark mass eigenstates through the Cabibbo–Kobayashi–Maskawa matrix. We therefore define $\theta_u$ and $\theta_d$ as the complex phases of the quark masses.

One possibility is that although the complex-valued quark masses might be expected to produce *CP* violating effects, these are in turn removed by an underlying symmetry of QCD. Indeed, we can consider the action of a chiral rotation with angle $\alpha_q$ on our quarks, which acts as

$$q \to e^{i\alpha_q \gamma^5} q, \qquad \bar{q} \to \bar{q} e^{i\alpha_q \gamma^5}. \tag{1.6}$$

In a classical theory, if we chirally rotate both the up quark and down quark by angles, $\alpha_u$ and $\alpha_d$, respectively, then we obtain the lagrangian

$$\mathcal{L}^{\text{classical}}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \sum_{q\in\{u,d\}} \left[ i\bar{q} \not{D} q - m_q \bar{q} e^{i(\theta_q + 2\alpha_q)\gamma^5} q \right]. \tag{1.7}$$

In the massless quark limit, the lagrangian is invariant under chiral rotations of the quarks, a property we refer to as a $U(1)_A$ symmetry. In the presence of quark masses, choosing $\alpha_q = \theta_q/2$ would remove the quark phases, thereby eliminating *CP* violation from the QCD sector. However, this $U(1)_A$ symmetry of the lagrangian is anomalous, meaning that it is broken in the quantized theory. After attempting to rotate away the quark mass phases in quantized QCD, we instead obtain

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4} G^a_{\mu\nu} G^{a,\mu\nu} + \frac{g^2(\theta_u + \theta_d)}{32\pi^2} G^a_{\mu\nu} \tilde{G}^{a,\mu\nu} + \sum_{q\in\{u,d\}} \left[ i\bar{q} \not{D} q - m_q \bar{q} q \right], \tag{1.8}$$

where $\tilde{G}_{\mu\nu} = \epsilon_{\mu\nu\alpha\beta} G^{\alpha\beta}/2$ is the Hodge dual of the gluon field strength tensor and $g$ is the QCD

coupling constant. Even though we have removed the CP violation from the quark mass terms, this $G\tilde{G}$ term, which is a total derivative, is also CP-violating, showing us that *CP* violation is a inescapable feature of our QCD lagrangian.

Observing that our theory in Eq. 1.5 was hiding a $G\tilde{G}$ interaction term motivates a more general QCD lagrangian of the form

$$\mathcal{L} = -\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} - \frac{g^2\theta}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a,\mu\nu} + \sum_{q\in\{u,d\}}\left[i\bar{q}\not{D}q - m_q\bar{q}e^{i\theta_q\gamma^5}q\right], \tag{1.9}$$

where $\theta$ is an angular phase setting the strength of the new $G\tilde{G}$ interaction. After rotating away the phases from the quark masses, we obtain

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}G^a_{\mu\nu}G^{a,\mu\nu} - \frac{g^2\bar{\theta}}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a,\mu\nu} + \sum_{q\in\{u,d\}}\left[i\bar{q}\not{D}q - m_q\bar{q}q\right], \tag{1.10}$$

where we have defined $\bar{\theta} = \theta - \theta_u - \theta_d$. It is then clear that generalizing our argument to include all known quark flavors merely requires including additional $\theta_q$ within the definition of $\bar{\theta}$. With the lagrangian cast in this form, we can see that the strength of CP violating effects in QCD, such as an NEDM, is determined by $\bar{\theta}$. The full field-theoretic calculation for the NEDM in terms of $\bar{\theta}$ was performed in [35], finding

$$d \approx 3.6 \times 10^{-16}\,\bar{\theta}\,e \times \text{cm}. \tag{1.11}$$

This calculation, in combination with existing NEDM constraints, requires $\bar{|}\theta| \lesssim 10^{-10}$. This very small value of $\bar{\theta}$ is surprising as it requires $\theta$, a free parameter of the theory which could take values between 0 and $2\pi$, to precisely cancel the sum of the phases of the quark masses. This mystery of why such a unnecessarily precise cancellation resulting in the absence of *CP* violation in QCD would be realized in nature is then referred to as the Strong *CP* Problem.

### 1.2.3 Solving the Strong *CP* Problem and Dark Matter with the Axion

The empirically small value of the theta angle that is without anthropic resolution [36] suggests the possibility of as-of-yet undiscovered dynamics that demand $\bar{\theta} = 0$.[6] The discovery of those dynamics would then resolve the Strong *CP* Problem. A possible solution was introduced by Roberto Peccei and Helen Quinn, who suggested the introduction of a new degree of freedom called the axion, which would enter the CP-violating interaction term as

$$\mathcal{L} \supset -\frac{g^2(\bar{\theta} - a/f_a)}{32\pi^2}G^a_{\mu\nu}\tilde{G}^{a,\mu\nu} \tag{1.12}$$

[6] By this we mean that our existence as observers of the universe seems possible for any value of $\bar{\theta}$.

where $a$ is the axion field, and $f_a$ is the axion decay constant. As a result of the axion's interactions with gluons, the axion experiences a potential at low energies of the form

$$V(a) = m_a^2 f_a^2 \left[1 - \cos\left(\bar{\theta} - \frac{a}{f_a}\right)\right] \tag{1.13}$$

where $m_a$ is the QCD-induced axion mass [37]. In order to minimize its potential, the axion will take on a vacuum expectation value (vev) of $a = f_a\bar{\theta}$, eliminating the interaction term that appears in Eq. 1.12. This process is known as the Peccei-Quinn (PQ) mechanism, and it represents the leading candidate for the solution to the Strong *CP* Problem [38, 39].

Arguably the simplest axion model that solves the Strong *CP* Problem was introduced in [40] and is known as the KSVZ axion. In this model, we consider a new quark $\psi$ which interacts with gluons as well as a complex scalar $\phi$ described by

$$\mathcal{L} \supset i\bar{\psi}\not{D}\psi + (\phi\bar{\psi}_L\psi_R + \phi^\dagger\bar{\psi}_R\psi_L) + |\partial\phi|^2 - m^2|\phi|^2 + \frac{m^2}{2f_a^2}|\phi|^4, \tag{1.14}$$

where $\psi_{L/R}$ are the left- and right-handed two-component spinors of the quark, $m$ is the complex scalar mass, and $f_a$ is a dimensionful constant that sets the strength of the complex scalar's quartic interaction. The introduction of a new quark provides us a new anomalous $U(1)$ symmetry through which phases can be rotated in and out of $\bar{\theta}$. An inspection of the scalar potential reveals that the theory will undergo a spontaneous symmetry breaking in which the magnitude of the scalar acquires a vev of $\phi_0 = f_a$ while its phase is still free to vary. Neglecting fluctuations of the scalar magnitude about its vev, we can parametrize the complex scalar in terms of the axion phase $a$ by $\phi = \phi_0 \exp(ia/f_a)$. Our lagrangian will then contain

$$\mathcal{L} \supset i\bar{\psi}\not{D}\psi + \bar{\psi}f_a e^{ia\gamma^5/f_a}\psi, \tag{1.15}$$

where we have neglected the axion kinetic term. We have now realized a dynamical phase for the quark mass term, which will enter $\bar{\theta}$ precisely as in Eq. 1.12, allowing us to solve the Strong *CP* Problem. Alternative constructions, such as the DFSZ axion [41, 42] or the PQWW axion [43, 44] offer additional model realizations.[7]

Shortly after the PQ mechanism was proposed and realistic models for realizing the axion were introduced, it was observed that small fluctuations of the axion field about its vev could be detected and could also account for the DM [43–46], spurring excitement for axions as the solution to two outstanding problems in fundamental physics. We note, however, that other mechanisms could solve the Strong *CP* Problem, such as the Nelson-Barr mechanism, a massless up quark, and

[7] While the KSVZ and the DFSZ axions represent current axion model benchmarks, the PQWW is now excluded by various experiments.

mirror SM sectors [47–50].

### 1.2.4 Additional Axion Motivations

There exist many additional theoretical motivations for axion DM beyond just the Strong *CP* problem. Typically, however, they require sacrificing the assumption that the QCD axion which solves the Strong *CP* Problem is the same as the axion which comprises the DM. This is not to say that invoking alternate motivations for the axion requires an alternate solution for Strong *CP* Problem; there might exist several species of axions, one of which could be the QCD axion, with the remaining having alternate roles in particle physics, including accounting for the DM. This precise scenario may even be expected in String Theory, which predicts a spectrum of axions across a broad range of masses [51, 52]. Axions may also be features of Grand Unified Theories of the three[8] fundamental forces [53] or explain the universe's matter-antimatter asymmetry [54]. While a full review of the diverse set of theories that include axions or the outstanding problems they could solve is beyond the scope of this thesis, suffice it to say that compelling models give us every reason to believe they might be realized in the spectrum of nature.

## 1.3 Searching for Axion Dark Matter

Having laid out the multitude of reasons axions are considered a well-motivated DM candidate, we transition to a discussion of efforts toward detecting axion DM. We begin with a review of the status of experimental and observational constraints on the axion parameter space. We then provide an overview of the strategies used in search of axions through their coupling to photons in direct and indirect detection efforts, both of which are considered in work presented in this thesis. Finally, we conclude the section with a brief discussion of the basic techniques used to analyze data from direct and indirect detection searches that address the unique statistical challenges faced in axion detection.

### 1.3.1 Existing Constraints on Axion Dark Matter

Physically, axion detection is challenging because the range of viable masses for axion DM spans over ten orders of magnitude with a diverse set of physical observables that require various technologies for detection. This problem was exacerbated by the comparatively greater focus on WIMPs, which have been considered a leading DM candidate for decades, and it is only recently

[8] I have excluded gravity from this accounting as there are no obvious prospects for its inclusion in a realistic quantized theory, nor does the inclusion of gravity appear to be necessary for the unification of the remaining three forces.

that considerable efforts have been invested in axion searches. This makes it a particularly exciting time to search for axions. A raft of new experiments are coming online alongside the long-running ADMX experiment, allowing for the very real possibility that axion DM will be either discovered or excluded in the relatively near future.[9]

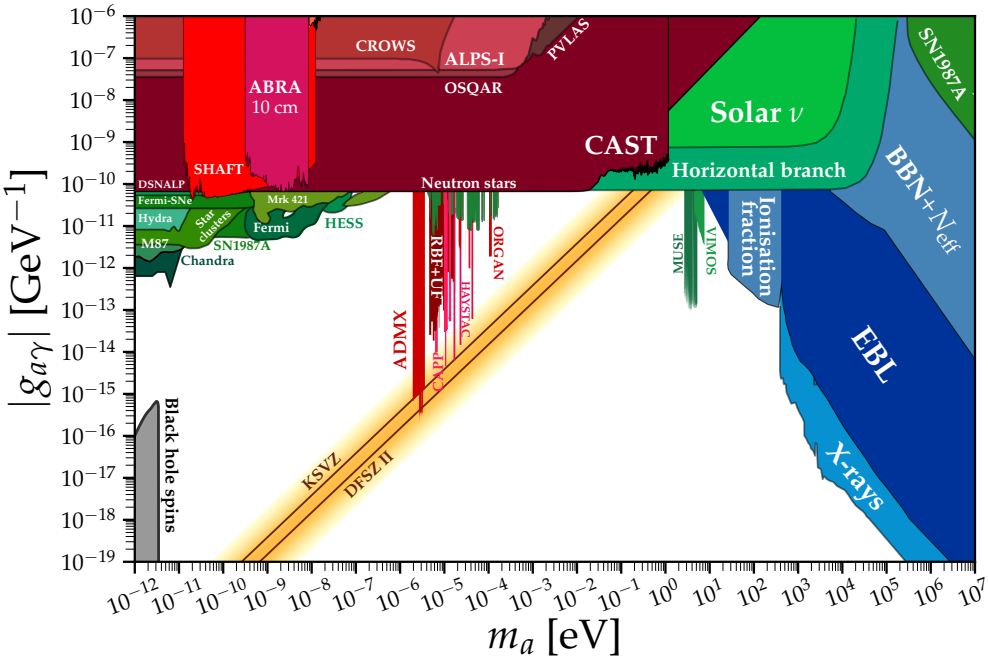



Figure 1.4: The axion parameter space in terms of the axion mass, $m_a$, and the strength of its coupling to photons, $g_{a\gamma\gamma}$. Green, blue, red, and gray regions are excluded by various searches. The orange band indicates the mass-coupling relation for an axion which solves the Strong *CP* Problem. White space is unconstrained. Several of these constraints have been produced by works included in this thesis. See https://cajohare.github.io/AxionLimits/ and accompanying references for details.

To illustrate the current status of the axion parameter space, we have included Fig. 1.4, which depicts the the plane of axion mass, $m_a$, and axion photon coupling strength, $g_{a\gamma\gamma}$. While axions may couple to many SM particles, we feature the axion-photon coupling as it is the coupling on which the work presented in this thesis is exclusively focused. Many existing exclusions are colorfully indicated in the figure, but for continued search efforts, the most relevant feature is the expanse of white in the figure, corresponding to unconstrained but theoretically viable axion DM parameter space. Likewise, the QCD axion band that depicts the expected strength of the axion-photon coupling for an axion that solves the Strong *CP* Problem, indicated in orange, remains unprobed above $10^{-4}$ eV and below $10^{-6}$,requiring the development and execution of new search

[9]"Relatively" is doing some heavy lifting here. The effort will take decades, but the relevant perspective is that it took over 40 years from the proposal of the PQ mechanism for a comprehensive detection program to begin in earnest.

strategies. In practice, this band is merely illustrative, as the axion-photon coupling is a model-dependent quantity. For instance, the two canonical QCD axion models, the KSVZ and DFSZ axions, predict a one-to-one relationship between the axion mass and its photon coupling, but alternative models which realize larger or much smaller photon couplings at a given mass are also possible. Discussion of the parameter space in terms of the similarly unconstrained axion couplings to other SM particles, including neutrons, protons, and electrons, can be found in [55–57].

### 1.3.2 Searching for the Axion through its Photon Coupling

With a some exceptions, such as the CASPEr experiment [58], the arguably best prospects for axion detection come from probes of the axion-photon coupling. This coupling is generically expected for axions or axion like particles, and appears as an additional term in the electromagnetism lagrangian as

$$\mathcal{L} = -\frac{1}{4} g_{a\gamma\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu} = -g_{a\gamma\gamma} a (\mathbf{E} \cdot \mathbf{B}) \tag{1.16}$$

where $a$ is the axion field, $F$ the electromagnetic field strength tensor, and $\tilde{F}$ its dual; this term can be alternatively formulated in terms of the axion field's coupling to the dot product of the electric field $\mathbf{E}$ and the magnetic field $\mathbf{B}$. This coupling of axions to photons is generically induced by the interaction of the axion with charged quarks. Schematically, we can think of this term as an interaction that occurs between a single axion and two photons. However, because axions are very light, if they are the DM, their number density is quite high, allowing for their description in the macroscopic limit as a classical field rather than individual particles.[10] From this perspective, we can consider the inclusion of an axion in our physical theory as a new classical scalar field that modifies the well-known Maxwell's equations for electromagnetism. These equations are:

$$-\ddot{a} + \nabla^2 a = m_a^2 a - g_{a\gamma\gamma} \mathbf{E} \cdot \mathbf{B} \tag{1.17}$$

$$\boldsymbol{\nabla} \cdot \mathbf{E} = \rho - g_{a\gamma\gamma} \mathbf{B} \cdot \boldsymbol{\nabla} a \tag{1.18}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \tag{1.19}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\dot{\mathbf{B}} \tag{1.20}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \dot{\mathbf{E}} + \mathbf{J} - g_{a\gamma\gamma} \left( \mathbf{E} \times \boldsymbol{\nabla} a - \dot{a} \mathbf{B} \right) \tag{1.21}$$

where $\rho$ is the charge density, $\mathbf{J}$ is the current density, and dotted quantities indicate their partial differentiation with respect to time [59].[11]

[10] Identical reasoning in a more familiar context leads us to consider classical fields rather than the dynamics of individual photons when discussing electromagnetism.

[11] We have chosen to present Maxwell's equations in vacuum here, but they are straightforwardly modified to their form in various media through use of constitutive equations.

We refer to this set of partial differential equations (PDEs) as the equations of axion electrodynamics, and they encapsulate various physical effects by which axions can be detected through their interactions with photons.[12] The modification to Gauss' law in Eq. 1.18 causes an electric field to be sourced by the axion field acting as an effective charge density in the presence of a magnetic field. Similarly, the modification to Ampere's law in Eq. 1.21 tells us that an axion field, in the presence of either an electric or magnetic field, will act as an effective current density and induce a magnetic field.[13] Finally, from Eq. 1.17, we see that the axion behaves as a time-oscillating field propagating through space that can be sourced in the presence of nonvanishing $\mathbf{E} \cdot \mathbf{B}$. Because the axion field is time-varying, so too will be the electromagnetic fields it sources. Moreover, by rewriting Maxwell's equations as a pair of second-order PDEs, it can be shown that a time-varying axion field in an electromagnetic background can induce propagating electromagnetic fields that can be detected as coherent photon waves. This process can be thought of as the conversion of axions to photons [61–63].

The physical processes primarily considered in this thesis are the induction of time-varying magnetic fields and propagating photon emission by the axion DM background. Axions may also be probed by "light shining through walls" experiments that generate axions from an $\mathbf{E} \cdot \mathbf{B}$ background that are then subsequently converted to photons [64]. We note that the study in Ch. 8 is qualitatively similar, though the axions are produced through the slightly more complicated Primakoff process. We will now describe the relevance of these mechanisms for both laboratory-based direct detection and astrophysical indirect detection.

#### 1.3.2.1 Direct Detection

Laboratory-based direct detection experiments seek to identify the physical observables associated with the presence of axion DM through precision measurements of nearly background-free environments. These observables are expected to be present, if weak, due to measurements of a local DM density in the solar neighborhood of $0.4\,\mathrm{GeV/cm^3}$ [65, 66]. As we see from the equations of axion electrodynamics, the axion can only induce a signal in the presence of background electromagnetic fields, with the strength of the interaction scaling linearly with the strength of the background field. Most experimental approaches provide a background magnetic field since large magnetic fields are more readily realized in the laboratory than large electric fields.

A class of experiments known as axion haloscopes represent the most mature axion detection strategy and search for the conversion of $\mu$eV-scale axions into detectable microwave photons in

[12] Other effects, like the possible decay of axions into two photons, are not conveniently captured in this formalism but may also be of relevance.

[13] Typically, dominant effects are realized by the $\dot{a}\mathbf{B}$ term as the contribution of $\mathbf{E} \times \boldsymbol{\nabla} a$ is suppressed by the nonrelativistic velocities of axion DM. In the presence of a relativistic axion background, this may no longer be true, see, *e.g.*, [60].

static $\mathcal{O}(1-10)$ T background magnetic fields. Notably, ADMX [67], HAYSTAC [68], and CAPP [69] have reached or are near benchmark QCD axion sensitivity. These experiments compensate for the small probability of axion conversion by engineering tunable resonant cavities that enhance the signal power. Cavity-based approaches have limited sensitivity to $m_a \gtrsim 50\,\mu$eV, and larger axion masses are targeted by the MADMAX experiment, which enhances the signal of higher-mass axion conversion with dielectric materials [70].

Lumped-element experiments, which instead aim to measure the time-varying magnetic field induced by an axion in the presence of a static background magnetic field, are of primary relevance to this thesis. A simplified description adapted from [71] is as follows. In the absence of an electric field, the solution to the axion's equation of motion at a fixed location is approximately

$$a(t) = \frac{\sqrt{2\rho_{\rm DM}}}{m_a}\sin(m_a t), \tag{1.22}$$

where we have determined the normalization of the field from the local DM density $\rho_{\rm DM}$. This axion field acts as an effective current of the form

$$\mathbf{J}_{\rm eff} = g_{a\gamma\gamma}\sqrt{2\rho_{\rm DM}}\cos(m_a t)\mathbf{B_0} \tag{1.23}$$

where $\mathbf{B}_0$ is the static background field in the modified form of Ampere's law. Casting Ampere's law in its integral form, rather than differential form as in Eq. 1.21, we then see that the axion induces a time-varying magnetic flux $\Phi$ which can be measured with a superconducting pickup loop with expected magnitude

$$\Phi(t) \sim g_{a\gamma\gamma}B_0 V_B\sqrt{2\rho_{\rm DM}}\cos(m_a t) \tag{1.24}$$

where $V_B$ is roughly the volume in which the static magnetic field of strength $B_0$ appears. This signal can then be readily identified as a monochromatic spectral excess at frequency $f = m_a/2\pi$.[14]

An interesting feature of lumped element searches is that they can operate as a broadband search, which is to say that long integration times and high-frequency readout make them sensitive to axions in the broad mass range of $10^{-14}$ to $10^{-6}$ eV in a single data collection. This contrasts with resonant haloscope experiments, which, by merit of their resonance, are only sensitive to a small range of masses with each data collection and must step through their narrower range of mass sensitivity with repeated tunings of their experiment. In addition to their broadband sensitivity, lumped element searches can simultaneously operate in a resonant readout mode that provides even greater sensitivity to targeted regions of parameter space. Several experiments, including

[14] In fact, the signal is only quasi-monochromatic as it is broadened by the finite velocity dispersion of DM in the Milky Way galactic halo, which we will study in Chs. 3 and 4.

ABRACADABRA [71–73], SHAFT [74, 75], and DM-Radio [76], are using the lumped element strategy, and much of the content of this thesis is devoted to work on statistical frameworks for these searches along with my work as a member of the ABRACADABRA collaboration.

#### 1.3.2.2 Indirect Detection

An alternative to laboratory-based direct detection is to instead search for astrophysical signals by pointing telescopes at locations on the sky we expect to be bright in the signal of DM. There are several notable advantages to the indirect detection strategy. First, the freedom to point our telescopes at any point in the sky allows us to target locations where DM is more abundant. While direction detection experiments are limited in their sensitivity by the local DM density, indirect detection enables to probe locations like the Galactic Center, where DM may be as much as $10^9$ times more abundant. Moreover, extreme astrophysical environments can realize conditions more conducive to generating axion signals than we can engineer in the lab. For instance, while axion conversion experiments may be capable of generating $10\,\mathrm{T}$ magnetic fields on the meter scale, neutron stars can have magnetic fields as strong as $10^{11}\,\mathrm{T}$ over kilometer scales. Moreover, the broad range of viable axion masses results in the possibility of axion signals at many energy scales, which can, in turn, be probed by telescopes operating at the appropriate energy. In this thesis, we consider signals that would appear in radio or X-ray, but DM may also produce signals at other frequencies.

However, there are several challenges associated with indirect detection. First, astrophysical systems are very far away from us; other than the sun, the nearest star to us is over a parsec away, with most sources of DM signals another thousand times further away, drastically reducing the intensity of any signal, which scales like $1/d^2$. Moreover, astrophysical environments have considerably larger modeling uncertainties and confounding backgrounds than carefully engineered laboratories. Finally, in the specific case of axion DM, because we need the DM signal to reach Earth, we depend on propagating photon signals and are forced to abandon the strategy of measuring time-varying axion-induced magnetic fields used to great success in lumped element searches and depend exclusively on the conversion process.

Given the diverse array of astrophysical observables that can be probed, we defer a more detailed discussion to individual chapters within this thesis, which treat searches for DM from three qualitatively different channels. We will merely remark that robust indirect detection searches require a careful balance in selecting on-sky locations that could contain bright DM signals but are also characterized by backgrounds that can be accurately modeled and do not introduce statistical or systematic uncertainties larger than a candidate signal. This notion smoothly transitions us to the final topic of discussion in this section.

### 1.3.3 Statistics for Dark Matter Detection

Direct and indirect searches alike collect data, and a key step in the search for a DM is a statistical analysis. Adopting a frequentist perspective, we develop a null hypothesis that the data do not contain a DM signal and an alternate hypothesis, that the data does indeed contain a DM signal, and then attempt to quantify the statistical evidence in favor of adopting the alternate hypothesis.[15] If we do not make a discovery, we then seek to set limits on the DM signal, akin to saying that had the signal been brighter than some threshold, we would have expected to see it. In practice, we construct nested hypotheses, infer model parameters with a technique known as maximum likelihood estimation, determine detection significances from the likelihood ratio, and construct confidence intervals to set limits, making use of a number of results in frequentist statistics which are excellently reviewed in [77]. In this section, we will provide a schematic overview of these procedures, followed by a brief discussion of the statistical challenges associated with DM detection.

#### 1.3.3.1 The Basics of Frequentist Analysis

Suppose we are operating a telescope and recording the number of photons at particular energies, obtaining a dataset $\mathbf{d} = \{d_1, d_2, ..., d_n\}$ corresponding to $d_1$ counts at energy $E_1$, $d_2$ counts at energy $E_2$, and so on. In this scenario, we would expect the number of photons at each energy to be drawn from a Poisson distribution characterized by some asymptotic event rate $c_1$ at energy $E_1$, etc. We would then first construct a null model $\mathcal{M}_0$ which attempts to predict the values of $c_1, ..., c_n$. A very simple model hypothesis might be that the event rate at each energy is the same, which is to say $c_1 = c_2 = ... = c_n = b$ and that our model $\mathcal{M}_0$ is now parametrized by the unknown background rate parameter $b$. We could then ask, for an assumed value of $b$, how likely are the data we observed. This likelihood denoted $\mathcal{L}(d|\mathcal{M}_0(b))$, is then given by

$$\mathcal{L}(d|\mathcal{M}_0(b)) = p(d_1|b) \times p(d_2|b) \times ... \times p(d_n|b) = \prod_i^n p(d_i|b), \tag{1.25}$$

where $p(d|b)$ is the probability of drawing $d$ counts from a Poisson distribution with count rate $b$. Since the number of observed counts at each energy is assumed to be independent, the likelihood is then directly computed by multiplying the individual probabilities of the data given the model at each energy. If we treat $b$ as an unknown parameter we want to infer, we can ask what value of $b$ maximizes the likelihood of the observed data. That value of $b$, denoted $\hat{b}$, is the maximum likelihood estimate of the parameter under the null hypothesis.

We can similarly treat our alternate hypothesis, which includes the possibility of a DM signal.

[15] Viable statistical analyses can also be performed with bayesian methodologies, which will not be discussed at length in this thesis and are relatively less commonly used in particle physics.

We will call this model $\mathcal{M}_1$, and it will be somewhat different from $\mathcal{M}_0$. The DM physics tells us the "shape" of the signal, which is to say that if DM is indeed present, we would expect more photon counts in some energies and fewer in others. We will now additionally include a new parameter, $A$, which describes the total number of photons we expect in our model to come from DM. We can now compute the likelihood for our alternate hypothesis as

$$\mathcal{L}(d|\mathcal{M}_1(A,b)) = p(d_1|As_1+b)\times p(d_2|As_2+b)\times ... = \prod_i^n p(d_i|As_i+b), \tag{1.26}$$

where $s_i$ is describing the shape of the DM photon spectrum. Just as we determined a maximum likelihood for $b$ under the null, we can determine $\hat{\hat{A}}$ and $\hat{\hat{b}}$ by maximizing the likelihood of our alternate hypothesis under the joint choices of $A$ and $b$.[16]

We have now computed our maximum likelihoods under each hypothesis, and the simplest question we could ask is whether data are more or less likely under the alternate hypothesis than the null hypothesis. Quantitatively, this question asks if the ratio of the maximum likelihoods, denoted $\Lambda$ and given by

$$\Lambda = \frac{\mathcal{L}(d|\mathcal{M}_1(\hat{\hat{A}},\hat{\hat{b}}))}{\mathcal{L}(d|\mathcal{M}_0(\hat{b}))}, \tag{1.27}$$

is greater than 1. However, there is a subtlety. The alternate hypothesis was constructed to contain the null hypothesis, which is realized in the $A \to 0$ limit. As a result, the alternate hypothesis can never result in a maximum likelihood that is less than the maximum likelihood under the null hypothesis, and $\Lambda \geqslant 1$. Moreover, because the alternate hypothesis has one more free parameter than the null hypothesis, we would typically expect that $\Lambda > 1$, even if the "true" underlying physical model from which the data were drawn was indeed the null model, due to statistical error.

The more informed question we could ask is if the likelihood ratio is so much greater than $1$ that it exceeds our threshold for belief that the data came from a realization of the null hypothesis, leading us to accept the alternate hypothesis. This question can be answered in a precise way by invoking Wilks' theorem [78], which tells us that the quantity $\mathrm{TS}$ defined by

$$\mathrm{TS} = -2\log(\Lambda) \tag{1.28}$$

asymptotically follows as a $\chi^2$-distribution with one degree of freedom under the assumption that the null hypothesis is the true description of the data.[17] The conventional threshold for accepting the alternate hypothesis in particle physics is the $5\sigma$ threshold, corresponding to when $\mathrm{TS} > 25$.

[16] Since $\hat{b}$ and $\hat{\hat{b}}$ were determined under different hypotheses, generically $\hat{b} \neq \hat{\hat{b}}$.

[17] More generally, for nested hypotheses, the TS will follow a $\chi^2$-distribution where the number of degrees of freedom is equal to the number of model parameters in the alternate hypothesis but not in the null hypothesis.

This is merely a sophisticated way of saying that we will only accept the alternate hypothesis if the $\mathrm{TS}$ is so large that if the null model did perfectly describe the data and we were to repeat our experiment $10^6$ times, we would never expect to see such a $\mathrm{TS}$ larger than the one we have observed.

Once we have determined our $\mathrm{TS}$ and decided whether or not to accept the alternate hypothesis, it is a relatively straightforward task to determine statistical uncertainties on our maximum likelihood estimated parameters. If we have not detected DM in our experiment, Fisher information considerations allow us to estimate a $95^{\mathrm{th}}$ percentile upper limit on the DM signal strength parameter $A$ by

$$A_{95} \approx \hat{\hat{A}} \left( 1 + \frac{1.65}{\sqrt{\mathrm{TS}}} \right). \tag{1.29}$$

In a realistic context, this signal strength parameter is determined by many factors, but assuming a well-characterized instrumental sensitivity and DM signal generation mechanism, we may extract a limit on the DM coupling from $A_{95}$. We also note in passing that this discussion has been structured to be didactic rather than rigorous. In careful analyses, null and alternate hypotheses are constructed before examining the data; otherwise, one risks building models biased either for or against discovery. Moreover, in many instances, the application of Wilks' theorem is invalid for one or many reasons, such as parameter degeneracies, failure of the data to reach the asymptotic limit, and improperly nested hypotheses, that result in nongaussian likelihoods. Much of the work presented in this thesis is focused on *a posteriori* validations of statistical analyses performed with Wilks' theorem and the development of more appropriate statistical tests and limit-setting procedures, *e.g*, through Monte Carlo simulation and data-driven estimates of systematic uncertainties.

#### 1.3.3.2 Statistical Challenges for Dark Matter Detection

While the previous section laid out a relatively straightforward methodology for the statistical analysis of data collected in DM searches, working with non-idealized data is rarely straightforward. Moreover we have seen that the DM parameter space, even when restricted to just the axion, is broad. As a result, we may not have the luxury of multiple independent, complementary probes, and without high-performing analysis frameworks with well-understood sensitivities, we risk never discovering a DM particle that was within our reach. Critically, the likelihood-ratio tests we have discussed are merely comparative tests of the goodness-of-fits of null and alternate hypotheses. If neither hypothesis provides good descriptions of the data, we can expect spurious discoveries or failures to discover real signals, and addressing this problem requires developing physically motivated models that accurately describe the potentially complex phenomena contributing to statistical and systematic backgrounds.

Moreover, the search for DM has rapidly become a big data problem. Many of the DM searches

discussed in this thesis produce terabyte-sized datasets and contain somewhere between hundreds and millions of statistically independent locations of data. Thus, not only must our analysis frameworks be accurate, they must be efficient while also enabling the quantification of uncertainties and the accurate interpretation of global significances of candidate detection. For instance, in the data collected by ABRACADABRA, there exist nearly $10^7$ statistically possible signal locations, which, in the perfectly modeled limit, we would expect to contain $\sim 5$ excesses with $\mathrm{TS} > 25$. Moreover, the analysis methodologies should be sufficiently flexible enough to describe the variations in background effects that may appear at those numerous possible signal locations while simultaneously preserving detection sensitivity. Finally, our analyses must be sophisticated in using experimental and observational controls to diagnose systematic failures our modelling assumptions and detector performance. Continuing with the ABRACADABRA example, one of our single largest sources of unanticipated background power was the collection of AM radio stations operating within the state of Massachusetts, and our failure to diagnose them would have represented the difference between presenting null results and announcing the discovery of dozens of strongly-coupled axions.

#### 1.3.3.3 A Summarizing Remark and Overall Perspective

We have spent nearly equal time discussing the statistical methods and challenges for axion detection as we have in reviewing the theoretical motivations, signal generation mechanisms, and experimental and observation prospects for axion DM. This is not an accident. The objective of the work presented in this thesis is three-fold: the identification of promising detection strategies for axion DM (or DM more generally), the development of the robust analysis frameworks for the search of those signals, and the execution of novel searches for axions using the multitude of existing and planned probes, even those which were not originally designed with DM in mind but nonetheless may reveal the nature of DM to us.

## 1.4 Organization of this Thesis

In *Chapter 2*, we use high-performance computing techniques to simulate the production of axion DM in the early universe beginning at times after the end inflation but before the breaking of the PQ symmetry and ending when the axion has acquired its zero-temperature mass through the QCD phase transition. By performing these simulations for a range of possible axion masses, we find a relationship between the axion mass and its late-time cosmological abundance. This allows a specific prediction for the mass of an axion that comprises the DM, providing direction to experiments and observations which seek to detect axion DM. We also study the spectrum of overdensity perturbations that arise in the axion field due to nonlinearities in the axion's equations

of motion and $\mathcal{O}(1)$ differences in the axion field at the time it acquires at mass. This spectrum of overdensities suggests that axion DM may produce small-scale gravitationally bound structures known as axion minihalos with masses at and below $10^{-12}\,M_{\odot}$.
Based on:

- *Early-Universe Simulations of the Cosmological Axion* with Malte Buschmann and Benjamin R. Safdi; *Phys. Rev. Lett.*, 124(16):161103, 2020; **arXiv:1906.00967** [79]

In *Chapter 3*, our attention shifts to axion direct detection, and we study the statistics of measurements of local axion field amplitude performed by generic laboratory-based experiments. We treat the axion field as a gaussian random field measured at a single spatial location with a coherence time determined by astrophysical DM velocity distribution. Using this as a starting point, we develop a statistical analysis framework for axion direct detection data. We also explore the sensitivity of direct detection experiments to that underlying velocity distribution, a prospect known as axion astronomy, showing how the expected modulation of an axion signal can be used to validate a candidate detection and how velocity substructure in the axion phase-space distribution can enhance detection significance.
Based on:

- *Revealing the Dark Matter Halo with Axion Direct Detection* with Nicholas L. Rodd and Benjamin R. Safdi; *Phys. Rev. D,* 97(12):123006, 2018; **arXiv:1711.10489** [80]

In *Chapter 4*, we continue our work on the statistics of axion direct detection to explore the improvements in sensitivity that can be achieved with a network of axion detection experiments, a prospect that may be soon achieved by rapidly maturing experimental collaborations. This represents a generalization of the work discussed in Chapter 3, where just as the astrophysical DM velocity distribution defines a temporal coherence scale, it also defines a coherence length that correlates the simultaneous measurements made by spatially separated detectors. We show that for detectors separated by distances greater than the coherence length, the sensitivity to the axion scales with the number of detectors like $N^{1/4}$ while sensitivity for detectors within the coherence length scales like $N^{1/2}$. For detectors separated by distances comparable to the coherence length, increased sensitivity to the axion velocity distribution may be obtained using the analysis framework which we develop.
Based on:

- *Dark Matter Interferometry* with Yonatan Kahn, Rachel Nguyen, Nicholas L. Rodd, and Benjamin R. Safdi; *Phys. Rev. D,* 103(7):076018, 2021; **arXiv:2009.14201** [81]

In *Chapter 5*, our consideration of the statistics of axion direct detection advances beyond a purely theoretical one within the ABRACADABRA collaboration. We design and implement the

ABRACADABRA-10 cm prototype detector, which we use to collect approximately one month of data used to search for a small time-varying magnetic flux induced by the coupling of the axion field to the background magnetic field following the statistical procedures developed in *Chapter 3*. This search produced new laboratory-based constraints on sub-$\mu$eV axion DM and paves the way for future searches with lumped-element detection strategies.
Based on:

- *First Results from ABRACADABRA-10 cm: A Search for Sub-$\mu$eV Axion Dark Matter* with Jonathan L. Ouellet et al.; *Phys. Rev. Lett.,* 122(12):121802, 2019; **arXiv:1810.12257** [72]

- *Design and implementation of the ABRACADABRA-10 cm axion dark matter search* with Jonathan L. Ouellet et al.; *Phys. Rev. D,* 99(5):052012, 2019; **arXiv:1901.10652** [82]

In *Chapter 6*, we continue our work within the ABRACADABRA collaboration to search for axions with an improved ABRACADABRA-10 cm prototype detector that achieves greater sensitivity to the local axion field. We also develop improved analysis controls designed to address the increased significance of systematic backgrounds associated with our greater sensitivity. This work produces leading constraints on the axion DM hypothesis for neV mass axions.
Based on:

- *The search for low-mass axion dark matter with ABRACADABRA-10 cm* with Chiara Salemi et al.; Accepted in *Phys. Rev. Lett*; **arXiv:2102.06722** [73]

In *Chapter 7*, our focus again shifts, this time to the indirect detection of axion DM. Through their interaction with photons, axions can convert to photons in the strong magnetic fields of neutron star magnetospheres, producing nearly monochromatic radio frequency emission. In this work, we use dedicated radio observations of several neutron star dense targets, including the Galactic Center, M54, M31, as well as two isolated neutron stars with the Effelsberg 100-m Radio Telescope and the Robert C. Byrd Green Bank Telescope in L- and S-Band radio frequencies for searches for an axion conversion signal. We find no evidence for axion DM, producing novel leading constraints on previously unprobed parameter space and setting the stage for future radio searches.
Based on:

- *Green Bank and Effelsberg Radio Telescope Searches for Axion Dark Matter Conversion in Neutron Star Magnetospheres* with Benjamin R. Safdi et al.; *Phys. Rev. Lett.*, Phys. Rev. D, 103(7):076018, 2021; **arXiv:2004.00011** [83]

In *Chapter 8*, we continue to consider the indirect detection of axions using X-ray observations of super star clusters (SSCs). SSCs are home to hot young stars that can be luminous in axions produced in thermal processes inside the stellar cores, and these axions can, in turn, convert

to X-ray photons as they propagate through the Galactic Magnetic Field. Because the axions are produced in the stars before generating an observable X-ray signal, this probe does not depend on axions comprising all of the DM, unlike all other searches discussed in this thesis. We combine stellar modeling performed with the `Modules for Experiments in Stellar Astrophysics` code package and analyze archival data collected by the NuSTAR X-ray telescope to set constraints on the axion-photon coupling for axions with masses below 50 peV.
Based on

- *X-ray Searches for Axions from Super Star Clusters* with Christopher Dessert and Benjamin R. Safdi; *Phys. Rev. Lett.,* 125(26):261102, 2020; **arXiv:2008.03305** [84]

In *Chapter 9*, we make use of 20 years of X-ray data collected by the *XMM-Newton* telescope totaling 547 Ms of exposure to search for signals of decaying DM. We find no evidence for and thereby setting leading constraints on decaying dark matter in the 5-16 keV mass range. Most of the discussion in this chapter is cast in the language of searches for sterile neutrinos that decay in the Milky Way galactic halo to produce diffuse X-ray spectral line emission with a spatial morphology determined by the DM spatial distribution as they represent the benchmark model for keV-scale for decaying DM. However, this work has been included within this thesis on axion detection for two reasons: these constraints on decaying DM also place strong limits on the axion-photon coupling for keV-scale axions, which may decay to two photons, and because the observation and analysis strategies used in sterile neutrino searches are highly similar to those used in axion searches. In particular, this chapter represents the first application of a non-parametric inference tool called Gaussian process modeling to astrophysical searches for DM and paves the way for its application to future analyses in the search for axion DM.
Based on

- *Deep search for Decaying Dark Matter with XMM-Newton Blank-Sky Observations* with Marius Kongsore, Christopher Dessert, Yujin Park, Nicholas L. Rodd, Kyle Cranmer, and Benjamin R. Safdi; *Phys. Rev. Lett.,* 127:051101, 2021; **arXiv:2102.02207** [85]

Finally, in *Chapter 10*, we offer brief concluding remarks.

# CHAPTER 2

# Simulations of Axion Production in the Post-Inflationary Misalignment Scenario

The quantum chromodynamics (QCD) axion is a well-motivated dark-matter (DM) candidate capable of producing the present-day abundance of DM while also resolving the strong *CP* problem of the neutron electric dipole moment [38, 39, 43–46, 86]. The axion is an ultralight pseudo-scalar particle whose mass primarily arises from the operator $aG\tilde{G}/f_a$, with $a$ the axion field, $G$ the QCD field strength, $\tilde{G}$ its dual, and $f_a$ the axion decay constant. Below the QCD confinement scale, this operator generates a potential for the axion; when the axion minimizes this potential it dynamically removes the neutron electric dipole moment, thus solving the strong *CP* problem. In the process the axion acquires a mass $m_a \sim \Lambda_{\rm QCD}^2/f_a$, with $\Lambda_{\rm QCD}$ the QCD confinement scale. The standard ultraviolet completion of the axion low-energy effective field theory is that the axion is a pseudo-Goldstone boson of a symmetry, called the Peccei-Quinn (PQ) symmetry, which is broken at the scale $f_a$ [40–42, 87, 88].

The cosmology of the axion depends crucially on the ordering of PQ symmetry breaking and inflation. If the PQ symmetry is broken before or during inflation, then inflation produces homogeneous initial conditions for axion field and generically the cosmology is relatively straightforward [89]. In this work we focus on the more complex scenario where the PQ symmetry is broken after reheating. Immediately after PQ symmetry breaking, the initial axion field is uncorrelated on scales larger than the horizon, with neighboring Hubble patches coming into causal contact in the subsequent evolution of the Universe. This leads to complicated dynamical phenomena, such as global axion strings, domain walls, and non-linear field configurations called oscillons (also referred to as axitons) [90–96].

We perform numerical simulations to evolve the axion field from the epoch directly before PQ symmetry breaking to directly after the QCD phase transition. Once the field has entered the linear regime after the QCD phase transition, we analytically evolve the free-field axion to matter-radiation equality. The central motivations for this work are to (i) quantify the spectrum of small-scale ultracompact minihalos that emerges through the non-trivial axion self-interactions

and initial conditions, and (ii) to determine the $m_a$ that leads to the correct DM density in this scenario.

The post-inflation PQ symmetry breaking cosmological scenario has been the subject of considerable numerical and analytic studies. It has been conjectured that this cosmology gives rise to ultra-dense compact DM minihalos with characteristic masses $\sim 10^{-13}$-$10^{-11}$ $M_\odot$, though we show that the typical masses are actually smaller than this, and initial DM overdensities of order unity [91–93, 96–99]. In this work we compute the minihalo mass function precisely, combining state-of-the-art numerical simulations with a self-consistent cosmological picture. Understanding this mass function is important as it affects the ways that we look for axions in this cosmological scenario. For example, it has been claimed that microlensing by minihalos and pulsar timing surveys [100] may constrain the post-inflation PQ symmetry breaking axion scenario [99], but these analyses rely crucially on the form of the mass function at high overdensities and masses. The axion minihalos may also impact indirect efforts to detect axion DM through radio signatures [101–106].

A precise knowledge of the $m_a$ that gives the observed DM density is of critical importance for axion direct detection experiments [67, 70–72, 76, 82, 107–111]. We find $m_a = 25.2 \pm 11.0$ $\mu$eV, which is within range of *e.g.* the HAYSTAC program [110]. Our axion mass estimate is similar to that found in recent simulations [112] but disagrees substantially with earlier semi-analytic estimates [113–119]. The minihalo mass function is also important for interpreting the results of the laboratory experiments. If a large fraction of the energy density of DM is in compact minihalos, it is possible that the expected DM density at Earth is quite low or highly time dependent, which means that direct detection experiments would need to be more sensitive than previously thought or use an alternate observing strategy.

The original simulations that tried to estimate the minihalo mass function were performed in [92] on a grid of size $100^3$. Ref. [92] found oscillons (soliton-like oscillatory solutions) that contribute to the high-overdensity tail of the mass function. Note that oscillons are analogous to the breather solutions found in the Sine-Gordon equation (see *e.g.* [120]). Recently [96] performed updated simulations on a grid of size $8192^3$. Our results expand on and differ from those presented in [96] in many ways, such as through our initial state that begins before the PQ phase transition, measurement of the overall DM density, evolution to matter-radiation equality, and accounting of non-Gaussianities.

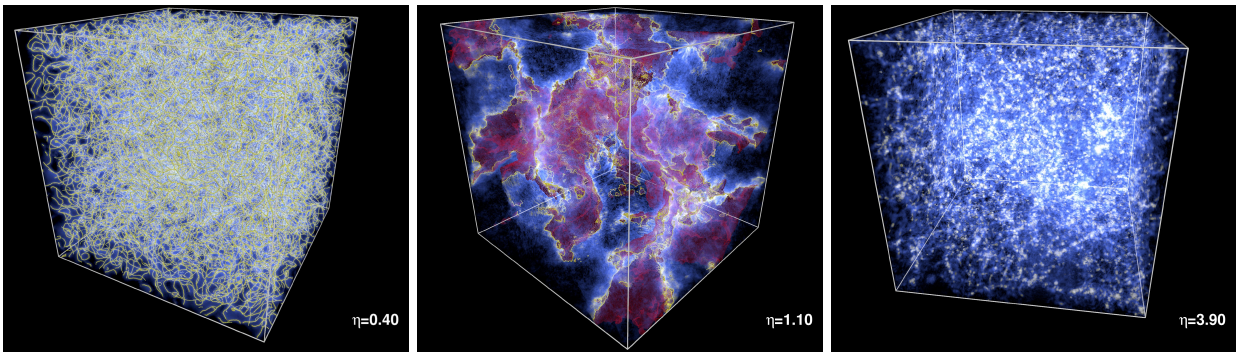



Figure 2.1: Each panel illustrates the string network (yellow strings), domain walls (red mesh), and energy density of the axion field (blue-white intensity) before (left), during (middle), and after (right) the QCD phase transition (see animation).

## 2.1 Simulation Setup

We begin our simulations with a complex scalar PQ field $\Phi$, with Lagrangian

$$\mathcal{L}_{PQ} = \frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}\left(|\Phi|^2 - f_a^2\right)^2 - \frac{\lambda T^2}{6}|\Phi|^2 - m_a(T)^2 f_a^2[1 - \cos \mathrm{Arg}(\Phi)], \tag{2.1}$$

with $T$ the temperature, $\lambda$ the PQ quartic coupling strength, and $m_a(T)$ the temperature-dependent axion mass generated by QCD [121]. The parametrization of the temperature-dependent mass is adopted from the leading-order term in the fit in [116]. Explicitly, the axion mass is parametrized by

$$m_a(T)^2 = \min\left[\frac{\alpha_a \Lambda^4}{f_a^2 (T/\Lambda)^n},\, m_a^2\right], \tag{2.2}$$

for $\alpha_a = 1.68 \times 10^{-7}$, $\Lambda = 400\,\mathrm{MeV}$ and $n = 6.68$. The growth of the mass is truncated when it reaches its zero-temperature value, which occurs at $T \approx 100\,\mathrm{MeV}$ independent of the axion decay constant. The zero-temperature mass is given by $m_a \approx 5.707 \times 10^{-5} (10^{11}\ \mathrm{GeV}/f_a)$ eV [122]. We present results using a similar but more recent mass parametrization from [123] in the Appendices; the difference between these results is used to quantify a source of systematic uncertainty.

For the PQ-epoch simulations we begin well before the breaking of the PQ symmetry at a time when the PQ field is described by a thermal spectrum. We fix $\lambda = 1$ for definiteness. The simulation is performed by evolving the equations of motion on a uniformly spaced grid of side-length $L_{PQ} = 8000$ in units of $1/(a_1 H_1)$, with $a_1$ ($H_1$) the scale factor (Hubble parameter) at the temperature when $H_1 = f_a$, at a resolution of $1024^3$ grid-sites. We use a standard leap-frog algorithm in the kick-drift-kick form with an adaptive time-step size and with the numerical Laplacian calculated by the seven-point stencil. It is convenient to use the rescaled conformal time $\tilde{\eta} = \eta/\eta_1$, where $\eta_1$ is the conformal time at which point $H(\eta_1) \equiv H_1 = f_a$. The simulation begins at $\tilde{\eta}_i = 0.0001$

and proceeds with initial time-step $\Delta\tilde{\eta}_i = 0.004$ until $\tilde{\eta} = 250$, after which a variable time-step calculated by $\Delta\tilde{\eta}_i(250/\tilde{\eta})$ is used to maintain temporal resolution of the oscillating PQ fields. Convergence was tested by re-running small time intervals of the simulation at smaller time steps. The PQ fields evolve from their initial thermal configuration until the PQ phase transition occurs at $\tilde{\eta} \approx 280$, after which the radial mode $|\Phi/f_a|$ acquires its vacuum expectation value (VEV). We simulate until $\tilde{\eta}_f = 800$ in order to proceed to a time at which fluctuations around the radial mode VEV have become highly damped.

Note that the difference in $\tilde{\eta}$ between $\tilde{\eta} = 1$ and the PQ phase transition is proportional to $\sqrt{m_{\rm pl}/f_a}$, with $m_{\rm pl}$ the Planck mass. The actual choice of $f_a$ here does not play an important role since we evolve the axion-string network into the scaling regime. In the left panel of Fig. 2.1 we show the final state of our simulation at the completion of the PQ simulation. The string network is seen in yellow, with the blue colors indicating regions of higher than average axion density. The length of the simulation box at this point is around $8000/(a(\tilde{\eta}_f)H(\tilde{\eta}_f))$, and we indeed find that there is around one string per Hubble patch as would be expected in the scaling regime.

We use the final state of the PQ-epoch simulation as the initial state in our QCD-epoch simulation. To do so we assume that the axion-string network remains in the scaling regime between the two phase transitions (see, *e.g.*, [117]). Recently [124] found evidence for a logarithmic deviation to the scaling solution and we confirm this behavior in the appendices. However, we perform tests to show that this deviation to scaling likely has a minimal impact on both the minihalo mass function and on the DM density, though we still assign a systematic uncertainty to our DM density estimate from the scaling violation.

Anticipating requiring greater spatial resolution for late-times in our QCD simulation, we increased the resolution of our simulation to $2048^3$ grid-sites with a nearest-neighbor interpolation algorithm. We re-interpreted the physical dimensions of our box from side-length $L_{\rm PQ} = 8000$ in PQ spatial units to $L_{\rm QCD} = 4$ in units of $1/(a_1 H_1)$. These units are defined such that $H_1 \equiv H(\eta_1^{\rm QCD}) = m_a(\eta_1^{\rm QCD})$ at conformal time $\eta_1^{\rm QCD}$. Further, we use the dimensionless parameter $\hat{\eta} = \eta/\eta_1^{\rm QCD}$. While our PQ simulation ended at $\tilde{\eta}_f = 800$ in PQ units, the start time in the QCD phase transition is taken to be $\hat{\eta}_i = 0.4$ in the QCD units. Modes enter the horizon as their co-moving wavenumber becomes comparable to the co-moving horizon scale, which scales linearly with $\eta$. Therefore, by maintaining the ratio $L_{\rm PQ}/\tilde{\eta}_f = L_{\rm QCD}/\hat{\eta}_i$, we preserve the status of our modes with respect to horizon re-entry.

We then evolve the equations of motion with our initial step size now chosen to be $\Delta\hat{\eta}_i = 0.001$. As before, we adaptively refine our time step size, using time-step $\Delta\hat{\eta}_i(1.8/\hat{\eta})^{3.34}$ after $\hat{\eta} = 1.8$, to maintain resolution of the oscillating axion field. We simulate until $\hat{\eta}_f = 7.0$, periodically checking if all topological defects have collapsed. When this occurs, we switch to axion-only equations of motion for computational efficiency, since past this point the radial mode does not

play an important role.

The conformal time $\hat{\eta}_c$ at which the mass growth was cut off corresponds to the physical value of the axion decay constant since it relates the temperature $T_1$ at which the axion begins to oscillate and the cutoff temperature $T_c \approx 100$ MeV at which the axion reaches its zero-temperature mass. We performed simulations at five values of $\hat{\eta}_c$ uniformly spaced between $2.8$ and $3.6$. These values are chosen to access different values of $f_a$ while still preserving a hierarchy between $\hat{\eta}_c$ and our simulation end time in order to provide sufficient time for the field to relax. At each of the five values of $\hat{\eta}_c$, we performed simulations at five different values of the parameter $\tilde{\lambda}$, defined by $\tilde{\lambda} \equiv \lambda f_a^2 / m_a(\hat{\eta}_1)^2$. This parameter can be interpreted as the squared mass of the radial PQ mode relative to the axion mass, at conformal time $\hat{\eta}_1$. In order for excitations of the radial mode to be well-resolved in our simulation, we require that the resolution of our simulation $\Delta\bar{x}$, with $\bar{x} = a_1 H_1 x$ and $x$ the spatial coordinate, be such that $1/(\hat{\eta}\tilde{\lambda}^{1/2}\Delta\bar{x}) > 1$, making simulations for realistic axion parameters $\tilde{\lambda}$ impossible. We break the relation between $\tilde{\lambda}$ and $f_a$ and consider $\tilde{\lambda} = [1024, 1448, 3072, 3584, 5504]$ in order to study the impact of this parameter. We provide additional arguments in the appendices supporting that while not physical these $\tilde{\lambda}$ values should preserve the correct timescale for physical defect network collapse (see also [125]).

We illustrate three important phases of the QCD-epoch simulation in Fig. 2.1. The left-most panel is the initial state discussed previously in the context of the PQ-epoch simulation final-state. When $m_a(\hat{\eta}) = 3H(\hat{\eta})$ at $\hat{\eta} \approx 1.22$, strings grow longer and become less numerous, with domain walls forming on surfaces bounded by the strings. This is illustrated in the middle panel, with red colors indicating domain walls. As the temperature continues to decrease with increasing $\hat{\eta}$, strings and domain walls tighten and decrease in size until they collapse. By $\hat{\eta} \gtrsim 2.0$, the network collapses in its entirety. Shortly thereafter, we observe the formation of oscillons [92, 96, 126]. We note that the oscillon field configuration is relativistic, so that near the origin of the oscillons the oscillation wavelength is $\sim m_a(\hat{\eta})^{-1}$, which is rapidly shrinking with increasing time. After the zero-temperature mass is reached, oscillons stop shrinking and slowly dissipate at varying rates until the full field enters the linear regime. White regions in the right-most panel of Fig. 2.1 denote regions of high axion energy density, which are mostly inhabited by oscillons.

At the end of the simulation, the field has relaxed into the linear regime (*e.g.*, axion self-interactions are unimportant), but the field remains mildly relativistic because axion radiation is produced during the string-network collapse and during the oscillon collapse. It is therefore important to continue evolving the axion field until a time nearer to matter-radiation equality to allow the field to become non-relativistic everywhere and also to allow the compact but high-momentum overdensities to spread out. We perform this evolution analytically by exactly solving the linear axion equations of motion in Fourier space. We end this evolution shortly before matter-radiation equality ($T \sim$ keV), at which time proper velocities have frozen out but local radiation domination

is preserved at all locations in our simulation box so that gravitational effects remain negligible.

## 2.2 Analysis and Results

We provide Supplementary Data [127] containing the final state from our most realistic QCD-epoch simulation, after having performed the evolution to near matter-radiation equality. Note that the axion field after the QCD phase transition is highly non-Gaussian and phase-correlated at small scales and cannot accurately be reconstructed from the power spectrum. In fact considering that we find large overdensities $\delta$ ($\delta \sim 10$), with $\delta = (\rho - \bar{\rho})/\rho$ and $\bar{\rho}$ ($\rho$) the average (local) DM density, the field could not possibly be Gaussian at these scales, considering that Gaussian random fields have symmetric over and under-densities but under-densities with $\delta < -1$ would have negative DM density.

We may try to estimate the present-day mass function by performing a clustering analysis on the final states. In particular, we expect that the large overdensities will detach from the cosmic expansion, due to reaching locally matter-radiation equality before the rest of the Universe, and collapse onto themselves under gravity. Thus by clustering the 3-D spatial energy density distribution from the simulation slightly before matter-radiation equality and quantifying the distribution of masses and overdensities that we find, we can make predictions for the spectrum of minihalo masses and concentrations today.

From the final-state we construct an overdensity field $\delta(x)$, and we identify overdensities as closed regions of positive $\delta$. Under this definition 50% of the total mass is in overdensities. In practice, we identify these regions by first finding all positive local maxima, then recursively identifying all neighboring grid sites that are larger than 20% of the corresponding local maxima. We assign to each overdensity with at least 80 grid sites a mass $M$ and a single mean concentration parameter $\delta$. The final mass function is not strongly dependent on the specific choice of the 20% and 80 grid sites threshold.

An illustration of our clustering procedure is shown in Fig. 2.2. In that figure we show a 2-dimensional slice through the overdensity field for our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$. Note that in the left panel we show the field at $\hat{\eta} = 7$ at the end of the QCD simulation while in the right panel we show the same slice slightly before matter-radiation equality, denoted by $\hat{\eta}_{\rm MR} = 10^6$ and corresponding to $T \sim$ keV. While a large overdensity left over from oscillon decay, along with corresponding rings of relativistic axion radiation, is visible in the left panel, that structure largely disperses in the subsequent evolution to $\hat{\eta}_{\rm MR}$. Two-dimensional slices through the boundaries of the clustered regions are shown in red in Fig. 2.2.

We characterize the minihalo mass function through the distribution $d^2 f/d(\log M)/d\delta$, where $f$ represents the fraction of mass in overdensities of mean overdensity $\delta$ and mass $M$ with respect to

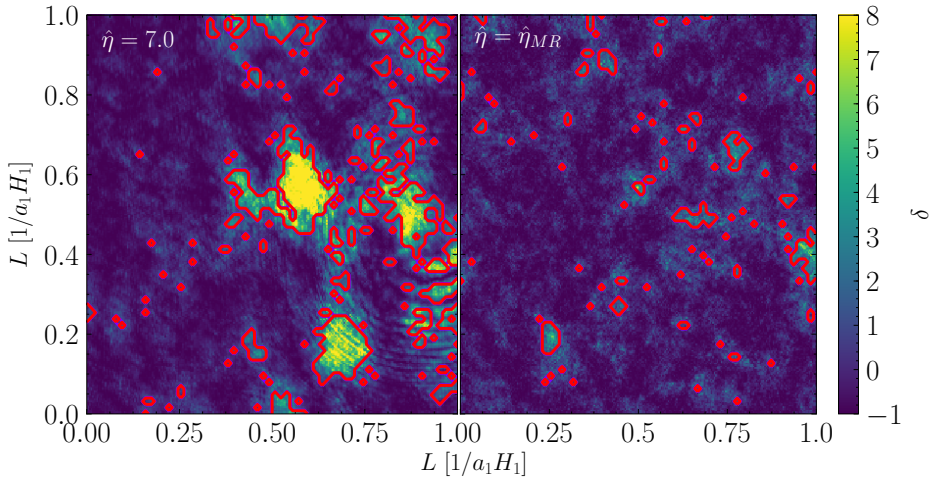



Figure 2.2: (Left) A portion of a 2-D slice through the overdensity field $\delta(x)$ at the end of the QCD stage of our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$. Large overdensities and rings of relativistic radiation arise from oscillon decay. Slices through the clustered minihalos are outlined in red. (Right) As in the left panel, except the field is evolved to matter-radiation equality. The large overdensities largely disperse and the field is everywhere non-relativistic.

the total mass in minihalos. We compute the mass function for all of the 25 simulations at varying $\tilde{\lambda}$ and $\hat{\eta}_c$. To perform the extrapolation to the physical $f_a$ ($\hat{\eta}_c$), we use the following procedure. First, we normalize the total DM density found in the simulation at $\hat{\eta}_{\rm MR}$ to the value that would give the observed DM density today. Then we perform the clustering algorithm to determine $d^2 f/d(\log M)/d\delta$. We rescale all of the masses by $\left[(a_1 H_1)^{\rm sim}/(a_1 H_1)^{\rm target}\right]^3$, where $(a_1 H_1)^{\rm sim}$ is the simulated horizon size at $\hat{\eta} = 1$ and $(a_1 H_1)^{\rm target}$ is the horizon size at the target $f_a$. The shift accounts for the fact that the characteristic scale of the overdensities is expected to be set by the horizon volume when the axion field begins to oscillate (see, *e.g.*, [96, 99] and the appendices). The effect of this shift is to move all of the masses to lower values, since the target $m_a$ is larger than those we simulate. The resulting mass function for our most realistic simulation is shown in Fig. 2.3. As we show in the appendices, after applying the mass shift the mass functions appear to give relatively consistent results between the different $\hat{\eta}_c$, though the agreement is not perfect at high $M$. As a result, we cannot exclude the possibility that simulating to the target $\hat{\eta}_c$ would give different results, especially at high masses, compared to our extrapolations. On the other hand, the effect of $\tilde{\lambda}$ appears to be minimal, since this parameter only affects the decay of the string network. We additionally caution that these minihalo masses have been defined in a non-standard way by calculating the total overdensity mass. For realistic projections associated with minicluster observables, *N*-body simulation using the late-time axion field as an initial state evolving under

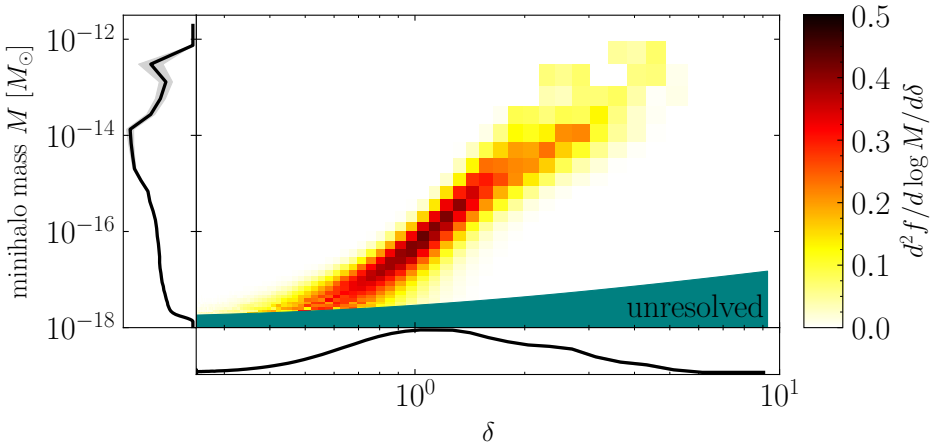



Figure 2.3: Differential mass distribution for axion minihalos for our most realistic simulation, as described in Fig. 2.2, computed by clustering the overdensity field at $\hat{\eta}_{\rm MR}$. The shaded "unresolved" region denotes the parameter space that is beyond our resolution limit. Small statistical uncertainties are displayed as grey error bands.

gravitation is necessary. However, our accounting, which has neglected gravity, is useful in that it provides an estimate of the minimum mass of gravitationally collapsed structure that arises in this scenario.

We may also compare our determinations of the total DM density produced during the QCD phase transition to previous analyses (see *e.g.* [112–119]). Our results are summarized in Fig. 2.4, where we show the DM density today that we find for our top four $\hat{\eta}_c$, converted to $f_a$, for our most physical $\tilde{\lambda}$. The uncertainties in our $\rho_a$ measurements are determined from the variance between the different $\tilde{\lambda}$ simulations, and while some small dependence on $\tilde{\lambda}$ is expected, we find that this dependence is subdominant to statistical noise and no trend is detectable in our data. We also include a conservative $10\%$ systematic uncertainty that accounts for our unphysical fixing of the effective number of degrees of freedom $g_*$ throughout our simulation, a 15% systematic uncertainty from violations to scaling between the PQ and QCD phase transitions, and the uncertainty on the measured value of $\Omega_a$ in our Universe [128] (see the appendices for details).

In Fig. 2.4 we compare our results to the best-fit simulation result from [112], which like us numerically evolved the axion-string system through the QCD phase transition, albeit with a different formalism, and also the semi-analytic calculations from [118]. Our results are in reasonable agreement with those in [112] and significantly disagree with those in [118]. Note that we self-consistently account for all production mechanisms for axion DM in our simulation, including string decay in the few decades before the QCD phase-transition. It is the late-time axion production, right before the QCD phase transition, which is most important since it is the least

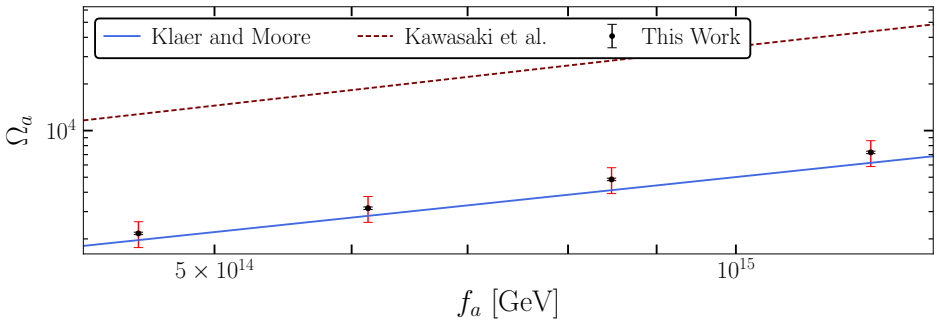



Figure 2.4: The DM density $\Omega_a$ as a function of the axion decay constant $f_a$, with statistical uncertainties (black) and correlated systematic uncertainties (red) indicated, for our top four simulations. We compare our results to those in [112] (Klaer and Moore), which agree relatively well with our own, and [118] (Kawasaki et al.), which predicts significantly higher $\Omega_a$ relative to what we find.

redshifted [118]. The source of the discrepancy could be due in part to the fact that by artificially separating the production mechanisms, [118] over-counted the DM density produced (see [112]). Additionally, the highly non-linear axion dynamics at the QCD epoch likely violate the number-conserving assumptions made by [118].

We may estimate the $f_a$ that gives the correct DM density by fitting our results to a power-law $\Omega_a \sim f_a^\alpha$. We find the best-fit index $\alpha = 1.24 \pm 0.04$, only including statistical uncertainties, which is marginally compatible with the analytic calculations in [112, 118] that predict $\alpha = (n+6)/(n+4) \approx 1.187$. Fixing $\alpha$ to the theoretical value, we find $\Omega_a = (0.102 \pm 0.02) \times (f_a/10^{11}\text{GeV})^{1.187}$, now incorporating the correlated systematic uncertainties, which leads to the prediction that the correct DM density is achieved for $f_a = (2.27 \pm 0.33) \times 10^{11}$ GeV ($m_a = 25.2 \pm 3.6$ $\mu$eV) in agreement with [112]. Note that if we fit for $\alpha$ instead of fixing $\alpha$ to the theoretical value we find $m_a = 17.4 \pm 4.5$ $\mu$eV; the difference between the two $m_a$ estimates could be due to a systematic difference between the theoretically predicted $\alpha$ and the actual dependence of $\Omega_a$ on $f_a$. In light of this we use the difference between the two $m_a$ estimates as an estimate of the systematic uncertainty from the extrapolation to $f_a$ below those simulated. We additionally include a $\sim$27% uncertainty on $m_a$ from uncertainties in the mass growth of the axion (see the appendices for details), leading to the prediction $m_a = 25.2 \pm 11.0$ $\mu$eV.

## 2.3 Conclusion

We performed high-resolution simulations of axion DM in the cosmological scenario where the PQ symmetry is broken after inflation, starting from the epoch before the PQ phase transition and

evolving the field until matter-radiation equality. After matter-radiation equality one should still evolve the axion field gravitationally down to lower redshifts, which we plan to do in future work. Our mass function is an estimate of the resulting mass function one would find after simulating the gravitational collapse. It is possible that the true halos will be slightly larger in mass due to *e.g.* accretion of surrounding DM.

We may try to estimate the halo sizes based upon when we expect the halos to collapse gravitationally. Under the assumption, for example, that the final density profile is a constant-density sphere of radius $R$ (which is likely not a good approximation but still is useful to get a sense of the halo sizes), then the halo density today was argued to be approximately $\rho \approx 140\rho_{\rm eq}\delta^3(\delta+1)$, where $\rho_{\rm eq}$ is the DM density at matter-radiation equality [93]. This implies, for example, that a $M = 10^{-14}\ M_\odot$ subhalo with an initial average overdensity $\delta = 3$ will have a characteristic size of $\sim 1 \times 10^6$ km. The implications for direct and indirect axion detection efforts (*e.g.*, non-trivial time dependence) are likely substantial and will be the subject of future work. One immediate implication, however, is that the axion minihalos are likely out of reach for microlensing and pulsar timing surveys [100], given the small minihalo masses.

# CHAPTER 3

# The Statistics of Axion Direct Detection

The local distribution of dark matter (DM) leaves a unique fingerprint on an emerging signal at axion direct detection experiments. While it has long been recognized that the local phase-space distribution of DM may be partially uncovered with direct-detection experiments searching for heavy DM candidates with masses $m_{\rm DM} \gtrsim$ MeV (for a recent review, see [129]), the role of the DM distribution at axion direct detection experiments, where $m_{\rm DM} \lesssim$ meV, remains less explored. In this work, we develop a likelihood-function-based analysis framework for analyzing the output of axion DM direct detection experiments. Using this framework, we explore in detail the impact of the DM phase-space distribution on the experimental sensitivity to the axion; in the presence of a signal, we show that many aspects of the full time-dependent phase-space distribution can be uncovered.

The need for understanding how the DM phase-space distribution is manifest in axion direct detection experiments has taken on a new sense of urgency recently due to a multitude of new experimental efforts. In addition to the long-running ADMX experiment [107, 130, 131], there has been a raft of new ideas for directly detecting axion DM, including ABRACADABRA [71], CASPEr [58], CULTASK [132], DM Radio [76, 111], MADMAX [70, 133–135], HAYSTAC [108–110], nEDM [136, 137], ORGAN [138], QUAX [139–141], TASTE [142], and more [143–165]. Our statistical framework allows us to better quantify limits and detection thresholds for the proposed experiments. Moreover, it also shows how various features of the DM distribution, for example annual modulation, gravitational focusing, and potential substructure such as local DM streams, can affect the sensitivity of these experiments and how they can be searched for in the data.

The resurgence of effort towards detecting axion DM is driven by a combination of factors, including the increasing tension that heavier DM candidates are facing from null searches, technological advancements that make axion searches more feasible, and new ideas for how to detect axion DM in the laboratory. However, axion DM is also a focus point due to its strong theoretical motivation. The quantum chromodynamics (QCD) axion was originally invoked to solve the strong CP problem of the neutron electric dipole moment [38, 39, 43, 44]. It was later realized that

the QCD axion behaves like cold DM for cosmological and astrophysical purposes [45, 46, 86]. The axion interacts with the electromagnetic sector through the following operator:

$$\mathcal{L}_a = -\frac{1}{4} g_{a\gamma\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu} \,, \tag{3.1}$$

where $F_{\mu\nu}$ is the electromagnetic field strength, $a$ is the axion field, and $g_{a\gamma\gamma}$ is the coupling.[1] We may parametrize the coupling as $g_{a\gamma\gamma} = g\alpha_{\rm EM}/(2\pi f_a)$, where $f_a$ is the axion decay constant, $\alpha_{\rm EM}$ is the electromagnetic fine structure constant, and $g$ is a model dependent parameter, which takes a value $-1.95$ $(0.72)$ for the KSVZ [40, 87] (DFSZ [41, 42]) QCD axion, although the space of models covers an even broader range (see, *e.g.*, [166]). The axion decay constant determines the axion mass through the coupling of the axion to QCD:

$$m_a \approx \frac{f_\pi m_\pi}{f_a} \,, \tag{3.2}$$

which is given in terms of the pion mass and decay constant, $m_\pi$ and $f_\pi$, respectively. Depending on the detailed cosmological scenario, the QCD axion may make up all of the DM for axion masses roughly in the range $\sim 10^{-12}$ eV to $\sim 10^{-5}$ eV (see [89] for a review). Lower masses are disfavored by requiring the axion decay constant, which is the scale of new physics that generates the axion, to be sub-Planckian. At higher masses it becomes more difficult to generate the required abundance of DM through the misalignment mechanism and the decay of topological defects (see, *e.g.*, [167]). In addition to the QCD axion, it is also possible to have more general axion-like DM particles that still couple to electromagnetism, but not to QCD. The mass of these axion-like particles is a free parameter, since there is no contribution from QCD; however, axion-like particles do not address the strong CP problem.

Most axion direct detection experiments exploit the fact that axion DM may be described by a coherently-oscillating classical field $a$ that acts as a source of $F_{\mu\nu}\tilde{F}^{\mu\nu}$. The oscillation frequency of $a$ is set by its mass $m_a$, while the coherence of the oscillations is set by the local DM velocity distribution. Locally, we expect the velocity dispersion of the bulk DM halo to be $\sim 10^{-3}$ in natural units, which leads to the expectation that the axion coherence time is $\tau \sim 10^6 \times (2\pi/m_a)$. Consequently, the axion sources a coherent signal that experiments can repeatedly sample by taking time-series data sensitive to the possible interactions of the axion. For example, in ADMX, which is the only experiment so far to constrain part of the QCD axion parameter space,[2] the coherent

[1] Throughout this work we will consider exclusively the electromagnetic coupling, but the framework we introduce can be straightforwardly extended to nucleon couplings.

[2] This, of course, depends on the exact definition of what constitutes a QCD axion. Recent studies have suggested the window could be broader than what we discuss in this work, see, *e.g.*, [168, 169]. Under such extended definitions, results from the HAYSTAC experiment may already probe the QCD parameter space [108].

axion background sources electromagnetic modes in a resonant cavity. The experiment tunes the resonant frequency of the cavity to scan over different possible masses. Most axion experiments make use of high-$Q$ oscillators or cavities to build up the otherwise small signal. However, some experiments, such as ABRACADABRA and MADMAX, can operate in a broadband mode that allows multiple masses to be searched for simultaneously, albeit with slightly reduced sensitivity.

Resonant experiments, such as ADMX, typically analyze their data by comparing the power output from the resonator, measured across the frequency bandwidth of the signal as determined by the coherence time, to the expectation under the null hypothesis using, for example, the Dicke radiometer equation [170], supplemented with Monte Carlo simulations as described in [130, 171]. In this work, we present a likelihood-function based approach to analyzing the data at resonant and broadband axion experiments that takes as input the Fourier components of the time-series data, with frequency spacing potentially much smaller than the bandwidth of the signal. We show that the velocity distribution of the local halo is uniquely encoded in the spectral shape of the Fourier components, within the frequency range set by the coherence time, and that it may be extracted from the data in the event of a detection.

We present an analytic analysis of the likelihood function using the Asimov dataset [172], which also allows us to calculate the sensitivity of axion experiments to DM substructure such as cold DM streams and a co-rotating dark disk. For example, we show that soon after the discovery of axion DM from the bulk DM halo, the DM component of the Sagittarius stream, which has been extensively discussed in the context of electroweak-scale direct detection [173–176], should become visible in the data through the likelihood analysis. Moreover, we may use the formalism to accurately predict exclusion and discovery regions analytically.

Most previous studies of axion direct detection have not addressed the question of how to extract measures of the local phase-space distribution from the data. In [177], it was demonstrated that effects of the non-zero axion velocity will need to be accounted for in future versions of the MADMAX experiment. Ref. [178] recently performed simulations to show how the sensitivity of ADMX changes for different assumptions about the velocity distribution, such as the possibility of a co-rotating dark disk or cold flows from late infall, using the analysis method used by ADMX in previous searches (see, for example, [179, 180]). In [181] (see also [182]) it was pointed out that the width of the resonance should modulate annually due to the motion of the Earth around the Sun, which slightly shifts the DM velocity distribution. Recently, [183] took an approach similar to that presented in this work and considered a likelihood-based approach to annual modulation and reconstructing the halo velocity distribution. We extend this approach to accurately account for the statistics of the axion field, to include previously-neglected but important phenomena such as gravitational focusing [184] induced by the Sun's gravitational potential, and to analytically understand, using the Asimov formalism [172], the effect of DM substructure.

## 3.1 Organization of this Chapter

We organize the remainder of this work as follows. To begin with, in Sec. 3.2 we derive a likelihood for axion direct detection. The result is derived for both broadband and resonant experimental configurations. Section 3.3 determines the expected limit and detection thresholds from this likelihood. In Sec. 3.4 we discuss our results in the context of an axion population following a time independent bulk halo. Finally, Sec. 3.5 extends the discussion of the axion phase space to include annual modulation, gravitational focusing, and the possibility of local DM substructure such as cold streams. We note that the analysis framework presented in this work is also provided in the form of publicly available code and can be accessed at `https://github.com/bsafdi/AxiScan`.

## 3.2 A Likelihood for Axion Direct Detection

In this section we derive a likelihood that describes how the statistics of the local DM velocity distribution are transformed into signals at axion direct detection experiments. The main result that will be used throughout the rest of the paper is the likelihood presented in (3.29); however, there will be several intermediate steps. In particular, in the first subsection we show how to write the local axion field as a sum over Rayleigh-distributed random variables, as specified in (3.10). In the following subsection we will show that when coupled to an experiment sensitive to the axion, if data is taken in the form of a power spectral density (PSD), it will be exponentially distributed, as given in (3.24). In the main body we will only derive the distribution of the signal, but in App. B.1 we will show that the background only, and signal plus background distributions, are both exponentially distributed also. Combining these, we then arrive at a form for the likelihood function.

In the initial derivation of the likelihood we will focus on how our formalism applies to a broadband experiment. However, the modification to a resonant framework is straightforward and we present the details in the final subsection.

### 3.2.1 The Statistics of the Local Axion Field

Our goal in this section is to build up the local axion field from the underlying distribution of fields describing individual axions. Thus as a starting point let us consider an individual axion-like particle, which we think of as a non-relativistic classical field.[3] If we assume that there are $N_a$

[3]Individual axion-like particles should technically be described as quantum objects not classical fields. Nevertheless the local occupancy numbers of these quantum particles is enormous. For example, taking axion dark matter with $m_a \sim 10^{-10}$ eV, the number of axions within a de Broglie volume is $\sim 10^{36}$. Accordingly the distinction is unimportant since formally when we say single particles we really mean a collection of particles in the same state with high

such particles locally that make up the local DM density $\rho_{\rm DM}$, then we can write down the field describing an individual particle as

$$a_i(v,t) = \frac{\sqrt{2\rho_{\rm DM}/N_a}}{m_a}\cos\left[m_a\left(1+\frac{v_i^2}{2}\right)t+\phi_i\right], \tag{3.3}$$

where $i \in 1, 2, \ldots, N_a$ is an index that identifies this specific axion particle, $m_a$ is the axion mass, $v_i$ is the velocity of this axion, and $\phi_i \in [0, 2\pi)$ is a random phase. The phase coherence of the full axion field constructed from the sum each of these particles is dominated by the common mass they share and to a lesser extent by velocity corrections which are drawn from a common DM velocity distribution. Beyond this we take the fields to be entirely uncorrelated, which is represented by the random phase. Axion self interactions could induce additional coherence. However, given the feeble expected strength of these interactions we assume such contributions are far subdominant to those written.

From here to build up the full axion distribution we need to sum (3.3) over all $i$. We proceed, though, through an intermediate step that takes advantage of the fact that there will be many particles with effectively indistinguishable speeds. As such let us partition the full list of $N_a$ particles into subsets $\Omega_j$, which contain the $N_a^j$ particles with speeds between $v_j$ and $v_j + \Delta v$, where $\Delta v$ is small enough that we can ignore the difference between their speeds. In this way the contribution from all particles in subset $\Omega_j$ is given by

$$a_j(t) = \sum_{i\in\Omega_j}\frac{\sqrt{2\rho_{\rm DM}}}{m_a\sqrt{N_a}}\cos\left[m_a\left(1+\frac{v_j^2}{2}\right)t+\phi_i\right]. \tag{3.4}$$

Note that it is only the random phase that differs between elements of the sum:

$$\begin{aligned}&\sum_{i\in\Omega_j}\cos\left[m_a\left(1+\frac{v_j^2}{2}\right)t+\phi_i\right]\\ =&\mathrm{Re}\left\{\exp\left[im_a\left(1+\frac{v_j^2}{2}\right)t\right]\left(\sum_{i\in\Omega_j}\exp\left[i\phi_i\right]\right)\right\}.\end{aligned} \tag{3.5}$$

To proceed further, we recognize that the sum over phases is equivalent to a 2-dimensional random walk; this allows us to write

$$\sum_{i\in\Omega_j}\exp\left[i\phi_i\right] = \alpha_j e^{i\phi_j}, \tag{3.6}$$

where $\phi_j \in [0, 2\pi)$ is again a random phase and $\alpha_j$ is a random number describing the root-

enough occupancy number such that the ensemble is described by a classical wave. For simplicity, however, we refer to these classical building blocks as "particles."

mean-squared distance traversed in a 2-dimensional random walk of $N_a^j$ steps. These distances are governed by the Rayleigh distribution, which takes the form

$$P[\alpha_j] = \frac{2\alpha_j}{N_a^j} e^{-\alpha_j^2/N_a^j} \,. \tag{3.7}$$

For future convenience, we remove $N_a^j$ from the distribution by rescaling $\alpha_j \to \alpha_j \sqrt{N_a^j/2}$, so that we can complete our result for this velocity component as follows:

$$\begin{aligned} a_j(t) &= \alpha_j \frac{\sqrt{\rho_{\rm DM}}}{m_a} \sqrt{\frac{N_a^j}{N_a}} \cos\left[ m_a \left(1 + \frac{v_j^2}{2}\right) t + \phi_j \right] \,, \\ P[\alpha_j] &= \alpha_j e^{-\alpha_j^2/2} \,. \end{aligned} \tag{3.8}$$

The final step to obtain the full local axion field is to sum over all $j$. Before doing so, however, we note the important fact that the speeds, $v_j$, are being drawn from the local DM speed distribution, $f(v)$. A simple ansatz for $f(v)$ is given by the standard halo model (SHM):[4]

$$\begin{aligned} f_{\rm SHM}(v|v_0, v_{\rm obs}) = &\frac{v}{\sqrt{\pi} v_0 v_{\rm obs}} e^{-(v+v_{\rm obs})^2/v_0^2} \\ &\times \left( e^{4v v_{\rm obs}/v_0^2} - 1 \right) \,, \end{aligned} \tag{3.9}$$

where in conventional units $v_0 \approx 220$ km/s is the speed of the local rotation curve, and $v_{\rm obs} \approx 232$ km/s is the speed of the Sun relative to the halo rest frame.[5] As shown in Sec. 3.5, small variations on this simple model can induce large changes to the expected experimental sensitivity, but $f_{\rm SHM}(v)$ is likely to approximately describe the bulk of the local DM speed distribution and so gives a good initial proxy for $f(v)$. As a first use of $f(v)$, we can rewrite $N_a^j$ in terms of $f(v)$, as from the definition of $j$ we have $N_a^j = N_a f(v_j) \Delta v$. With this we arrive at the main goal of this section, a form for the local axion distribution:

$$\begin{aligned} a(t) = &\frac{\sqrt{\rho_{\rm DM}}}{m_a} \sum_j \alpha_j \sqrt{f(v_j) \Delta v} \\ &\times \cos\left[ m_a \left(1 + \frac{v_j^2}{2}\right) t + \phi_j \right] \,, \end{aligned} \tag{3.10}$$

where note the sum over $j$ is effectively a sum over velocities, and again we emphasize that each

[4] We note in passing that data from the *Gaia* satellite is likely to lead to updates to this simple model [185, 186]. Further, there is also likely a cut-off at the Galactic escape velocity, $\sim$550 km/s, though this will not play an important role in the analyses in this work.

[5] When manipulating the velocity distribution, we will often work in natural units.

$\alpha_j$ is a random number drawn from the distribution given in (3.8).

### 3.2.2 Coupling the Axion to a Broadband Experiment

We now discuss how to quantify the coupling of the DM axion field to an experiment sensitive to the coupling in (3.1), using the form of the local axion field given in (3.10). Then, we write down a likelihood function that may be used to describe the experimental data. Here we focus on determining the statistics of the signal alone; combining the signal with background is straightforward and described in more detail in App. B.1. To make the discussion concrete, we frame the problem in the context of the recently proposed ABRACADABRA experiment [71], operating in the broadband readout mode. We emphasize, however, that the results we derive are much more general and are applicable to any experiment which seeks to measure time-series data based upon the local axion field. An example of this generality is provided in the next section, where we extend the formalism to the resonant case.

Let us briefly review the operation of ABRACADABRA, a 10-cm version of which is currently under development [187]. This experiment exploits the fact that the coupling between the axion and QED, given by the operator in (3.1), induces the following modification to Ampère's circuital law

$$\nabla \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} - g_{a\gamma\gamma}\left(\mathbf{E} \times \nabla a - \mathbf{B}\frac{\partial a}{\partial t}\right) . \tag{3.11}$$

The final term in this equation implies that in the presence of a magnetic field and axion DM, there is an effective current induced that follows the primary laboratory magnetic field lines and oscillates at the axion frequency. ABRACADABRA sources this effective current via a toroidal magnet, which generates a large static magnetic field. The axion then generates an oscillating current parallel to the magnetic field lines, which in turn sources an oscillating magnetic flux through the center of the torus. By placing a pickup loop in the center of the torus, this oscillating magnetic field will induce an oscillating magnetic flux of the form

$$\Phi_{\rm pickup}(t) = g_{a\gamma\gamma} B_{\rm max} V_B m_a a(t) \,, \tag{3.12}$$

where $B_{\rm max}$ is the magnetic field at the inner radius of the torus, and $V_B$ is a factor that accounts for the geometry of the toroidal magnet and pickup loop and has units of m$^3$. In the broadband configuration, the pickup loop, which is taken to have inductance $L_p$, is inductively coupled to a DC SQUID magnetometer of inductance $L$, which will then see a magnetic flux of

$$\Phi_{\rm SQUID} \approx \frac{\alpha}{2}\sqrt{\frac{L}{L_p}}\Phi_{\rm pickup} \,, \tag{3.13}$$

where $\alpha$ is an $\mathcal{O}(1)$ number characterizing how the SQUID geometry impacts the mutual inductance of the SQUID and pickup loop circuit. A typical value we will use in calculations is $\alpha = 1/\sqrt{2}$. The coupling will also induce a frequency independent phase difference between the pickup loop and magnetometer fluxes, but as we show below such an overall phase will not contribute to the measured PSD and so we do not keep track of it.

In this way, through repeated measurements of the magnetic flux detected by the SQUID, ABRACADABRA is able to build up a time series of data proportional to the local axion field. If the experiment is sampling the magnetic flux at a frequency $f$ over a time period $T$, then it will collect a total of $N = f\,T$ data points separated by a time spacing $\Delta t = 1/f$. Storing all of the experimental data may pose a challenge.[6] In Sec. 3.3 we will introduce a stacking procedure to cut down on the amount of stored data while maintaining the same level of sensitivity, but for now we will put this issue aside and assume that all the data is stored and analyzed. Combining (3.10), (3.12), and (3.13), we find that

$$\Phi_n = \sqrt{A}\sum_j \alpha_j \sqrt{f(v_j)\Delta v} \times \cos\left[m_a\left(1+\frac{v_j^2}{2}\right) n\Delta t + \phi_j\right], \tag{3.14}$$

where $n \in 0, 1, \ldots, N-1$ indexes the measurement at time $t = n\Delta t$, and for future convenience we have defined

$$A \equiv \frac{\alpha^2}{4}\frac{L}{L_p} g_{a\gamma\gamma}^2 B_{\rm max}^2 V_B^2 \rho_{\rm DM}\,. \tag{3.15}$$

$A$ is proportional to the terms that dictate the size of the axion signal in the experiment, and the specific form here is peculiar to ABRACADABRA. We note that $A$ caries the SI units of $\mathrm{Wb}^2$, which conveniently makes it dimensionless in natural units.

To pick the axion signal out of this time-series data, given the signal is oscillating almost at a specific frequency $m_a$ plus small corrections coming from the velocity components, it is convenient to instead consider the discrete Fourier transform of the data:

$$\Phi_k = \sum_{n=0}^{N-1} \Phi_n e^{-i2\pi kn/N}\,, \tag{3.16}$$

where now $k \in 0, 1, \ldots, N-1$. In practice it is more useful to work with the PSD of the magnetic flux, given by

$$S_{\Phi\Phi}^k = \frac{(\Delta t)^2}{T}\left|\Phi_k\right|^2 = A\frac{(\Delta t)^2}{T}\left|\sum_{n=0}^{N-1}\sum_j \alpha_j\sqrt{f(v_j)\Delta v}\cos\left[\omega_j n\Delta t + \phi_j\right] e^{-i2\pi kn/N}\right|^2. \tag{3.17}$$

[6] To quantify this, if we take the realistic values of $f = 100$ MHz and $T = 1$ year, this amounts to almost 13 PB of data.

Note that in the second equality we defined $\omega_j \equiv m_a\left(1+v_j^2/2\right)$. For the moment, it is helpful to rewrite the PSD as a function of the angular frequency $\omega$, which we can do by noting that $k = \omega T/(2\pi) = \omega \Delta t N/(2\pi)$, giving

$$S_{\Phi\Phi}(\omega) = A\left|\sum_j \alpha_j \sqrt{\frac{f(v_j)\Delta v}{T}}\Delta t \sum_{n=0}^{N-1}\cos\left[\omega_j n\Delta t+\phi_j\right]e^{-i\omega n\Delta t}\right|^2 . \tag{3.18}$$

Our experimental resolution to frequency differences is dictated by the time the experiment is run for, specifically $\Delta f = 1/T$. Then, given the definition of $\omega_j$, for large enough $T$ we have approximately $1/T \approx m_a v_j \Delta v/(2\pi)$, and so

$$S_{\Phi\Phi}(\omega) = A\left|\sum_j \Delta v\,\alpha_j \sqrt{\frac{f(v_j)m_a v_j}{2\pi}}\Delta t \sum_{n=0}^{N-1}\cos\left[\omega_j n\Delta t+\phi_j\right]e^{-i\omega n\Delta t}\right|^2 . \tag{3.19}$$

In a realistic experimental run, $T$ will usually be much larger than any other time scale in the problem considered so far. Exceptions to this occur when there are ultra-coherent features in the dark matter distribution, which we discuss in detail in Sec. 3.5. Putting the exceptions aside for now, we can approximate $T \to \infty$, which means we can also treat $\Delta v \to dv$, $\Delta t \to dt$, and replace the sum over $j$ with an integral over $v$ as follows:

$$S_{\Phi\Phi}(\omega) \approx A\left|\int dv\,\alpha_v \sqrt{\frac{f(v)m_a v}{2\pi}}dt \sum_{n=0}^{N-1}\cos\left[\omega_v n dt+\phi_v\right]e^{-i\omega n dt}\right|^2 . \tag{3.20}$$

Note in the above result we have a subscript $v$ on $\alpha_v$ and $\phi_v$, indicating that for every value of $v$ in the integral we have a different random draw of these numbers.

At this point, to make further progress we focus specifically on the sum over $n$ in the second line above. In detail,

$$\begin{aligned}
&dt\sum_{n=0}^{N-1}\cos\left[\omega_v n dt+\phi_v\right]e^{-i\omega n dt}\\
=&\frac{dt}{2}\left\{e^{i\phi_v}\frac{1-\exp\left[i\left(\omega_v-\omega\right)T\right]}{1-\exp\left[i\left(\omega_v-\omega\right)dt\right]}\right.\\
&\left.+e^{-i\phi_v}\frac{1-\exp\left[-i\left(\omega_v+\omega\right)T\right]}{1-\exp\left[-i\left(\omega_v+\omega\right)dt\right]}\right\}\\
\approx&\frac{e^{i(\phi_v+(\omega_v-\omega)T/2)}}{2}\left\{\frac{\sin\left[\frac{1}{2}(\omega_v-\omega)T\right]}{\frac{1}{2}(\omega_v-\omega)}\right.\\
&\left.+e^{-i(2\phi_v+\omega_v T)}\frac{\sin\left[\frac{1}{2}(\omega_v+\omega)T\right]}{\frac{1}{2}(\omega_v+\omega)}\right\},
\end{aligned} \tag{3.21}$$

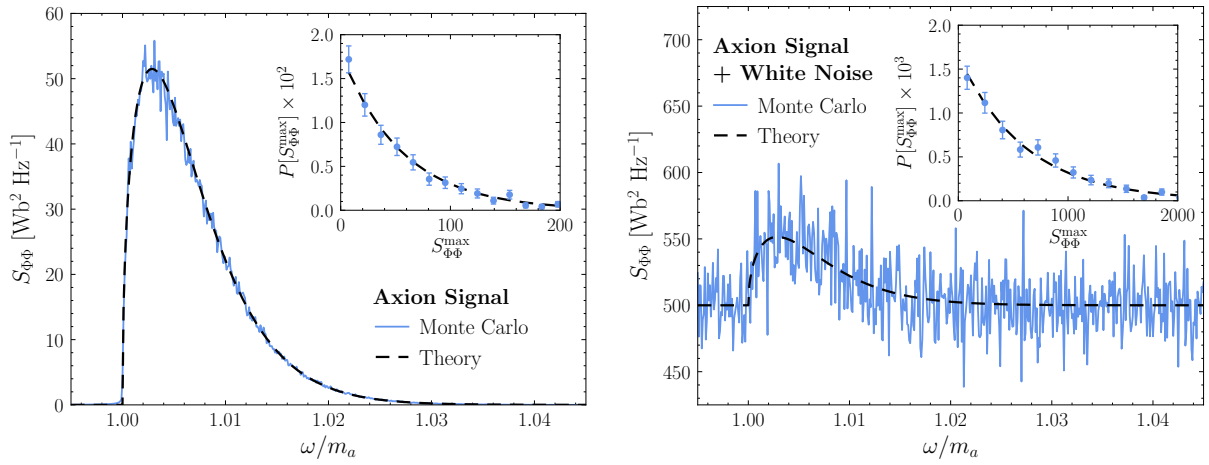



Figure 3.1: (Left) A comparison between the mean of 500 Monte Carlo simulations of a signal only PSD dataset (blue) and the analytic expectation given in (3.26) (black). The inset shows the distribution of the 500 simulated $S_{\Phi\Phi}$ versus the predicted exponential distribution, as in (3.24), at the frequency where the signal distribution is maximized, $\omega/m_a \approx 1.003$. This example was generated assuming the unphysical but illustrative parameters $A = 1$ Wb$^2$, $m_a = 2\pi$ Hz, and $v_0 = v_{\rm obs} = 220{,}000$ km/s. Importantly the simulations were generated by constructing the full axion field starting from (3.3), and so the agreement between theory and Monte Carlo is a non-trivial confirmation of the framework. (Right) As on the left, but with Gaussian distributed white noise added into the time-series data with variance $\lambda_B/\Delta t$, and taking $\lambda_B = 500$ Wb$^2$ Hz$^{-1}$. Again we see the theory prediction in good agreement with the average data, whilst at an individual frequency point the simulated data is exponentially distributed. See text for details.

where in the final step we expanded using $(\omega_v \pm \omega)dt \ll 1$. Then, taking the $(\omega_v \pm \omega)T \to \infty$ limit we can use the result that $\lim_{\epsilon\to 0} \sin(x/\epsilon)/x = \pi\delta(x)$ to rewrite the terms in angled brackets in terms of Dirac-$\delta$ functions which we can use to perform the integral over speeds. There are terms associated with both positive and negative frequencies, but as we have $\omega_v > 0$ we only keep the positive result, and so conclude:

$$dt \sum_{n=0}^{N-1} \cos\left[\omega_v n dt + \phi_v\right] e^{-i\omega n dt} \approx \pi e^{i(\phi_v + (\omega_v - \omega)T/2)} \delta(\omega_v - \omega)\,. \tag{3.22}$$

With the above arguments we may perform the velocity integral in (3.20), obtaining

$$S_{\Phi\Phi}(\omega) = A\,\frac{\pi f(v)}{2m_a v}\alpha^2\Big|_{v=\sqrt{2\omega/m_a - 2}}\,. \tag{3.23}$$

Note that $\omega \approx m_a$, up to corrections that are $\mathcal{O}(v^2)$; where the distinction is not important, we write $m_a$ instead of $\omega$, as in the denominator above. Further, in (3.23) we have dropped the subscript $v$ from $\alpha$, as it is just a single Rayleigh distributed number as given in (3.8). Since $\alpha^2$ is exponentially

distributed, this then implies that the PSD is also exponentially distributed:

$$\begin{aligned} P\left[S_{\Phi\Phi}(\omega)\right] &= \frac{1}{\lambda(\omega)} e^{-S_{\Phi\Phi}(\omega)/\lambda(\omega)} \,, \\ \lambda(\omega) \equiv \langle S_{\Phi\Phi}(\omega)\rangle &= A \left.\frac{\pi f(v)}{m_a v}\right|_{v=\sqrt{2\omega/m_a - 2}} \,. \end{aligned} \tag{3.24}$$

Recall that $A$, which is effectively dictating the strength of the axion signal, has units of $\mathrm{Wb}^2$, so $S_{\Phi\Phi}$ carries units $\mathrm{Wb}^2/\mathrm{Hz}$, or in natural units $\mathrm{eV}^{-1}$.

In any real experiment there will also be background sources of noise in the dataset. For most sources we can think of this as mean zero Gaussian distributed noise in the time domain.[7] For example, in ABRACADABRA the main background sources are expected to be noise within the SQUID for the broadband configuration or thermal noise in the resonant circuit [71]. Both of these are well described by normally-distributed noise sources, and so they fall under this class of backgrounds. In ADMX the dominant background is also thermal noise, and the Gaussian nature of this source has been discussed in Refs. [188, 189]; indeed, in [189] they noted the power due to thermal noise in the experiment should be exponentially distributed. It is likely that most other noise sources will also be normally distributed. However, it may well be possible that certain axion direct detection experiments do suffer from background sources that are not well described by Gaussian noise. In such a case the framework we present in this work will not go through directly, but the same logic can be used to derive a new likelihood that accounts for the specific background distribution. Restricting ourselves to the Gaussian approximation, then, as demonstrated in App. B.1, if we have a series of Gaussian distributed backgrounds of variance $\lambda_B^i/\Delta t$, where $i$ indexes the various backgrounds, then the PSD formed from the combinations of all these will again be exponentially distributed with mean

$$\langle S_{\Phi\Phi}^{\mathrm{bkg}}(\omega)\rangle = \lambda_B \equiv \sum_i \lambda_B^i \,. \tag{3.25}$$

It is important to note that in general $\lambda_B$ will be a function of $\omega$, reflecting an underlying time variation in the backgrounds.

Given that the individual signal and background only cases are exponentially distributed, it is perhaps not surprising that the combined signal plus background is exponentially distributed also. This fact is demonstrated in App. B.1, however we point out here that the correct way to think about this is that the two are combined at the level of the time-series data, not at the level of the

[7] If the mean of the background distribution is non-zero, then this will only impact the $k = 0$ mode of the PSD. For reasons discussed in App. B.1, we will not include this mode in our likelihood, and as such we are only sensitive to the variance of the distributions, and so can choose them to have mean zero without loss of generality.

PSD. To highlight this, the sum of two exponential distributions is not another exponential. Taking this fact, we arrive at the result that the full PSD will be exponentially distributed, with mean

$$\lambda(\omega) = A \left. \frac{\pi f(v)}{m_a v} \right|_{v=\sqrt{2\omega/m_a - 2}} + \lambda_B \,. \tag{3.26}$$

As noted above, in the broadband mode noise within the SQUID magnetometer is expected to be the dominant source of background for ABRACADABRA, making it a useful example to keep in mind. At high frequencies this noise source becomes frequency independent, with magnitude:

$$\sqrt{\lambda_B} \sim 10^{-6} \Phi_0 / \sqrt{\text{Hz}} \,, \tag{3.27}$$

which is written in terms of the flux quantum, $\Phi_0 = h/(2e) \approx 2.1 \times 10^{-15}$ Wb. As such the typical value for the background is

$$\lambda_B \approx 4.4 \times 10^{-42} \text{ Wb}^2 \text{ Hz}^{-1} = 1.6 \times 10^5 \text{ eV}^{-1} \,. \tag{3.28}$$

With this example in mind, we will often assume we have a frequency independent background in our analysis to simplify results, but the formalism can in general account for an arbitrary dependence. Despite this we note that in a real DC SQUID, there will also be a contribution to the noise scaling as $1/f$, that should dominate below $\sim 50$ Hz. We refer to [71] for a more detailed discussion of these backgrounds.

To demonstrate how mock datasets compare to the theoretical expectations derived above, in Fig. 3.1 we show the comparison directly, with (right) and without (left) background noise. In both cases we show the PSD as a function of frequency averaged over 500 realization of the simulated data. In the main figures we see that the frequency dependence of the mean of the signal only and signal plus background distributions, constructed from the simulations, are well described by the analytic relation in (3.26). The insets demonstrate that at a given frequency the simulated data is exponentially distributed in both cases, as predicted by (3.24). The agreement is a non-trivial check of the validity of the framework. We emphasize that the Monte Carlo simulations are constructed in the time domain using (3.3) in the signal case and by drawing mean zero Gaussian noise with variance $\lambda_B/\Delta t$ for the background at each time step. To generate these results we picked numerically convenient rather than physically realistic values. Specifically we used $A = 1$ Wb$^2$, $m_a = 2\pi$ Hz, $\lambda_B = 500$ Wb$^2$ Hz$^{-1}$, and we assumed the signal was drawn from an SHM as given in (3.9), but with $v_0 = v_{\text{obs}} = 220{,}000$ km/s instead of the physical values. However, we emphasize that these values were chosen for presentation purposes only and that we have explicitly verified that the formalism above is also valid for more realistic signal and background parameters.

Knowing how the data is distributed means we can now write down a likelihood function to constrain a signal and background model $\mathcal{M}$, with model parameters $\boldsymbol{\theta}$, for a given dataset $d$. The dataset is given, in the case of ABRACADABRA, by $N$ measurements of the magnetic flux in the SQUID at time intervals $\Delta t$. This data is then converted into a PSD distribution $S^k_{\Phi\Phi}$, measured at $N$ frequencies given by $\omega = 2\pi k/T$, for $k \in 0, 1, \ldots, N-1$. The likelihood function for the model $\mathcal{M}$ then takes the form[8]

$$\mathcal{L}(d|\mathcal{M}, \boldsymbol{\theta}) = \prod_{k=1}^{N-1} \frac{1}{\lambda_k(\boldsymbol{\theta})} e^{-S^k_{\Phi\Phi}/\lambda_k(\boldsymbol{\theta})} \,, \tag{3.29}$$

where we have used an index $k$ to denote quantities evaluated at a frequency $\omega = 2\pi k/T$. Note that the $\boldsymbol{\theta}$ completely specify the model expectation given in (3.26). Specifically, $\boldsymbol{\theta}$ includes parameters controlling the background contribution in $\lambda_B$, the DM halo velocity distribution $f(v)$, and the axion coupling $g_{a\gamma\gamma}$ that appears in $A$. We have boxed this expression to emphasize its outsized importance within the context of this work. All results subsequently derived will represent particular evaluations of the likelihood (or ratios of the likelihood) under assumptions about the experimental sensitivity parameter $A$ or the local DM velocity distribution $f(v)$. In the following section, we will show how to use this likelihood to set a limit on or claim a discovery of the axion, as well as constrain properties of the axion velocity distribution in the event of a detection. First, however, we describe how the formalism above is modified for a resonant readout.

### 3.2.3 Coupling to a Resonant Experiment

The discussion above was premised upon a broadband experimental set up. The broadband circuit has the advantage of being able to search across a broad range of axion masses with the same dataset. A common alternative is the resonant framework, where the resonant frequency is tuned to the axion mass under consideration before reading out the signal [190]. Resonant experiments provide increased sensitivity at the frequencies under consideration. The resonators may include physical resonators, such as that used by the ADMX experiment, or resonant circuits as used, for example, in Ref. [147].

In this section we demonstrate how the framework above is modified in these cases, and importantly will find that the same likelihood function applies, with a simple modification to the

[8]The omission of $k = 0$ from the likelihood is deliberate. As described in App. B.1, the background is in fact not exponentially distributed for this value. In addition the signal cannot contribute to the $k = 0$ mode, as this would correspond to probing the velocity distribution at an imaginary value. As such the $k = 0$ or DC mode is only probing a constant contribution to the background, which we can always simply set to be zero and neglected, implying that we lose no sensitivity by simply excluding this case. Moreover, in practice it is likely only necessary to include $k$ modes corresponding to frequencies in the vicinity of the mass under question.

expected PSD given in (3.26). As a consequence, this will show that the various applications of the likelihood framework that we demonstrate throughout the rest of this work are applicable to resonant experiments, even though our examples will generally be couched in the language of a broadband framework for simplicity.

To avoid the discussion becoming too abstract, we will again work with the concrete set up of ABRACADABRA, this time in the resonant mode. We assume, for simplicity, a simple resonant circuit, where the pickup loop is connected to an RLC circuit that is inductively coupled to the SQUID, though more complicated circuits, such as feedback damping circuits [111, 191, 192], may be preferable in practice [71]. However, the analysis formalism described below should apply to any resonant circuit where thermal noise is the dominant noise source.

Our starting point is the magnetic flux due to the axion through the pickup loop, $\Phi_{\rm pickup}$, as given in (3.12). Instead of directly inductively coupling the pickup loop to the SQUID, this time we run the pickup loop through an RLC circuit with inductance $L_i$, resonant frequency $\omega_0$, and quality factor $Q_0$. The strategy is to vary $\omega_0$ over time in order to probe a range of axion masses; we will discuss a strategy for how to choose the time variation later in this work. Note that the quality factor also determines the bandwidth of the circuit, and so choosing a $Q_0$ corresponding to the width of the signal or better is preferable, though we leave a detailed optimization of the resonant strategy to future work. If we inductively couple this circuit directly to the SQUID, then the flux received will be

$$\Phi_{\rm SQUID} = \alpha Q_0 \sqrt{\mathcal{T}(m_a)} \frac{\sqrt{L L_i}}{L_T} \Phi_{\rm pickup} \,, \tag{3.30}$$

where we ignore constant phase shifts. Note that we have defined the total inductance of the pickup loop and the RLC circuit as $L_T \equiv L_i + L_p$ and also a transfer function for the RLC circuit:

$$\mathcal{T}(\omega) \equiv \frac{1}{\left(1 - \omega_0^2/\omega^2\right)^2 Q_0^2 + \omega_0^2/\omega^2} \,. \tag{3.31}$$

Following through the same steps as in the broadband case, we find that now our expected signal PSD is

$$\begin{aligned} \lambda^{\rm res}(\omega) &= A^{\rm res} Q_0^2 \mathcal{T}(\omega) \frac{\pi f(v)}{m_a v} \,, \\ A^{\rm res} &\equiv \alpha^2 \frac{L L_i}{L_T^2} g_{a\gamma\gamma}^2 B_{\rm max}^2 V_B^2 \rho_{\rm DM} \,, \end{aligned} \tag{3.32}$$

where again velocities are evaluated at $v = \sqrt{2\omega/m_a - 2}$. Comparing the expected resonant signal PSD, $\lambda^{\rm res}(\omega)$, with the expected broadband result, $\lambda(\omega)$ given in (3.24), we see that other than the additional frequency dependence in $\mathcal{T}(\omega)$ the two only differ in experimental prefactors.

In the resonant case we also need to rethink what constitutes the dominant background source. In particular, the addition of a resistor in the RLC circuit will generate a new source of background:

Johnson–Nyquist or thermal noise. This background is again expected to be normally distributed, with a variance $\lambda_B^{\rm therm}/\Delta t$ and

$$\lambda_B^{\rm therm}(\omega) = 2\alpha^2 k_b T \frac{L L_i}{L_T}\frac{\omega_0}{\omega^2} Q_0 \mathcal{T}(\omega)\,, \tag{3.33}$$

where $T$ is in this context the temperature of the circuit. At the resonance frequency, for typical values of the parameters of interest, it may be verified that thermal noise dominates the intrinsic noise in the SQUID [71, 111]. Accordingly, we neglect the background from the SQUID noise, and our full resonant model prediction is given by:[9]

$$\begin{aligned}\lambda^{\rm res}(\omega) &= \left[A^{\rm res} Q_0 \frac{\pi f(v)}{m_a v} + \tilde{\lambda}_B^{\rm therm}(\omega)\right] Q_0 \mathcal{T}(\omega)\,,\\ \tilde{\lambda}_B^{\rm therm}(\omega) &\equiv 2\alpha^2 k_b T \frac{L L_i}{L_T}\frac{\omega_0}{\omega^2}\,.\end{aligned} \tag{3.34}$$

As we will see below, the fact that the transfer function is common to both the signal and background will mean its dependence vanishes when computing our experimental sensitivity. This point will be demonstrated in the next section.

Finally we note in passing several limitations with the simple configuration described above. Firstly above we envisioned using a DC SQUID, which should be functional for the frequency range 100 Hz to $\sim$10 MHz. At higher frequencies, the SQUID noise may begin to dominate over the thermal noise; moving to an AC SQUID can stave off this transition to 1 GHz [111]. Beyond this an entirely different set up would be required to read out the flux through the pickup loop, one example being provided by a parametric amplifier. We refer to [111] for a detailed discussion of each of these regimes. Importantly, while more complicated circuits may lead to more complicated transfer functions in (3.31), so long as the frequency-dependent factors are common to both the signal and the noise, the analysis formalism described below goes through unchanged. Going forward, we assume that whenever discussing the resonant readout technique that we are in a thermal background dominated regime so the form of the transfer function is irrelevant.

## 3.3 Experimental Sensitivity

Armed with the likelihood given in (3.29), we will now determine the experimental sensitivity we can achieve.[10] Below we will firstly define a series of useful statistics that will be the basic tools

[9] In practice, we can often approximate $L_T \approx L_i$ for a resonant configuration.

[10] In this and subsequent sections, we will predominantly use a frequentist statistical framework when applying the likelihood. Nevertheless, we emphasize that our likelihood can be applied equally well within a Bayesian setting. In particular, in Sec. 3.5, we will use the Bayesian posterior as a tool for analyzing data in the presence of a putative

in our analysis. After this we will then use an Asimov based analysis, following [172], to study the expected background and signal distributions. We then introduce a procedure for stacking the data, which will reduce the computational demands associated with analyzing the enormous datasets axion direct detection experiments could potentially collect. Following on from this, we will show how to use the Asimov framework to estimate our expected upper limits and discovery threshold, fully accounting for the look elsewhere effect. Finally we will contrast our method to the simple $S/N = 1$ approach commonly used in the literature. An alternative analysis strategy to the one described in this section is to instead consider the average power in some frequency range near the expected signal location. Such an approach is less sensitive to the one presented here, and so we have relegated its discussion to App. B.2.

The starting point for our analysis is the likelihood $\mathcal{L}(d|\mathcal{M}, \boldsymbol{\theta})$. To claim a discovery or set limits on the axion, we need to know properties of the likelihood as a function of the coupling strength, which is effectively given by $A$, and the axion mass $m_a$. As such we separate out the parameters $\boldsymbol{\theta}$ into those of interest, $\{A, m_a\}$, and those describing the background, $\boldsymbol{\theta}_B$: $\boldsymbol{\theta} = \{A, m_a, \boldsymbol{\theta}_B\}$. Note that for now we fix the halo velocity distribution, though in the next two sections we generalize the model parameters to include ones that describe the DM velocity distribution. With this distinction, we can now set up our basic frequentist tool for testing the axion model, based on the profile likelihood:

$$\begin{aligned}\Theta(m_a, A) = 2[&\ln \mathcal{L}(d|\mathcal{M}, \{A, m_a, \hat{\boldsymbol{\theta}}_B\}) \\ &- \ln \mathcal{L}(d|\mathcal{M}_B, \hat{\boldsymbol{\theta}}_B)]\,,\end{aligned} \tag{3.35}$$

where in each of these terms $\hat{\boldsymbol{\theta}}_B$ denotes the values of the background parameters that maximize the likelihood for that dataset and model. Note in the second line we have defined the background-only model $\mathcal{M}_B$ that has $A = 0$ and model parameters $\boldsymbol{\theta}_B$.

In terms of this basic object we can now define two useful quantities. The first of these is a test statistic used for setting upper limits on $A$ and hence $g_{a\gamma\gamma}$:

$$q(m_a, A) = \begin{cases} \Theta(m_a, A) - \Theta(m_a, \hat{A}) & A \geq \hat{A}\,, \\ 0 & A < \hat{A}\,, \end{cases} \tag{3.36}$$

where $\hat{A}$ is the value of $A$ that results in the maximum value of $\Theta(m_a, A)$ at fixed $m_a$. The rationale for setting this test statistic to zero for $A < \hat{A}$ is that when setting upper limits, the best we can hope to do is constrain a parameter corresponding to one stronger than the best fit value. Observe that when $A \geq \hat{A}$, we have

$$q(m_a, A > \hat{A}) = 2[\ln \mathcal{L}(d|\mathcal{M}, \{A, m_a, \hat{\boldsymbol{\theta}}_B\} - \ln \mathcal{L}(d|\mathcal{M}, \{\hat{A}, m_a, \hat{\boldsymbol{\theta}}_B\}]\,, \tag{3.37}$$

signal.

and so this corresponds to the degradation in the likelihood as we increase $A$ beyond the best fit point. According to Wilks' theorem, the statistic $q$, at fixed $m_a$, is asymptotically a half-chi-squared distributed with one degree of freedom. It is a half and not full chi-squared distribution, as from the definition in (3.36), $q$ vanished by definition for $A < \hat{A}$. This implies, in particular, that for a given $m_a$, the 95% limit on $A$ will be set when $q(m_a, A_{95\%}) \approx -2.71$. Note also that when setting limits we allow $A$ to float negative.

The second object of interest is a test statistic for discovery, denoted TS, which quantifies by how much a model with an axion of a given mass provides a better fit to the data than one without it. This is defined as:

$$\mathrm{TS}(m_a) = \Theta(m_a, \hat{A})\,. \tag{3.38}$$

Below we will use the TS to quantify the 3 and 5$\sigma$ discovery thresholds, giving an accounting for the look elsewhere effect. But the intuition is that the larger the $\mathrm{TS}$ the more preferred the axion.

Importantly both $q$ and TS are defined in terms of $\Theta$, implying that through an understanding of this object we can determine everything about our two test statistics. As this will be the central object of interest, we will write out its form explicitly. Combining (3.35) with our form of the likelihood in (3.29), we arrive at:

$$\Theta(m_a, A) = 2\sum_{k=1}^{N-1}\left[S^k_{\Phi\Phi}\left(\frac{1}{\lambda_B} - \frac{1}{\lambda_k}\right) - \ln\frac{\lambda_k}{\lambda_B}\right]\,. \tag{3.39}$$

Recall that here $S^k_{\Phi\Phi}$ represents the data, whilst $\lambda_k$ and $\lambda_B$ represent the signal plus background and background only contributions respectively. We also reiterate that only $\lambda_k$ is a function of $m_a$ and $A$, and further that $\lambda_B$ can also be $k$ dependent if the background varies with frequency. Moreover, we stress that all $k$ modes need not be included in (3.39) in practice, but rather only the $k$ modes corresponding to frequencies in the vicinity of $m_a$.

### 3.3.1 Asymptotic Distribution of the Test Statistics

The object defined in (3.35) can be used immediately to quantify the preference for an axion signal in an experimental dataset, through the two test statistics defined above. Before looking at any data, however, it is often useful to know what the expected sensitivity is of an experiment using these statistics. Traditionally this is obtained via Monte Carlo simulations of the experiment, and through many realizations the expected distribution of $q$ and TS can be constructed. The problem is also analytically tractable, however, using the method of the Asimov dataset [172], which allows us to determine the asymptotic properties of the test statistics over many realization of the data. In this subsection we will exploit the Asimov approach to derive the asymptotic distribution of $\Theta$, and

then in subsequent sections we use this formalism to determine the expected limit and discovery potential of a prospective experiment.

The key step in the Asimov approach for our purposes is to take the dataset to be equal to the mean predictions of the model under question, neglecting statistical fluctuations. Consider the case where we have a dataset that contains a signal of the axion with signal strength $A_t$, where the subscript $t$ indicates this is the true value. In this case, the Asimov dataset is given by:

$$S_{\Phi\Phi}^{k,\text{Asimov}} \equiv \lambda_k^t = A_t \frac{\pi f(v)}{m_a v} + \lambda_B \,, \tag{3.40}$$

which is just (3.26) with $A \to A_t$. Note that this expression should be evaluated at $v = \sqrt{4\pi k/(m_a T) - 2}$, but here and below we leave the relation between $v$ and $k$ implicit. Now using this Asimov dataset, $\Theta$ becomes (suppressing the dependence on $m_a$):

$$\tilde{\Theta}(A) = 2 \sum_{k=1}^{N-1} \left[ \lambda_k^t \left( \frac{1}{\lambda_B} - \frac{1}{\lambda_k} \right) - \ln \frac{\lambda_k}{\lambda_B} \right] \,, \tag{3.41}$$

where $\tilde{\Theta}$ denotes the asymptotic form of $\Theta$. Importantly, one can check that this object is maximized exactly at $A = A_t$; in detail,

$$\max_A \tilde{\Theta}(A) = \tilde{\Theta}(A_t) \,. \tag{3.42}$$

Now if we assume that the experiment has been run long enough that the width of frequency bins is much smaller than the range over which $\lambda_k$ or $\lambda_B$ varies,[11] then we can approximate the sum over $k$ modes as an integral over velocity, just as we did in Sec. 3.2:

$$\begin{aligned} \tilde{\Theta}(A) =& \frac{T m_a}{\pi} \int dv\, v \left[ \left( A_t \frac{\pi f(v)}{m_a v} + \lambda_B \right) \right. \\ & \times \left( \frac{1}{\lambda_B} - \frac{1}{A \pi f(v)/(m_a v) + \lambda_B} \right) \\ & \left. - \ln \left( 1 + A \frac{\pi f(v)}{m_a v \lambda_B} \right) \right] \,. \end{aligned} \tag{3.43}$$

To further simplify the expression above, we note a signal will likely be much smaller than the background in any individual bin, such that $A\pi f(v)/(m_a v)$, $A_t \pi f(v)/(m_a v) \ll \lambda_B$. Expanding to leading order in $A$ and $A_t$, we then find

[11] Note that in general we would expect the signal to at least have a spread set by the velocity dispersion of the SHM, although in the presence of substructure the dispersion could be much smaller.

$$\tilde{\Theta}(A) \approx \frac{AT\pi}{m_a}\left(A_t - \frac{A}{2}\right)\int \frac{dv}{v}\frac{f(v)^2}{\lambda_B^2}\,, \tag{3.44}$$

where we have left $\lambda_B$ in the integral, as in general it will depend on frequency and hence velocity according to $\omega = m_a(1 + v^2/2)$. As before, we have boxed this equation due to its importance. In particular, we see already the dependence of our test statistic for detection on the local velocity distribution as it enters the integral in (3.44) that implies interesting results for axion direct detection. If we assume that the background is frequency independent, then this result tells us that the experimental sensitivity to the axion coupling $g_{a\gamma\gamma}^2$ scales as

$$g_{a\gamma\gamma}^2 \sim \frac{1}{\sqrt{\int_0^\infty dv\, \frac{f(v)^2}{v}}} \qquad (\text{Field})\,, \tag{3.45}$$

with the DM velocity distribution. This should be contrasted with the rate at WIMP[12] direct detection experiments, which scales with the mean inverse speed (see, for example, [193]). In particular, the limit on the DM cross-section $\sigma_{\rm DD}$ to scatter off ordinary matter, which generically scales with the coupling $g$ to ordinary matter as $g^2$, scales with the velocity distribution as

$$\sigma_{\rm DD} \sim \frac{1}{\int_{v_{\rm min}}^\infty dv\, \frac{f(v)}{v}} \qquad (\text{Particle})\,, \tag{3.46}$$

where $v_{\rm min}$ is the minimum speed required to cause the target nucleus in the detector to recoil at a given recoil energy. This cut off scales with the inverse reduced mass of the WIMP nucleon system, $v_{\rm min} \propto 1/\mu$, so that for lighter DM particles the rate is particularly sensitive to the upper end of the speed profile. In the axion case, the significance of an axion signal depends on an integral over the full speed profile. Importantly, the quadratic scaling of the integrand with the speed distribution implies that axion direct detection experiments are particularly sensitive to small scale structures in the speed profile, such as those that can be induced by local DM substructure. This stands in contrast to WIMP direct detection, where substructure is generally thought to only have a minimal impact, see, *e.g.*, [194].

We will explore the sensitivity of axion direct detection experiments to DM substructure in Sec. 3.5, but for now we illustrate the difference between axion and WIMP experiments noted above with a simple example. Suppose that there is a contribution to the local DM velocity distribution that can be modeled as a cold stream, with $f_{\rm str}(v) = 1/\delta v$ for $v_{\rm str} < v < v_{\rm str} + \delta v$ and zero otherwise. We assume that the stream width $\delta v \ll v_{\rm str}$, where $v_{\rm str}$ is the stream boost speed

[12]Here, we use weakly interacting massive particle (WIMP) direct detection to simply refer to the direct detection of massive DM particles at the ∼MeV scale and above, even if the particle models are not directly related to the WIMP paradigm.

in the Earth frame. Then, then in the WIMP case we find $\sigma_{\rm DD} \sim v_{\rm str}$, where we have assumed $v_{\rm str} > v_{\rm min}$. However, in the axion case there is an extra enhancement for small stream widths such that $g^2_{a\gamma\gamma} \sim \sqrt{v_{\rm str}\delta v}$. Note that this implies that as $\delta v$ decreases we can probe smaller values of $g_{a\gamma\gamma}$ in the axion case, while conversely decreasing $\delta v$ does not improve our sensitivity to $\sigma_{\rm DD}$ in the WIMP case.

Finally we note that if we repeated the analysis leading to (3.44) for the resonant case, we would instead have arrived at

$$\tilde{\Theta}^{\rm res}(A^{\rm res}) = \frac{Q_0^2 A^{\rm res} T\pi}{m_a}\left(A_t^{\rm res} - \frac{A^{\rm res}}{2}\right) \times \int \frac{dv}{v}\frac{f(v)^2}{(\tilde{\lambda}_B^{\rm therm})^2}\,, \tag{3.47}$$

which is essentially the same result but with the broadband quantities replaced with their appropriate resonant counterparts. Importantly, note that the transfer function and its associated frequency dependence has dropped out of this result because it involved a ratio of the signal to the background, both of which are linear in $\mathcal{T}(\omega)$. This justifies the claim that going forward our estimates for the resonant case can be obtained straightforwardly from the broadband results provided we make the substitutions:

$$\begin{aligned} A &\to Q_0 A^{\rm res}\,, \\ \lambda_B &\to \tilde{\lambda}_B^{\rm therm}\,. \end{aligned} \tag{3.48}$$

### 3.3.2 A Procedure for Stacking the Data

We would like a method to reduce the number of PSD components that need to be stored, without sacrificing sensitivity, given that if we are sampling at a high rate, for example $\sim$100 MHz or higher, over an extended time, the amount of data to be stored and analyzed could become substantial. As we will now show, stacking the PSD data provides exactly such a method.[13]

The central idea is to break the data up into $N_T$ subintervals of duration $\Delta T = T/N_T$, each with $\Delta N = N/N_T$ PSD components.[14] In each of these subintervals we calculate the PSD $S^{k,\ell}_{\Phi\Phi}$, where now $k$ only indexes the integers from 0 to $\Delta N - 1$, and we have the new index $\ell = 0, 1, \ldots, N_T - 1$ that identifies the relevant subinterval. Using this data, our likelihood takes the form

$$\mathcal{L}(d|\boldsymbol{\theta}) = \prod_{\ell=0}^{N_T-1}\prod_{k=1}^{\Delta N-1}\frac{1}{\lambda_k(\boldsymbol{\theta})}e^{-S^{k,\ell}_{\Phi\Phi}/\lambda_k(\boldsymbol{\theta})}\,. \tag{3.49}$$

Importantly, we assume that the model prediction in each subinterval is identical, which we com-

[13] We thank Jon Ouellet for conversations related to this point.

[14] The choice of notation here is used to emphasize that for $N_T \gg 1$ we have $\Delta T \ll T$ and $\Delta N \ll N$, but of course neither quantity should ever be thought of as infinitesimal.

ment on more below. With this assumption, it is natural to define a stacked PSD

$$\bar{S}^k_{\Phi\Phi} \equiv \frac{1}{N_T} \sum_{\ell=1}^{N_T-1} S^{k,\ell}_{\Phi\Phi} \,. \tag{3.50}$$

The averaged PSD components will be distributed as the average of a sum of exponentially distributed random variables with mean $\lambda_k$, which is given by the Erlang distribution:

$$P[\bar{S}^k_{\Phi\Phi}] = \frac{N_T^{N_T}}{(N_T-1)!} \frac{\left(\bar{S}^k_{\Phi\Phi}\right)^{N_T-1}}{\lambda_k^{N_T}} e^{-N_T \bar{S}^k_{\Phi\Phi}/\lambda_k} \,. \tag{3.51}$$

Using this stacked data, we can simplify (3.49) by removing the sum over $\ell$:

$$\mathcal{L}(d|\boldsymbol{\theta}) = \prod_{k=1}^{\Delta N-1} \frac{1}{\lambda_k(\boldsymbol{\theta})^{N_T}} e^{-N_T \bar{S}^k_{\Phi\Phi}/\lambda_k(\boldsymbol{\theta})} \,, \tag{3.52}$$

where in this result we can already see the reduction in computational requirements as it only involves a product over $\Delta N \ll N$ numbers, since the $\bar{S}^k_{\Phi\Phi}$ can be precomputed and updated as more data comes in.

Our next task is to determine how this choice will impact our sensitivity, using the test statistics defined in the previous subsections. It is sufficient to consider $\Theta(m_a, A)$, defined in (3.35) and from which the other statistics of interest can be derived. Doing so, we can repeat the Asimov analysis from the previous subsection to determine the asymptotic form of the stacked $\Theta$, given by

$$\tilde{\Theta}_{\rm stacked}(A) = \frac{A N_T \Delta T \pi}{m_a} \left(A_t - \frac{A}{2}\right) \int \frac{dv}{v} \frac{f(v)^2}{\lambda_B^2} \,. \tag{3.53}$$

Yet as $N_T \Delta T = T$, the stacked and unstacked form of $\tilde{\Theta}$ are identical. This implies that our stacking procedure, which for $N_T \gg 1$ dramatically reduces the required computation, has no impact on our sensitivity to an axion signal.

There is, however, a catch. Stacking implies that we are only sensitive to frequency shifts of size $\Delta f = 1/\Delta T$, which can be much larger than the shifts we were sensitive to in the full dataset, where $\Delta f = 1/T \ll 1/\Delta T$. This could mean, depending on the size of the frequency spacings, that ultra-cold local DM substructure is no longer resolved, and therefore the enhancement it would have given to the integral over velocity discussed above is lost. In this sense stacking can lead to a degradation in sensitivity, and so choosing a stacking strategy should be done with careful consideration of the features being searched for. To provide some intuition, if we are searching for an axion at a mass corresponding to a frequency $f$ and drawn from a velocity distribution with dispersion $v_0$, then the coherence time is $\sim 1/(f v_0 v_{\rm obs})$. To be able to fully resolve the

axion signal we would then want $\Delta T \gg 1/(f v_0 v_{\rm obs})$. For the SHM, and scanning in frequencies from 100 MHz down to 100 Hz, the coherence time varies from 20 ms up to 5 hours. In such a scenario, if data were collected for a year, many stacking procedures would be feasible. On the other hand if searching for the signal from a cold stream with a dispersion of $v_0 = 1$ km/s, then over the same frequency range the coherence time varies from 4 seconds up to 45 days. For the lowest frequencies in this case, any stacking procedure would be sacrificing sensitivity to such cold substructure. On the other hand, at the lowest frequencies high sampling rates are not necessary. Thus, a hybrid approach may be preferable in practice, where the data is stacked in Fourier space at high frequencies while at low frequencies the data is stacked in time (*i.e.* down-sampled) in order to reduce the data size without sacrificing the sensitivity to cold substructure at any possible axion mass.

Another relevant consideration is that due to the Earth's acceleration, lab-frame frequencies may shift throughout the day and year, which would invalidate our assumption that the model predictions are identical between subintervals. The rotational speed of the Earth's surface about its axis is roughly $0.46\cos(\delta)$ km/s, where $\delta$ is the latitude. This value is small enough that it can safely neglected for any cold flow with a velocity dispersion greater than this. The rotation of the Earth about the Sun, however, occurs at roughly 30 km/s and is thus harder to ignore when searching for cold substructure, as we discuss later in this work. Annual and daily modulation can lead to striking additional signatures, which we explore in detail in Sec. 3.5.

### 3.3.3 Expected Upper Limit

We are now in a position to write down the expected 95% limit on $A$. In the case of a limit, the appropriate Asimov dataset to use is a background only distribution, so that $A_t = 0$. Then by combining our definition of the likelihood profile in (3.36) with our Asimov result in (3.44), we arrive at the 95% limit where $q(m_a, A_{95\%}) = -2.71$, given by

$$\tilde{A}_{95\%} = \sqrt{2.71 \left[\frac{T\pi}{2m_a}\int \frac{dv}{v}\frac{f(v)^2}{\lambda_B^2}\right]^{-1}}\,. \tag{3.54}$$

Note that again the tilde indicates this is an Asimov, or median, quantity. This result is particularly relevant as it can be interpreted for any experiment provided the definition of appropriate $A$ and $\lambda_B$. For the particular example of ABRACADABRA we can insert the form of $A$ given in (3.15), yielding

$$\tilde{g}_{a\gamma\gamma}^{95\%} = \frac{2.71^{1/4}\sqrt{L_p/L}}{\alpha B_{\rm max} V_B \sqrt{\rho_{\rm DM}}} \times \left[\frac{T\pi}{32m_a}\int \frac{dv}{v}\frac{f(v)^2}{\lambda_B^2}\right]^{-1/4}\,. \tag{3.55}$$

One of the real powers of the Asimov analysis is that not only can we determine the median expected limit, we can also derive analytically the expected size of fluctuations away from the central value, without having to revert to Monte Carlo simulations. The details of this statistical procedure are discussed in [172]. As we are constructing power-constrained 95% one-sided limits, we obtain confidence intervals via

$$q(m_a, A_{95\%\pm N\sigma}) = -\left(\Phi^{-1}\left[0.95\right] \pm N\right)^2 , \tag{3.56}$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution (zero mean and unit variance), and $\Phi^{-1}$ is the inverse of this (so $\Phi^{-1}\left[0.95\right] \approx 1.64$). Note that if we take $N = 0$, then the above just reduces to $q(m_a, A_{95\%}) = -2.71$, but this more general result contains the information about the error bands in the expected limit. In this way, by replacing the 2.71 that appears in (3.55) with the appropriate value for the $N\sigma$ uncertainty band on the 95% limit, we can construct the median and uncertainty bands on $\tilde{g}^{95\%}_{a\gamma\gamma}$ analytically. For completeness, in App. B.3 we verify that the bands constructed in this manner agree with those generated using Monte Carlo simulations. Finally, to be conservative we use power-constrained limits [195], which in practice means we do not allow ourself to set a limit below our $1\sigma$ uncertainty band on the upper limit.

### 3.3.4 Expected Discovery Reach

In order to find evidence for a signal, we need to understand the expected distribution of the TS under the null hypothesis. The reason is that this distribution determines how likely the background is to produce a given TS value, and hence what threshold $\mathrm{TS}_{\mathrm{thresh}}$ we should set to establish the existence of a signal at a given confidence level. Once we have such a threshold test statistic, applying our Asimov results above to the case of discovery, we find we would be sensitive to discover a signal with the following strength

$$\tilde{g}^{\mathrm{thresh}}_{a\gamma\gamma} = \frac{\mathrm{TS}^{1/4}_{\mathrm{thresh}}\sqrt{L_p/L}}{\alpha B_{\mathrm{max}} V_B \sqrt{\rho_{\mathrm{DM}}}} \times \left[\frac{T\pi}{32 m_a}\int \frac{dv}{v}\frac{f(v)^2}{\lambda_B^2}\right]^{-1/4} . \tag{3.57}$$

Locally, the significance in favor of the axion model is expected to be approximated by $\sqrt{\mathrm{TS}}$ [172]; that is, a value $\mathrm{TS} = 25$ corresponds to approximately $5\sigma$ local significance. However, when scanning over multiple independent mass points, the look elsewhere effect must be accounted for in quoting values for the global rather than local significance. The look elsewhere effect may be determined through Monte Carlo simulations. However, in this section we will derive an analytic approximation to $\mathrm{TS}_{\mathrm{thresh}}$, which accounts for the look elsewhere effect, and as we will show provides an accurate representation to the output from such Monte Carlo studies. The result will be a mapping between the desired global significance threshold and the value of $\mathrm{TS}_{\mathrm{thresh}}$ that should

be taken, depending on the mass range scanned. We note that there are also other proposals in the literature for approaching this problem; for a recent one see, *e.g.*, [196].

Our starting point is to note that the asymptotic form of the survival function for the local TS under the null hypothesis is given by

$$S[\mathrm{TS}_{\mathrm{thresh}}] = 1 - \Phi\left(\sqrt{\mathrm{TS}_{\mathrm{thresh}}}\right) , \tag{3.58}$$

where $S[\mathrm{TS}_{\mathrm{thresh}}]$ is the probability that the TS, under the null hypothesis, takes a value greater than $\mathrm{TS}_{\mathrm{thresh}}$. This is derived explicitly in App. B.4 starting from the likelihood function, and it is equivalent to the statement that the asymptotic local significance is given by $\sqrt{\mathrm{TS}}$. However in any realistic experiment, we will look in a number of independent frequency windows corresponding to different axion masses. To account for this we need to note that in any of these windows there could be an upward fluctuation. To do so let us say that we look at $N_{m_a}$ independent mass points, and we want to set the threshold test statistic, $\mathrm{TS}_{\mathrm{thresh}}$, such that the probability that the background will not fake the signal in any bin is $1 - p$. To relate these two quantities, if we assume that $p$ is small enough, we can write the probability that at least one of the TSs, from the set over all mass points, is greater than $\mathrm{TS}_{\mathrm{thresh}}$ as

$$p = 1 - (1 - S[\mathrm{TS}_{\mathrm{thresh}}])^{N_{m_a}} \approx N_{m_a} S[\mathrm{TS}_{\mathrm{thresh}}] \, . \tag{3.59}$$

From here we can then substitute the survival function from (3.58), and expanding this out gives

$$\mathrm{TS}_{\mathrm{thresh}} = \left[\Phi^{-1}\left(1 - \frac{p}{N_{m_a}}\right)\right]^2 . \tag{3.60}$$

Using this result, as soon as we know $N_{m_a}$ we can determine $\mathrm{TS}_{\mathrm{thresh}}$ as it should be used in our formula for $\tilde{g}_{a\gamma\gamma}^{\mathrm{thresh}}$ in (3.57). To give some intuition, in the case where we ignore the look elsewhere effect and set $N_{m_a} = 1$, then the $3\sigma$ requirement is that $p \approx 1.35 \times 10^{-3}$, yielding $\mathrm{TS}_{\mathrm{thresh}} = 9$, as expected. Importantly, note that the $p$ values here correspond to that for 1-sided fluctuations [172].

In any realistic experiment, we expect $N_{m_a} \gg 1$. However, estimating the correct value for $N_{m_a}$ is complicated by the fact that we may scan over a continuum of different possible mass points in practice, though not all of the mass points have independent data. We expect a mass point as frequency $m_a$ to extend over a frequency bandwidth $\sim m_a v_0^2$, for the SHM. Thus, we expect to be able to characterize a set of independent mass point by the relation

$$m_a^{(i)} = m_a^{(0)} \left(1 + \alpha v_0 v_{\mathrm{obs}}\right)^i , \tag{3.61}$$

where $m_a^{(0)}$ is the first mass point, $i = 0, \ldots, N_{m_a} - 1$, and $\alpha$ is a number order unity that should be tuned to Monte Carlo simulations. Given the parameterization in (3.61), we may estimate the number of mass points by relating $m_a^0$ with the minimum frequency $f_{\rm min}$ and $m_a^{(N_{m_a}-1)}$ with the highest frequency $f_{\rm max}$; solving for $N_{m_a}$ in the limit $N_{m_a} \gg 1$ then gives

$$N_{m_a} \approx \frac{1}{\alpha\, v_0 v_{\rm obs}} \ln \frac{f_{\rm max}}{f_{\rm min}}\,. \tag{3.62}$$

In Fig. 3.2 we compare the analytic prediction in (3.60), combined with (3.62), with the result of 2.5 million Monte Carlo simulations. From the ensemble of simulations, we are able to compute the value of $p$ for each value of $\text{TS}_{\rm thresh}$. Note that in each simulation we scan for axion DM over a frequency range $f_{\rm max}/f_{\rm min} \approx 1.0007$; setting $v_0 = 220$ km/s and $v_{\rm obs} = 232$ km/s then gives, through (3.62), $N_{m_a} \approx 1.23 \times 10^3/\alpha$. The analytic results are found to agree well with the simulations for $\alpha \approx 3/4$; this value may also be understood by thinking more carefully about the extent of the SHM. Note that the real power of the analytic formalism is that once we have tuned the relations in (3.60) and (3.62) to Monte Carlo, in order to find the appropriate value of $\alpha$, we may extrapolate to smaller values of $p$, where the number of Monte Carlo simulations required to directly determine $\text{TS}_{\rm thresh}$ would be intractable.

To give some more realistic examples, if we assume the experiments scans from 100 Hz to 100 MHz, using the SHM values we obtain $N_{m_a} \sim 3 \times 10^7$. This then increases the $3\sigma$ ($5\sigma$) threshold TS to 40.9 (57.5). To contrast if instead our significance was dominated by a stream with dispersion roughly 20 km/s, then instead we would find $N_{m_a} \sim 4 \times 10^9$, and the $3\sigma$ ($5\sigma$) threshold TS becomes 50.3 (67.0).

### 3.3.5 Comparison with $S/N = 1$

In the absence of a full likelihood framework, a common method employed for estimating sensitivity is obtained by setting the signal equal to the expected background, or $S/N = 1$. For example, this approach was used in the original ABRACADABRA proposal [71] and also for the proposed CASPEr experiment [58]. In this section we want to contrast this simple estimate to the output from our full likelihood machinery.

Now, following these earlier references, in our notation the signal-to-noise ratio can be written as

$$S/N = |\Phi_{\rm SQUID}|\,(T\tau)^{1/4}/\sqrt{|\lambda_B|}\,, \tag{3.63}$$

where $\tau$ is the signal coherence time. This $S/N \propto T^{1/4}$ scaling occurs when the collection time is longer than the coherence time. If $T < \tau$, instead the significance grows as $S/N \propto T^{1/2}$, as demonstrated in [58]. In App. B.5 we demonstrate that this same scaling can also be seen directly

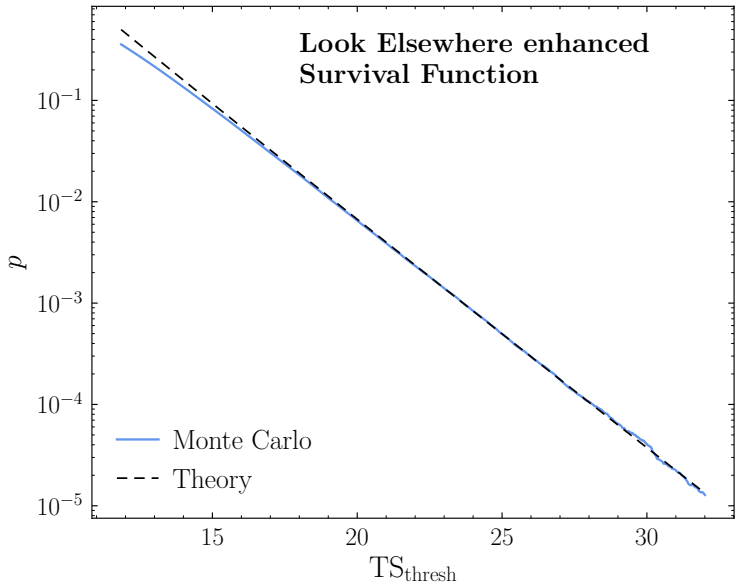



Figure 3.2: A comparison between the look elsewhere effect improved survival function approximate result derived between (3.60) and (3.62), and the equivalent values derived directly from Monte Carlo simulations. The good agreement between the two, especially at large $\mathrm{TS}_{\mathrm{thresh}}$ demonstrates that our approximate result is useful for estimating how often the background can fluctuate to fake the signal at a given confidence level. Note the values plotted here correspond to signals varying from 0 to $4\sigma$, for derived values of $\lambda_B$ given in (3.28) and 2.5 million Monte Carlo simulations. We do not extend the plot up to the $5\sigma$ value relevant for discovery, as this would require roughly 100 times as many simulations. This statement in itself already demonstrated the utility of our approximate analytic result.

from our likelihood.

In order to make a concrete comparison, we consider ABRACADABRA with the axion following only the bulk velocity distribution. In this case, the coherence of the bulk halo, as discussed above, will effectively ensure we always have $T \gg \tau$, implying the signal grows as $T^{1/4}$. To simplify (3.63), firstly consider $|\Phi_{\mathrm{SQUID}}|$. Combining (3.13) and (3.12), we have:

$$|\Phi_{\mathrm{SQUID}}| = \frac{\alpha}{2}\sqrt{\frac{L}{L_p}} g_{a\gamma\gamma} B_{\max} V_B m_a \left|a(t)\right| \,. \tag{3.64}$$

For the purposes of determining the average axion field over a time $T \gg \tau$, we can simply consider the axion field in the zero velocity limit, where

$$|a(t)| = \frac{\sqrt{2\rho_{\mathrm{DM}}}}{m_a} \left|\cos(m_a t)\right| = \frac{\sqrt{\rho_{\mathrm{DM}}}}{m_a} \,. \tag{3.65}$$

Note that since it is the PSD that is measured in practice, we calculate the average as

$\sqrt{|\cos^2(m_a t)|} = 1/\sqrt{2}$. The coherence time is determined by the kinetic energy $\frac{1}{2}m_a v^2$, which perturbs the axion frequency. Once the phase shift from this correction equals $\pi$, the field will be fully out of phase, so we take

$$\tau = \frac{2\pi}{m_a v_0 v_{\rm obs}}\,, \tag{3.66}$$

where again with the bulk halo in mind, we took values appropriate for the SHM. Finally, we assume that we have a frequency independent background PSD $\lambda_B$. Combining these results with the threshold $S/N = 1$, we obtain a sensitivity estimate of

$$g_{a\gamma\gamma} = \frac{2\sqrt{\lambda_B}\sqrt{L_p/L}}{\alpha B_{\rm max} V_B \sqrt{\rho_{\rm DM}}} \left(\frac{m_a v_0 v_{\rm obs}}{2\pi T}\right)^{1/4}\,. \tag{3.67}$$

We want to contrast this estimate with the exact value we obtain from the analysis method outlined in this section. For this purpose we take our result, but evaluated at some $\mathrm{TS}_{\rm req}$ which is schematic—it can be 2.71 for the case of a 95% limit, or ∼58 for a $5\sigma$ discovery accounting for the look elsewhere effect. If we assume $f(v)$ follows the SHM and further take $v_{\rm obs} = v_0$, then the equivalent result is:

$$g_{a\gamma\gamma} = \left(\frac{64\,\mathrm{TS}_{\rm req}\sqrt{2\pi}}{\mathrm{erf}\left[\sqrt{2}\right]}\right)^{1/4} \times \frac{\sqrt{\lambda_B}\sqrt{L_p/L}}{\alpha B_{\rm max} V_B \sqrt{\rho_{\rm DM}}} \left(\frac{m_a v_0^2}{2\pi T}\right)^{1/4}\,. \tag{3.68}$$

Note that the formula above is equivalent to the statement that

$$S/N \approx 1.8\,\mathrm{TS}_{\rm req}^{1/4}\,. \tag{3.69}$$

For example, the 95% expected upper limit would require $S/N = 2.31$, whilst a $5\sigma$ discovery accounting for the look elsewhere effect assuming the SHM, requires $S/N = 4.97$. We will see in the next section that the comparisons are similar for a resonant experiment also. In general the various thresholds are achieved with a larger signal than the naive $S/N = 1$ suggests. Nonetheless, the standard estimate is not a terrible approximation to the true results, especially considering that $S \sim g_{a\gamma\gamma}^2$. We emphasize, however, that there is a lot more that can be extracted from having the full likelihood framework, which we turn to in the subsequent sections.

## 3.4 Application to the Bulk Halo

In this section we apply the formalism developed so far to ABRACADABRA and ADMX. For this purpose we take a simple concrete example, where $f(v)$ describes only the bulk halo, which we further assume follows the SHM as defined in (3.9). Additionally we assume that over the

frequency band of the signal,[15] the mean of the background distribution in frequency space is approximately frequency independent. These assumptions imply that the integral appearing in (3.55) and (3.57) can be evaluated exactly:

$$\int \frac{dv}{v} \frac{f(v)^2}{\lambda_B^2} = \frac{\mathrm{erf}\left[\sqrt{2} v_{\rm obs}/v_0\right]}{\sqrt{2\pi} v_0 v_{\rm obs} \lambda_B^2}\,, \tag{3.70}$$

with $\sqrt{\lambda_B} \approx 10^{-6} \Phi_0/\sqrt{\mathrm{Hz}}$ as given in (3.27). In the following subsections, we will demonstrate explicitly how to construct projected limits and detection sensitivities, under the assumption of the SHM velocity distribution, and we will show in the event of a detection the parameters of the SHM may be determined using the likelihood framework. We will extend this framework to more realistic $f(v)$, including DM substructure, in the next section.

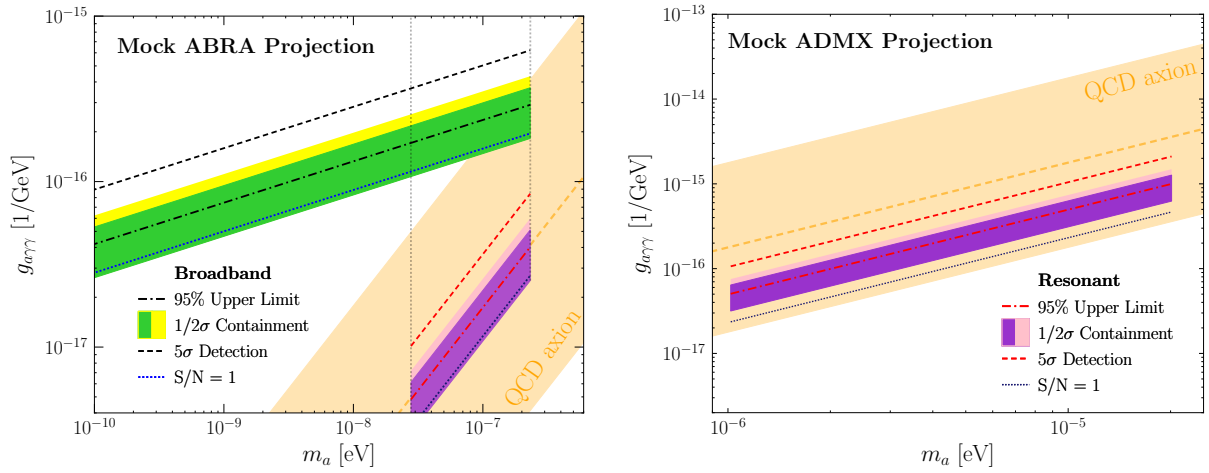



Figure 3.3: (Left) A comparison of the projected sensitivities for a hypothetical version of the ABRACADABRA (ABRA) experiment [71], with inner toroidal radius $R = 0.85$ m, an outer toroidal radius double this value, and a height $h = 4\,R$. A maximum magnetic field of 10 T is assumed, along with an interrogation time of 1 year. (Right) An equivalent comparison of projections for a future ADMX experiment. Here we take a total run time of 5 years, a volume of 500 L, quality factor of $10^5$, magnetic field of 7 T, and a system temperature of 148 mK. In both panels the exact sensitivities are contrasted with an estimate obtained from the signal-to-noise ratio, $S/N = 1$.

### 3.4.1 Sensitivity

In Fig. 3.3 we illustrate the formalism introduced in Sec. 3.3 for hypothetical future versions of the ABRACADABRA and ADMX experiments. To be specific, for ABRACADABRA we assumed

[15] By the frequency band we simply mean the range of frequencies over which the signal will be significant, which for the SHM is approximately $[m_a, m_a(1 + v_0 v_{\rm obs})]$.

that the radius of the pickup loop is identical to the inner radius of the torus, $R$, and also equal to the width of the torus, so that the total radius out to the outer edge of the toroid is $2R$. For concreteness, we took $R = 0.85$ m and then set the height of the torus to be $h = 4\,R$. For the remaining parameters we generally follow [71], taking $\alpha^2 = 0.5$, pickup-loop inductance $L_p = \pi R^2/h$, SQUID inductance $L = 1$ nH, and local DM density $\rho_{\rm DM} = 0.4\ {\rm GeV/cm^3}$. In the broadband mode we assume a flat spectrum of SQUID noise of $\sqrt{\lambda_B} = 10^{-6}\Phi_0/\sqrt{\rm Hz}$. In the resonant mode, we take a temperature of 100 mK and $Q_0 = 10^6$ for the RLC circuit. Note that we cut off our projections when the Compton wavelength of the axion is equal to the inner radius of the detector. The reason for this is that at high frequencies the magnetoquasistatic approximation used in the original analysis of [71], which we follow, breaks down. ABRACADABRA is still expected to set limits in this regime, albeit weaker, however in the absence of a detailed treatment we leave this region out.[16]

For ADMX, we use the projected values recently presented in [197], which updated the earlier projections from [107,198]. We take the volume $V = 500$ L, quality factor $Q = 10^5$, magnetic field $B = 7$ T, and system temperature $T_s = 148$ mK. So far, we have not described how our analysis framework is modified for the case of ADMX. Nevertheless, it is again a simple modification of the framework presented in Sec. 3.2. Starting from the power the axion field and thermal noise sources generate in the ADMX cavity, which is described in detail in a number of references, see *e.g.*, [151, 157, 183, 188, 199, 200], we find

$$\begin{aligned} A^{\rm ADMX} =& g_{a\gamma\gamma}^2 \frac{\rho_{\rm DM}}{m_a} Q B^2 V C_{010}\,, \\ \lambda_B^{\rm ADMX} =& k_B T_s\,, \end{aligned} \tag{3.71}$$

where $C_{010} \approx 0.692$ is the cavity form factor for the $\rm TM_{010}$ mode, which dominates for the ADMX configuration. In terms of these quantities, the mean PSD is given by

$$\lambda^{\rm ADMX}(\omega) = \left( A^{\rm ADMX}\ \frac{\pi f(v)}{m_a v}\bigg|_{v=\sqrt{2\omega/m_a - 2}} + \lambda_B^{\rm ADMX} \right) \times \mathcal{T}^{\rm ADMX}(\omega)\,,$$

where $\mathcal{T}^{\rm ADMX}(\omega)$ is the transfer function for the ADMX resonant cavity. The transfer function has support over a frequency interval of width $\sim\omega_0 Q^{-1}$, where $\omega_0$ is the resonant frequency, in analogy to (3.31). However, the exact form of this transfer function is not important for our purposes, since it is common to the noise and signal contributions. In addition to computing the sensitivity of ADMX using our likelihood framework, we also derive an $S/N = 1$ estimate for the sensitivity from the Dicke radiometer equation [170].

[16] Preliminary simulations indicate that good sensitivity is likely maintained to somewhat higher frequency values. We thank Kevin Zhou for these preliminary results.

In Fig. 3.3, the dashed curves represents the sensitivity for a $5\sigma$ discovery, using the formalism derived in Sec. 3.3.4, including the look elsewhere effect.[17] We also show the median expected 95% limit, as well as the 1 and $2\sigma$ bands on the expectations for these quantities, derived using the procedure described in Sec. 3.3.3. We reiterate that we present power-constrained limits [195], so that we do not allow ourselves to set limits stronger than the expected $1\sigma$ downward fluctuation. In addition we have also added the naive $S/N = 1$ estimated sensitivity line for the broadband mode, as given in (3.67). As shown in Sec. 3.3.5, the 95% limit and detection threshold differ only from the naive estimate by factors of order unity. The figure also includes the theoretically motivated region for the QCD axion in orange.

For the resonant results shown in Fig. 3.3, we adjusted the scanning strategy such that the mean limit under the null hypothesis is parallel to the QCD line in the $g_{a\gamma\gamma} - m_a$ plane. For ABRACADABRA, we chose a minimum mass $m_a = 2.8 \times 10^{-8}$ eV and a maximum mass $m_a = 2.3 \times 10^{-7}$ eV, and the total number of bins scanned in the resonant search was $1.3 \times 10^6$. A total scanning time of 1 year was used. The lowest-frequency bin was scanned for $T = 704$ s, while the highest-frequency bin was scanned for $T = 0.0175$ s; the amount of time spent at the $i^{\text{th}}$ mass scales as $T \propto (m_a^i)^{-5}$. Note that we have not considered the possibility of incorporating an additional broadband readout in the resonant scan to increase the sensitivity, though such an approach may be feasible. For ADMX, we instead scanned between masses of $1.0 \times 10^{-6}$ and $20.1 \times 10^{-6}$ eV, using a total of $1.8 \times 10^6$ mass bins. Here a total scanning time of 5 years was broken up as follows: the smallest and largest masses were scanned for 268 and 13.5 s, and now the time spent at the $i^{\text{th}}$ mass scales as $T \propto (m_a^i)^{-1}$.

To simulate what an actual limit would look like as derived from real data, we generate Monte Carlo data for the mock broadband ABRACADABRA experimental setup under the assumption that the axion explains all of DM with $m_a = 10^{-8}$ eV and $g_{a\gamma\gamma} = 2.21 \times 10^{-16}\,\text{GeV}^{-1}$. Fig. 3.4 shows the resulting limit in the vicinity of the true mass; the region has been magnified so that the bin to bin fluctuations can be seen. The figure shows that in general the limit moves around between the expected bands, however right at the center, at the location of the true mass, the limit weakens considerably.

### 3.4.2 Parameter Estimation

In this section, we show how to estimate the DM coupling to photons and aspects of the DM phase-space distribution in the event of a detection or a detection candidate. This is done in practice by

[17] We caution that in the resonant case, looking for upwards fluctuations in excess of the $5\sigma$ look elsewhere effect enhanced detection thresholds is unlikely to be the optimal discovery strategy. Instead, one could, for example, further interrogate masses where a $2\sigma$ upward fluctuation is observed. For example, ADMX implements exactly such a strategy, as described in [130]. We make no attempt to determine the ideal resonant discovery strategy in this work.

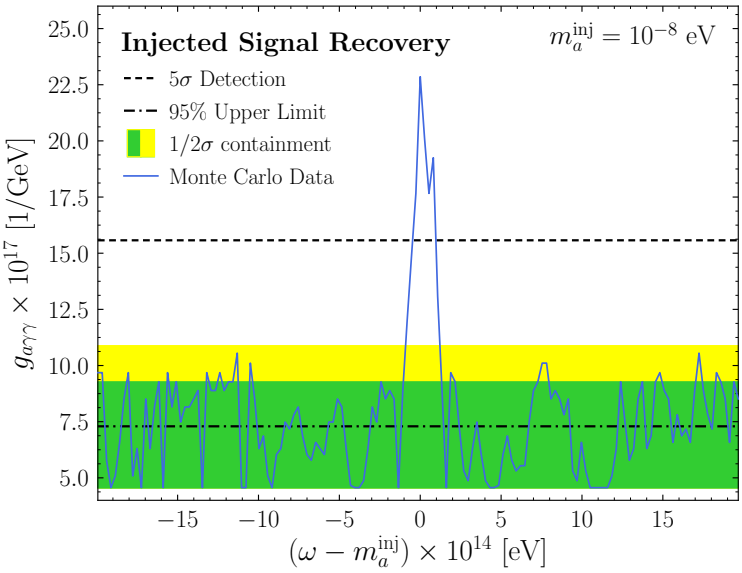



Figure 3.4: An actual limit obtained from a single Monte Carlo simulation, with the broadband readout, compared to the various expectations for the broadband ABRACADABRA framework used in Fig. 3.3. The data was simulated with an injected signal corresponding to $m_a = 10^{-8}$ eV and $g_{a\gamma\gamma} = 2.21 \times 10^{-16}$ GeV$^{-1}$, and indeed we can see that right near the frequency corresponding to the axion mass, we are unable to exclude the corresponding signal strength.

scanning over the likelihood function with the relevant degrees of freedom given to the parameters of interest. In this section, we show how to anticipate the uncertainties on the parameter estimates using the Asimov framework. We proceed in an analogous fashion to previous sections, where we studied the asymptotic form of the background only distribution; in this section, we study the asymptotic form of the likelihood in the presence of a signal.

As a starting point, consider estimating the signal strength $A$ from a dataset drawn from a distribution where the true value is $A_t$. Note that we use $A$ rather than $g_{a\gamma\gamma}$ only to simplify the expressions; the extension to the actual parameter of interest is straightforward. Recall we have actually already shown in the previous section that the asymptotic form of $\Theta$ given in (3.44) has the key property that it is maximized at the correct value of the signal strength, $A_t$.[18] We can determine the uncertainty on the estimated $A$ from the curvature around the maximum. In detail,

$$\sigma_A^{-2} = -\frac{1}{2}\partial_A^2 \tilde{\Theta}(A)|_{A=A_t} = \frac{T\pi}{2m_a}\int \frac{dv}{v}\frac{f(v)^2}{\lambda_B^2}\,, \tag{3.72}$$

where $\sigma_A$ is the expected uncertainty on the measurement. Using the SHM velocity distribution,

[18] Recall we assumed $A_{(t)}\pi f(v)/(m_a v) \ll \lambda_B$ in deriving that expression.

this simplifies to

$$\sigma_A = \sqrt{\frac{2\sqrt{2}\, m_a\, \lambda_B^2\, v_0 v_{\rm obs}}{T\sqrt{\pi}\,\mathrm{erf}\left[\sqrt{2}v_{\rm obs}/v_0\right]}} = \frac{A_t}{\sqrt{\rm TS}}\,. \tag{3.73}$$

From this we can see that, as expected, the uncertainty on the signal strength increases with the background, decreases with a longer experimental run time, and scales inversely proportional to the square root of the TS for detection. The last point is important because it says that the central value $A_t$ is $\sqrt{\rm TS}$ standard deviations away from zero, which matches our interpretation of $\sqrt{\rm TS}$ as the significance.

We can readily extend this strategy to the estimation of other signal parameters. For example, we can use this to estimate the best fit SHM parameters, $v_0$ and $v_{\rm obs}$, and their associated uncertainties. Let us denote by $f_t(v) = f_{\rm SHM}(v|v_0^t, v_{\rm obs}^t)$ the speed distribution given by the true SHM parameters, and then $f(v) = f_{\rm SHM}(v|v_0, v_{\rm obs}^t)$ represents the distribution for some arbitrary value of $v_0$. To repeat the Asimov analysis, we now use the dataset and model predictions given by

$$\begin{aligned} S_{\Phi\Phi}^{k,\rm Asimov} \equiv \lambda_k^t =& A_t \frac{\pi f_t(v)}{m_a v} + \lambda_B\,, \\ \lambda_k =& A_t \frac{\pi f(v)}{m_a v} + \lambda_B\,, \end{aligned} \tag{3.74}$$

respectively. Then, through the same process as above we arrive at

$$\tilde{\Theta}(v_0) = \frac{A_t^2 T\pi}{m_a}\int \frac{dv}{v}\frac{f(v)}{\lambda_B^2}\left(f_t(v) - \frac{f(v)}{2}\right)\,. \tag{3.75}$$

Again this asymptotic expression satisfies the central Asimov requirement that

$$\max_{v_0}\tilde{\Theta}(v_0) = \tilde{\Theta}(v_0^t)\,. \tag{3.76}$$

Beyond this, however, we can again estimate the uncertainty on the best fit velocity dispersion:

$$\sigma_{v_0}^{-2} = -\frac{1}{2}\partial_{v_0}^2\tilde{\Theta}(v_0)|_{v_0=v_0^t} = \frac{A_t^2 T\pi}{2m_a}\int\frac{dv}{v}\frac{\left(\partial_{v_0}f(v)|_{v_0=v_0^t}\right)^2}{\lambda_B^2}\,, \tag{3.77}$$

so that if we assume $\lambda_B$ is independent of frequency, we have

$$\sigma_{v_0} = \frac{v_0^t}{\sqrt{\rm TS}}\left(\frac{3}{4} + \frac{v_{\rm obs}^t\left(9v_0^{t2} - 4v_{\rm obs}^{t2}\right)e^{-2v_{\rm obs}^{t2}/v_0^{t2}}}{\sqrt{2\pi}v_0^{t3}\,\mathrm{erf}\left[\sqrt{2}v_{\rm obs}^t/v_0^t\right]}\right)^{-1/2} \approx 1.02\frac{v_0^t}{\sqrt{\rm TS}}\,.$$

Above, we have taken the SHM values for the approximate result. Applying the same strategy for $v_{\rm obs}$, we would find the maximum is again obtained at the true value, with the uncertainty now

given by

$$\sigma_{v_{\rm obs}} = \frac{v_0^t}{\sqrt{\rm TS}}\left(1 - \frac{4v_{\rm obs}^t e^{-2v_{\rm obs}^{t2}/v_0^{t2}}}{\sqrt{2\pi}v_0^t \,{\rm erf}\left[\sqrt{2}v_{\rm obs}^t/v_0^t\right]}\right)^{-1/2} \approx 1.11\frac{v_0^t}{\sqrt{\rm TS}}\,.$$

From these three results for parameter estimation using our likelihood we can see that in general if we are estimating a parameter $\alpha_t$, the estimated mean value will be $\mu_\alpha = \alpha_t$, and the uncertainty tends to scale as $\sigma_\alpha \sim {\rm TS}^{-1/2}$. Thus exactly as expected, the more significant the detection of axion, or specifically the larger the TS, the greater precision with which we can estimate parameters.

## 3.5 Impact of a Realistic and Time-Varying DM Distribution

In the previous sections, we have developed a framework for the analysis of a signal sourced by axion DM drawn from the SHM distribution $f_{\rm SHM}(v|v_0, v_{\rm obs})$. However, this neglects a number of effects that modify the DM speed distribution; in particular: annual modulation, gravitational focusing, and the possible presence of local velocity substructure. As we have verified by Monte Carlo simulations, the exclusion of these features from our analysis has a negligible effect on our ability to successfully constrain or discover an axion signal in our data, even when features excluded from the analysis are included in the data sets. Consequently, the framework of Sec. 3.4 is sufficient for the first stage of the data analysis. Nonetheless, since we do expect these effects to be manifest in a hypothetically discovered signal, they present opportunities to gain sharper insight on the local DM distribution. Moreover, because annual modulation and gravitational focusing result in distinct signatures expected to be present only in the presence of a genuine axion signal, the identification of these features would further strengthen any candidate detection. In addition, if we are within a cold stream or debris flow, a significant enhancement to the signal is possible. In this section, we specify the details of annual modulation, gravitational focusing, and velocity substructure and their inclusion in the DM speed distribution.

Because the signatures of annual modulation, gravitational focusing, and velocity substructure are necessarily time-dependent, we are forced to promote our likelihood to incorporate variation in time.[19] To do so, we will make use of the stacking procedure described in Sec. 3.3. We assume that the full dataset is broken into $N_T$ subintervals of duration $\Delta T = T/N_T$ containing $\Delta N = N/N_T$ PSD measurements. Now, however, we will assume that $\Delta T$ is sufficiently small that the speed distribution does not change appreciably within a given interval. As the distribution will change

[19] Cold velocity substructure is more subject to annual and daily modulation, which is why these effects are time-dependent in the Earth frame even if they are not in the Solar frame.

over the full collection time $T$, we have a different model prediction in each time interval given by:

$$\lambda_{k,\ell} = A\frac{\pi f(v, t_\ell)}{m_a v} + \lambda_B \,, \tag{3.78}$$

which leads to the following modified likelihood

$$\mathcal{L}(d|\boldsymbol{\theta}) = \prod_{\ell=0}^{N_T-1} \prod_{k=1}^{\Delta N-1} \frac{1}{\lambda_{k,\ell}(\boldsymbol{\theta})} e^{-S_{\Phi\Phi}^{k,\ell}/\lambda_{k,\ell}(\boldsymbol{\theta})} \,. \tag{3.79}$$

This is the form of the likelihood we will use throughout this section. Note that the $\ell$ dependence on the model prediction invalidates the stacking analysis performed in Sec. 3.3, though the data may still be stacked over time intervals that are sufficiently smaller than a year (day) for annual (daily) modulation.

### 3.5.1 Halo Annual Modulation

Before studying how annual modulation impacts the expected axion signal, we first review how it modifies the DM speed distribution.[20] Our starting point for this is the SHM distribution given in (3.9). Throughout the year the detector's speed in the Galactic halo frame, $v_{\rm obs}$, is expected to oscillate as the Earth orbits the Sun. In the lab frame, this results in an effectively time-dependent halo distribution $f_{\rm SHM}(v, t)$. All of the time dependence, neglecting that from gravitational focusing, which will be dealt with separately, can be accounted for by upgrading the relative detector-halo speed to a time-dependent parameter $v_{\rm obs}(t)$. To determine this speed, first note that $\mathbf{v}_{\rm obs}(t) = \mathbf{v}_\odot + \mathbf{v}_\oplus(t)$, where $\mathbf{v}_\odot$ and $\mathbf{v}_\oplus(t)$ are the velocity of the Sun with respect to the Galactic frame and the velocity of the Earth with respect to the Sun, respectively. These are specified by[21]

$$\begin{aligned} \mathbf{v}_\odot &= v_\odot(0.0473, 0.9984, 0.0301) \,, \\ \mathbf{v}_\oplus(t) &\approx v_\oplus \left(\cos\left[\omega(t-t_1)\right]\hat{\boldsymbol{\epsilon}}_1 + \sin\left[\omega(t-t_1)\right]\hat{\boldsymbol{\epsilon}}_2\right) \,, \end{aligned} \tag{3.80}$$

where the magnitudes are given by $v_\odot \approx 232.37$ km/s and $v_\oplus \approx 29.79$ km/s. We have further introduced $\omega \approx 2\pi/(365 \text{ days})$ as the period of the Earth's revolution, $t_1$ as the time of the vernal equinox (which occurred on March 20 in 2017), and the unit vectors $\hat{\boldsymbol{\epsilon}}_1$ and $\hat{\boldsymbol{\epsilon}}_2$ specifying the

[20] We refer to [201] for a comprehensive review of these details.

[21] Corrections to $\mathbf{v}_\oplus(t)$ are suppressed by the eccentricity of the Earth's orbit, given by $e \approx 0.016722$, and so can safely be neglected.

ecliptic plane. These vectors are given in Galactic coordinates by

$$\begin{aligned}\hat{\boldsymbol{\epsilon}}_1 &\approx (0.9940, 0.1095, 0.0031)\,,\\ \hat{\boldsymbol{\epsilon}}_2 &\approx (-0.0517, 0.4945, -0.8677)\,.\end{aligned} \tag{3.81}$$

We may then find the time-varying Galactic-frame speed

$$v_{\rm obs}(t) = \sqrt{v_\odot^2 + v_\oplus^2 + 2v_\odot v_\oplus \alpha \cos\left[\omega(t-\bar{t})\right]}\,, \tag{3.82}$$

given in terms of the parameters

$$\begin{aligned}\alpha &\equiv \sqrt{(\hat{\mathbf{v}}_\odot \cdot \hat{\boldsymbol{\epsilon}}_1)^2 + (\hat{\mathbf{v}}_\odot \cdot \hat{\boldsymbol{\epsilon}}_2)^2} \approx 0.491\,,\\ \bar{t} &\equiv t_1 + \frac{1}{\omega}\arctan\left(\frac{\hat{\mathbf{v}}_\odot \cdot \hat{\boldsymbol{\epsilon}}_2}{\hat{\mathbf{v}}_\odot \cdot \hat{\boldsymbol{\epsilon}}_1}\right) \approx t_1 + 72.5\ \text{days}\,.\end{aligned} \tag{3.83}$$

Whilst we have given the accepted values for the various parameters above, if a definitive axion signal was detected we could then take for example $v_\odot$, $\alpha$, and $\bar{t}$ as unknown parameters to be estimated from the likelihood. Their agreement with the accepted values would be a highly non-trivial test of the signal. We will show an example of this below, but before doing so we use the Asimov formalism to estimate how significant a signal we would need to detect annual modulation from the bulk halo.

Ignoring annual modulation, the detection significance of a SHM signal scales with the parameters of interest as

$$\text{TS} = \frac{A^2 T\pi}{2m_a\lambda_B^2}\frac{\text{erf}\left[\sqrt{2}v_{\rm obs}/v_0\right]}{\sqrt{2\pi}v_0 v_{\rm obs}}\,, \tag{3.84}$$

where here and throughout this section we assume the background is frequency independent over the width of the signal. The relevant question is, on average, at what value of TS do we detect annual modulation at a given significance? To estimate this, we calculate the test statistics between models with and without annual modulation included; in order to discover annual modulation we can think of the model without it included as the null hypothesis. We denote this test statistic by $\text{TS}_{\text{a.m.}}$. We can estimate the median value for $\text{TS}_{\text{a.m.}}$ as a function of the model parameters using the asymptotic form of $\Theta$ and the Asimov formalism; in this case, the Asimov dataset includes annual modulation. Specifically, we find

$$\text{TS}_{\text{a.m.}} = \frac{A^2 T\pi}{m_a\lambda_B^2}\int\frac{dv}{v}\left[f_t(v)^2 - f(v)\left(f_t(v) - \frac{f(v)}{2}\right)\right]\,. \tag{3.85}$$

Above, $f_t$ features annual modulation while $f$ does not. In order to simplify the calculation, we

define an expansion parameter:

$$\epsilon \equiv \frac{v_\odot v_\oplus}{v_\odot^2 + v_\oplus^2} \approx 0.126\,, \tag{3.86}$$

in terms of which we can write:

$$v_{\rm obs}(t) \approx v_{\rm obs}\left(1 + \epsilon\alpha\cos\left[\omega(t-\bar{t})\right]\right)\,, \tag{3.87}$$

with $v_{\rm obs} \approx 232$ km/s. Using this and averaging all time dependence over one period in the final result, we calculate the ratio of $\text{TS}_{\text{a.m.}}$ to TS in the SHM as

$$\begin{aligned}\frac{\text{TS}_{\text{a.m.}}}{\text{TS}} &= \frac{\alpha^2\epsilon^2 v_{\rm obs}^2}{2v_0^2}\left(1 - \frac{4v_{\rm obs}e^{-2v_{\rm obs}^2/v_0^2}}{\sqrt{2\pi}v_0 \text{erf}\left[\sqrt{2}v_{\rm obs}/v_0\right]}\right) \\ &\approx 0.00173\,.\end{aligned} \tag{3.88}$$

From the discussion above, we see that if it took a time $T$ to detect the axion at a given significance, it would take a time $580T$ to detect annual modulation at the same significance. Alternatively, as the test statistic scales like $g_{\alpha\gamma\gamma}^4$, the coupling for the threshold of discovery for annual modulation will be $\sim$5 times larger, on average, than the coupling for the threshold of discovery of a signal. On the other hand, in the resonant setup large increases in the TS are readily obtainable since after the axion mass is known we can stay at the correct frequency for an extended period instead of scanning over multiple frequencies.

In Fig. 3.5 we show the posterior distribution generated in a Bayesian framework from an analysis of the Asimov dataset with $g_{a\gamma\gamma}$ at the threshold for detection of annual modulation at $5\sigma$. Note that we float $A$, $m_a$, $v_0$, $v_\odot$, $\alpha$, and $\bar{t}$ as model parameters with linear-flat priors in the fit. All model parameters are seen to be well converged, including $m_a$ which is not shown in the figure. This analysis was performed using `Multinest` [202, 203] with 500 live points. The Asimov results are consistent with those found from an ensemble of simulated datasets, as expected.

### 3.5.2 Halo Gravitational Focusing

An additional source of annual modulation in the axion signal is sourced by the focusing of the axion flux by the Sun's gravitational potential. This effect is already known to have a significant impact on annual modulation in the context of WIMP direct detection, as pointed out in [184]. The intuition behind gravitational focusing is that in the frame of the Sun the DM velocity distribution appears as a wind. The gravitational field of the Sun focuses the DM "down-wind" of the Sun, leading to an enhanced rate when the Earth is "down-wind" relative to when the Earth is "up-wind." Here we investigate the impact of gravitational focusing on the corresponding axion signal.

In [184] an exact closed-form expression was used to model the perturbation to the DM phase-space distribution from the Sun's potential. The perturbed phase-space distribution is derived using Liouville's theorem and exactly solving for the trajectories of the DM particles in the gravitational field. However, in this work we take advantage of a perturbative result (to leading order in Newton's constant), valid when the DM speeds are much larger than the Solar escape velocity, that allows us to write [204]

$$f(v,t) = f_{\rm halo}(v,t) + f_{\rm GF}(v,t)\,, \tag{3.89}$$

where $f_{\rm halo}(v,t)$ is the unperturbed velocity distribution in the Earth frame, and where the perturbation by gravitational focusing $f_{\rm GF}$ is given by

$$f_{\rm GF}(v,t) \equiv -\frac{2GM_\odot}{x_\oplus(t)} \int \frac{v^2 d\Omega}{\pi^{\frac{3}{2}} v_0^5} \frac{e^{-(\mathbf{v}+\mathbf{v}_\oplus(t)+\mathbf{v}_\odot)^2/v_0^2}}{v} \times \frac{(\mathbf{v}+\mathbf{v}_\oplus(t)+\mathbf{v}_\odot)\cdot\left(\hat{\mathbf{x}}_\oplus(t) - \frac{\mathbf{v}+\mathbf{v}_\oplus(t)}{|\mathbf{v}+\mathbf{v}_\oplus(t)|}\right)}{1-\hat{\mathbf{x}}_\oplus(t)\cdot\left(\frac{\mathbf{v}+\mathbf{v}_\oplus(t)}{|\mathbf{v}+\mathbf{v}_\oplus(t)|}\right)}\,. \tag{3.90}$$

Note that in this equation, $v^2 d\Omega$ is written out explicitly to account for the measure. Here, $\mathbf{x}_\oplus(t)$ denotes the position of the Earth in the Solar frame; an explicit form for this in Galactic coordinates can be found in [201]. Note that $f(v,t)$ is no longer normalized to integrate to unity, but rather the change in $\int dv f(v,t)$ throughout the year indicates the fractional change in the DM density do to gravitational focusing. We have explicitly verified that the perturbative formalism for gravitational focusing is a good approximation to the exact formalism used in [184] for the SHM.

To determine the impact of gravitational focusing, we perform two analyses using the Asimov dataset at the $5\sigma$ detection threshold for annual modulation but this time including gravitational focusing. We analyze the Asimov data in the Bayesian framework including with two models; the first model does not account for gravitational focusing, while the second one does. The results of these analyses are shown in Fig. 3.6. The use of a limited number of live points is the most likely source of the residual disagreement between the injected and median value of $\bar{t}$ in the right panel. Note that in these analyses we only float $A$, $\alpha$, and $\bar{t}$ for simplicity. Neglecting gravitational focusing in the model (left panel) only leads to a approximately $2\sigma$ overestimate in the value of the $A$ parameter, while the central value of $\bar{t}$ is on average off by $\sim$10 days. On the other hand, when gravitational focusing is included in the model (right panel), the halo parameters and the normalization are correctly inferred.

### 3.5.3 Local DM Substructure

So far, we have only considered an axion signal sourced by dark matter contained within the bulk halo, but there additionally exist a number of well-motivated classes of velocity substructure

that have the potential to leave dramatic signatures in the direct detection data. One large class of substructure relates to the DM subhalos that are expected to be present in the Milky Way [205]. DM subhalos are believe to persist down to very small mass scales, potentially $\sim 10^{-6}\ M_\odot$ and below, due to the nearly scale-invariant spectrum of density perturbations generated during inflation. Low-mass DM subhalos have low velocity dispersions, and so if we happen to be sitting in a DM subhalo, even if it only makes up a small fraction of the local DM density, it could show up as a narrow spike in velocity space over the bulk SHM contribution. Even if we are not directly in a bound DM subhalo, we could still be affected by the tidally stripped debris that in-falling subhalos leave throughout the Galaxy. There are two types of tidally-stripped substructure, in velocity space, that are important for direct detection (for a review of the importance of tidal debris at WIMP experiments, see [193]): DM streams and debris flows.

As an in-falling subhalo descends through the potential of the Milky Way, the outer regions of the DM subhalo are expected to become tidally stripped and form an ultra-cold trailing stream [194, 205]. Such streams should trail from DM subhalos of all sizes, with smaller subhalos having colder streams. Eventually, the tidal debris dragged away from in-falling subhalos will become fully virialized. However, before that occurs the debris becomes homogeneously distributed in position space but remains coherent in velocity space, forming the substructure known as debris flow [206]. While it is unlikely that a DM substructure from in-falling subhalos dominates the local DM density [194, 205], as we show in this subsection, even if the substructure only makes up a small fraction of the local DM density, due to the coherence in velocity space the signature of substructure at axion experiments can be substantial and even dominate over the SHM contribution. This can be contrasted to the case in WIMP direct detection experiments, where substructure is expected to play an important role in annual modulation studies but not necessarily have a significant impact on the total rate [193, 201, 207]. DM streams were recently considered in the context of axion direct detection in [183].

One DM stream in particular has received a significant amount of attention with regards to WIMP direct detection and that is the potential DM component of the Sagittarius stream. The Sagittarius stream consists of a winding stream of stars wrapping through the Milky Way that is thought to have formed from tidal stripping of the Sagittarius dwarf galaxy. It is possible that the DM component of the Sagittarius stream contributes at the few percent level to the local DM density (see, *e.g.*, [194,205]). We follow [173, 175, 176] and model the stream as a boosted Maxwellian distribution with a narrow velocity dispersion of $v_0 = 10$ km/s and a stream velocity of $v_{\rm str} = (0, 93.2, -388)$ km/s, in Galactic coordinates. Further we assume that the Sagittarius stream constitutes 5% of the local DM. We will show that even though the stream may only be a small component of the local DM density, it can still leave an important signature in axion direct detection experiments, due to its small velocity dispersion.

Another possible source of DM substructure that has low velocity dispersion is a dark disk. Co-rotating thick dark disks are found to form in certain $N$-body simulations with baryons [208–211] due to the disruption of merging satellites galaxies that are pulled into the disk. In the simulations, the dark disks are found to be co-rotating with lag speeds and velocity dispersions both $\sim 50$ km/s. They may even dominate the local DM density [208, 210]; however, as we will see, even if the dark disk is only a small fraction of the local DM density, it can still leave a significant signature in the direct detection data due to the small velocity dispersion and lag speed.

To develop some intuition for how important substructure could be, let us take the oversimplified scenario in which the substructure of interest makes up a fraction $x$ of the local DM distribution and also follows the Maxwellian distribution with the same $v_{\rm obs}$ as in the SHM, but with a much smaller dispersion parameter $v_0^{\rm str}$. Then we can write

$$f(v) = (1-x) f_{\rm SHM}(v|v_0, v_{\rm obs}) + x f_{\rm SHM}(v|v_0^{\rm str}, v_{\rm obs})\,. \tag{3.91}$$

Using this we can explicitly calculate the expected test statistic (in favor of the model of the SHM plus the stream over the null hypothesis of no DM) of a signal with a frequency independent background, though we do not provide the expression here as its cumbersome form is not particularly informative. In Fig. 3.7 we show this TS plotted as a function of the fraction of the DM in the stream $x$ for various values of $v_0^{\rm str}$, normalized to the TS when no stream is present. The figure makes it clear that if the detector is within an ultra-cold DM stream the impact on the expected axion signal can be significant, even if the stream only makes up a small fraction of the DM. For example, if 5% of the local DM is in a stream with $v_0^{\rm str} \approx 0.1$ km/s, then the TS in favor of the model with DM is nearly 10 times larger when the stream is modeled versus when it is not. This emphasizes the importance of searching for cold DM substructure in addition to the SHM component.

Even though velocity substructures are not intrinsically time-dependent features, annual modulation is considerably more important for the detection of substructure, which is typically characterized by a speed dispersion less than the peak-to-peak variation of the Earth's velocity with respect to a given substructure frame. The result is an observational signature of a given substructure feature poorly localized in frequency data collected over a year. Therefore we need a more careful treatment than the one above, as we can only search for these features in a model framework which accounts for time-varying signals.

Under the assumption that velocity substructure can still be reasonably modeled by a boosted Maxwellian distribution, it is easily accommodated within our time-dependent model template.[22] The direction of the stream in the ecliptic plane is specified through the parameters $\alpha^{\rm sub}$ and $\bar{t}^{\rm sub}$,

[22] Even if the velocity distribution is not Maxwellian, the relevant signal template is a straightforward generalization of that presented here for a Maxwellian.

which are defined in analogy to (3.83) but where $\mathbf{v}_\odot^{\rm sub} = v_\odot^{\rm sub}\hat{\mathbf{v}}_\odot^{\rm sub}$ is the stream boost velocity in the Solar frame. The generalized velocity distribution, including gravitational focusing, for both the SHM and the substructure components is then given by

$$f = (1-x)f^{\rm SHM}(v|v_\odot, \alpha, \bar{t}, v_0) + xf^{\rm sub}(v|v_\odot^{\rm sub}, \alpha^{\rm sub}, \bar{t}^{\rm sub}, v_0^{\rm sub})\,, \tag{3.92}$$

where the superscripts "sub" and "SHM" denote the generalized substructure and SHM velocity distributions, respectively, after gravitational focusing has been accounted for. The generalization to multiple substructure components is straightforward.

The importance of annual modulation for cold substructure is illustrated in Fig. 3.8, where we show, in the left panel, the mean PSD assuming the Sagittarius stream parameters taken at two different times throughout the year. We have chosen the dates where the TS in favor of the stream is maximized, June 5, and minimized, November 23, both for 2017. Since the stream is narrow in frequency space, the sharp peaks at these two different times of year are almost completely non-overlapping. On the contrary, at frequencies where the stream does not contribute appreciably, annual modulation does not significantly affect the contribution from the SHM.

Just as we performed parameter estimations for the bulk halo component, we can also estimate the parameters defining the contribution of velocity substructure to the speed distribution. It should be noted that the parameter estimation for the bulk halo component can be substantially affected by the presence of velocity substructure if the substructure is not properly accounted for. An example of this can be seen in Fig. 3.9, where we have included a stream with Sagittarius-like parameters in the data, as given earlier, and used the Asimov dataset. However, we have not accounted for the stream in the model that is fit to the data. Note that the TS in favor of DM in this case is $\sim 10^4$. Our estimates for the SHM parameters $v_0$ and $v_\odot$ are significantly affected by the presence of the stream and disagree with the true values by multiple standard deviations. In contrast, in Fig. 3.10 we display the posterior distribution for a fit including a Maxwellian stream. Note that while both the SHM and the stream parameters are floated at the same time, we display the posteriors for the SHM and stream model parameters separately. In this case both the stream and the SHM model parameters are accurately estimated. Comparing the model that included the stream to that without, we find a TS value $\sim$400 in favor of the model with the stream over that without.[23]

Note that for our fiducial set of model parameters for the Sagittarius stream, we find that when the SHM is detected at $5\sigma$ significance (TS $\sim$ 58), including the look elsewhere effect, the stream may barely start to become visible at $\sim 1.6\sigma$ significance. We stress, however, that is possible that

[23]To simplify the analysis, we have neglected gravitational focusing in considering this Sagittarius-like stream. Gravitational focusing is more important at lower speeds, and therefore is generally less relevant for such a stream than it would be in considering, for example, a dark disk. We note, however, that if the stream is well-aligned with the ecliptic plane, it is possible to get large enhancements to the rate over short periods of time during the year [212–214], although such a configuration is not present for the Sagittarius stream.

other, colder DM streams would contribute more substantially even if they are a smaller fraction of the local DM density. While we illustrated the stream example for simplicity, the effects of the other types of velocity substructure may be worked out similarly. For example, we find that with our fiducial choice of parameters for the dark disk lag speed and velocity dispersion, the dark disk would be detectable at the same significance as the SHM even if the dark disk only makes up $\sim$20% of the local DM density. Moreover, the dark disk should be more affected by annual modulation and gravitational focusing than the SHM component, since the DM in the dark disk is on average slower moving in the Solar frame. The PSD template is illustrated, assuming the dark disk makes up 20% of the local DM density, in the right panel of Fig. 3.8. The dark disk leads to a significant increase in the PSD at low velocities, corresponding to frequencies near the axion mass. As in the stream case, we show the PSD at two different times of year, corresponding to the date of maximal TS, November 18, and minimal TS, June 5.

## 3.6 Conclusion

The QCD axion, and axion like particles more generally, is a well motivated class of DM candidates, and if it constitutes the DM of our universe, then the burgeoning experimental program searching for such DM could be on the verge of a discovery. With such possibilities it is important to be able to clearly and accurately quantify any emerging signal and set limits in their absence. The likelihood framework we have introduced allows for exactly this. In addition, through the use of the Asimov dataset, we have derived a number of analytic results that make quantifying these thresholds possible without recourse to Monte Carlo simulations.

In the event of an emerging signal, one would always worry about the possibility of unanticipated backgrounds. Nevertheless DM provides its own way of addressing this concern through unique fingerprints in the frequency and time domains. For example, we showed the form the local DM velocity distribution uniquely determines the frequency dependence of the PSD data, and that by exploiting this knowledge one is able to, through the likelihood framework, constrain properties of the local velocity distribution. Since the bulk of the DM halo is expected, locally, to follow a Maxwellian distribution with velocity dispersion set by the local rotation speed, correctly measuring the Maxwellian parameters will provide a non-trivial check of the nature of the signal. In the time domain, any true signal should undergo annual modulation, including the subtle effect of gravitational focusing, and we quantified how this may be verified using the likelihood formalism. Further, the likelihood is sensitive to the presence of local DM substructure such as cold streams, which can enhance the expected signal through an associated increase in the axion coherence time. For example, we showed that the Sagittarius stream could leave a unique signature in the PSD data. Nevertheless there are a great many possible types of DM substructure, beyond those considered

here, that could be present at the position of the Earth, and we leave a careful study of these to future work.

Taken together the results of this work provide a set of tools that will prove useful in moving towards a possible DM axion detection, and, if we should be so lucky, into the era of axion astronomy that would follow. Towards that end, we have provided an open-source code package at https://github.com/bsafdi/AxiScan for performing all the likelihood analyses discussed in this work and also simulating data at axion direct detection experiments for different background and signal models.

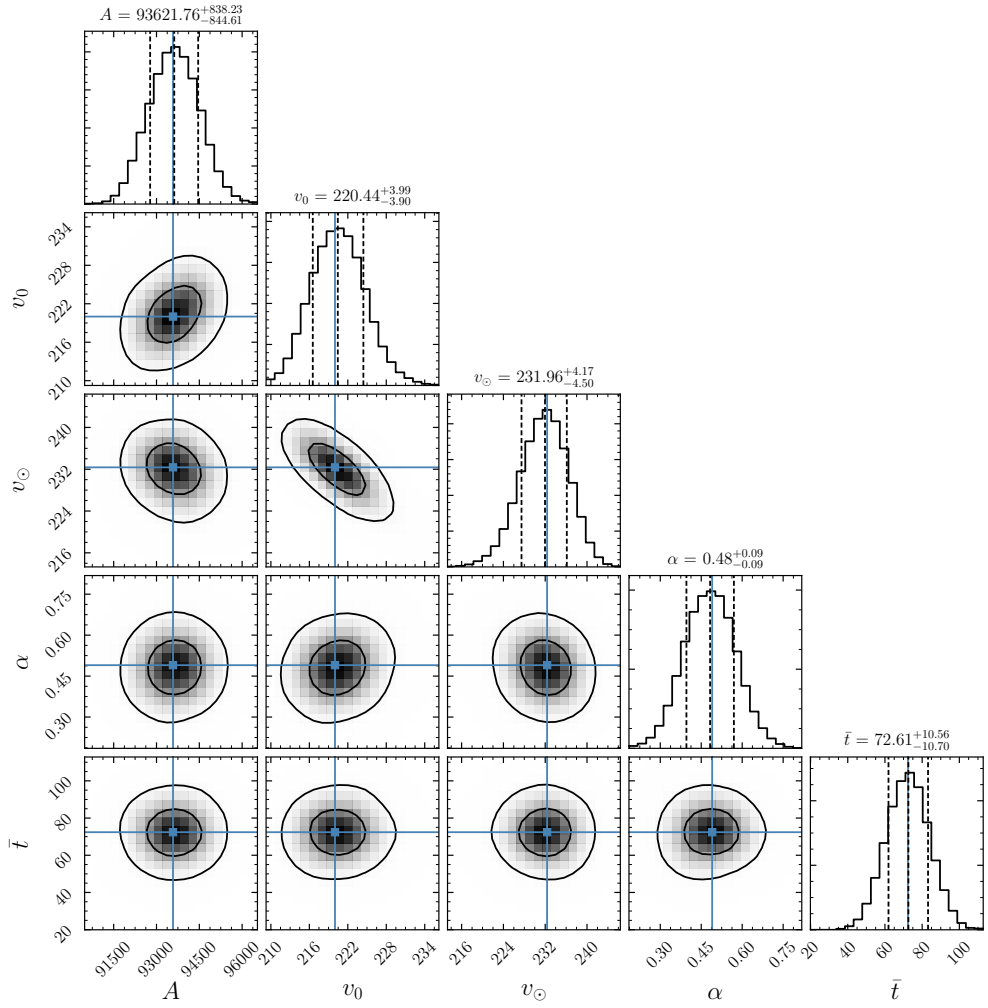



Figure 3.5: The posterior distribution for a model with annual modulation where the signal strength is at the threshold of annual modulation detection at $5\sigma$. The true parameter values are indicated in blue, with the $1\sigma$ confidence intervals on the parameter estimations indicated by the dashed black lines in the one parameter posteriors. The two parameter posteriors show the 1 and $2\sigma$ contours. The axion mass, $m_a$, was also scanned over, and is recovered accurately but not shown here. Note that this example uses the Asimov dataset. All times are measured in days and velocities in km/s, while the units of $A$ are arbitrary.

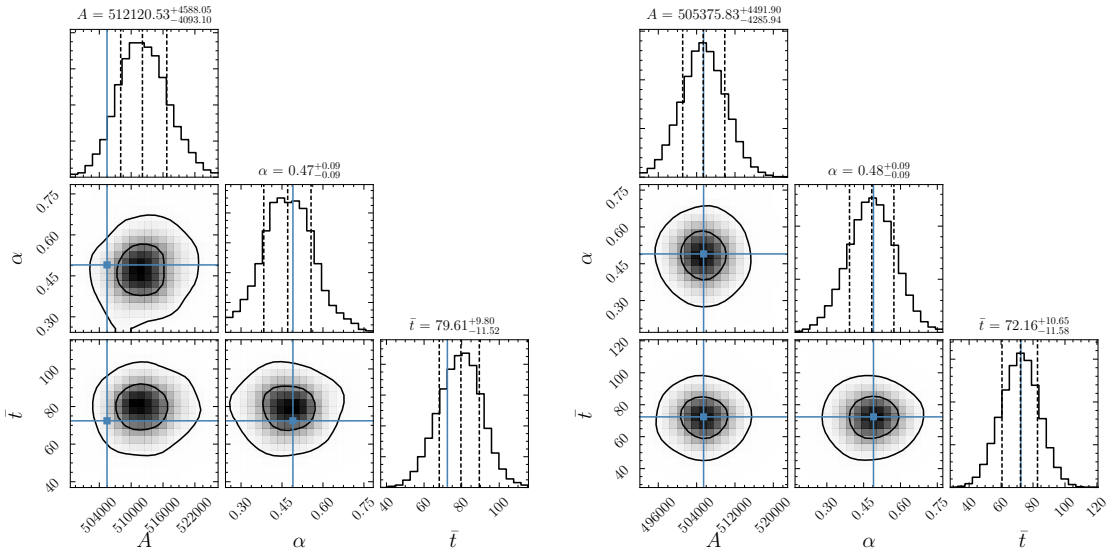



Figure 3.6: As in Fig. 3.5, but this time the data includes gravitational focusing and the model only includes the parameters $A$, $\alpha$ and $\bar{t}$. (Left) Gravitational focusing, while present in the Asimov data, is excluded from the model template. The estimations of $A$ and $\bar{t}$ are off at the $\sim 2\sigma$ and $\sim 1\sigma$ levels, respectively. (Right) As in the left panel but including gravitational focusing in the model template. As expected, the parameter estimation is quite accurate in this case.

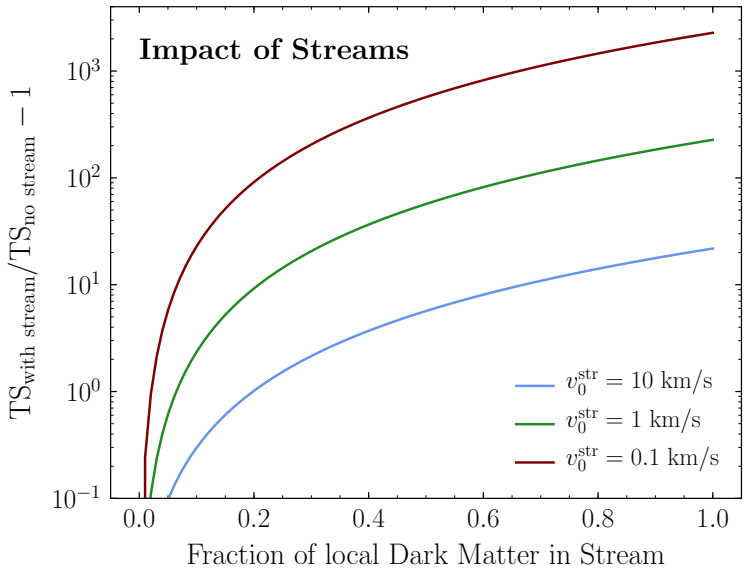



Figure 3.7: The enhancement expected in the TS in the presence of a coherent DM stream. The TS is shown as a ratio with respect to the case where only the bulk halo is present and as a function of the fraction of the local DM within the substructure.

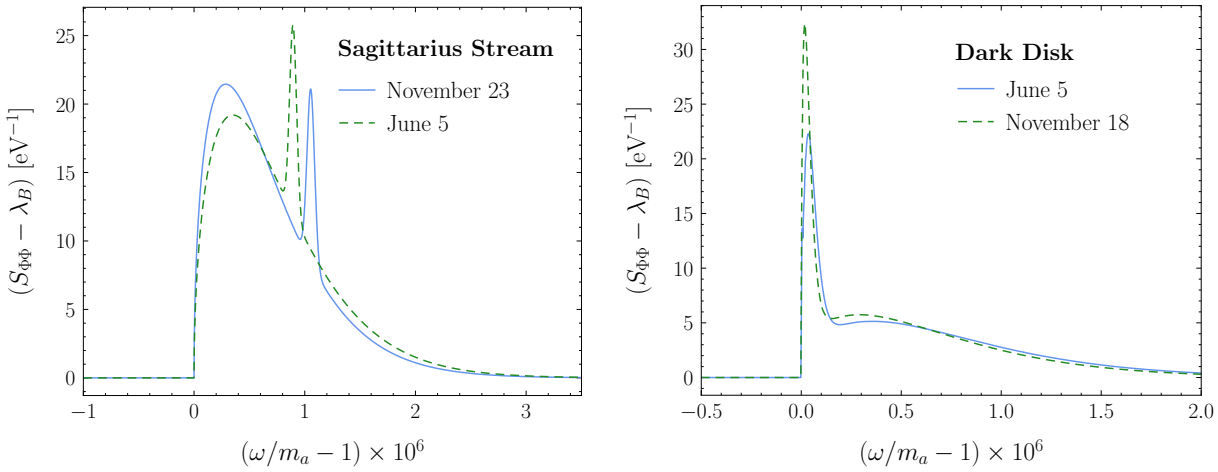



Figure 3.8: The axion contribution to the PSD as a function of frequency in the presence of DM substructure. (Left) We show the effect of a Sagittarius-like stream that makes up $\sim$5% of the local DM density at two different times of year, corresponding to the dates of maximum TS (June 5) and minimum TS (November 23), where all dates are for 2017. Annual modulation plays an important role for cold substructure because the Earth's orbital velocity may be larger than the substructure velocity dispersion. (Right) As in the left panel, but for a dark disk that makes up $\sim$20% of the local DM density. The dark-disk is co-rotating with the baryonic disk, with a lag speed $\sim$50 km/s, and so the contribution to the PSD is at lower speeds compared to the stream case. Gravitational focusing also plays an important role for the disk since the solar-frame velocities are relatively low. In this case the maximum and minimum TS occur on November 18 and June 5 respectively. For both of these panels, the signal is generated using $m_a = 1$ MHz, $A$ set to the value for the threshold for detection of the SHM, and $\lambda_B$ set to the minimum SQUID noise.

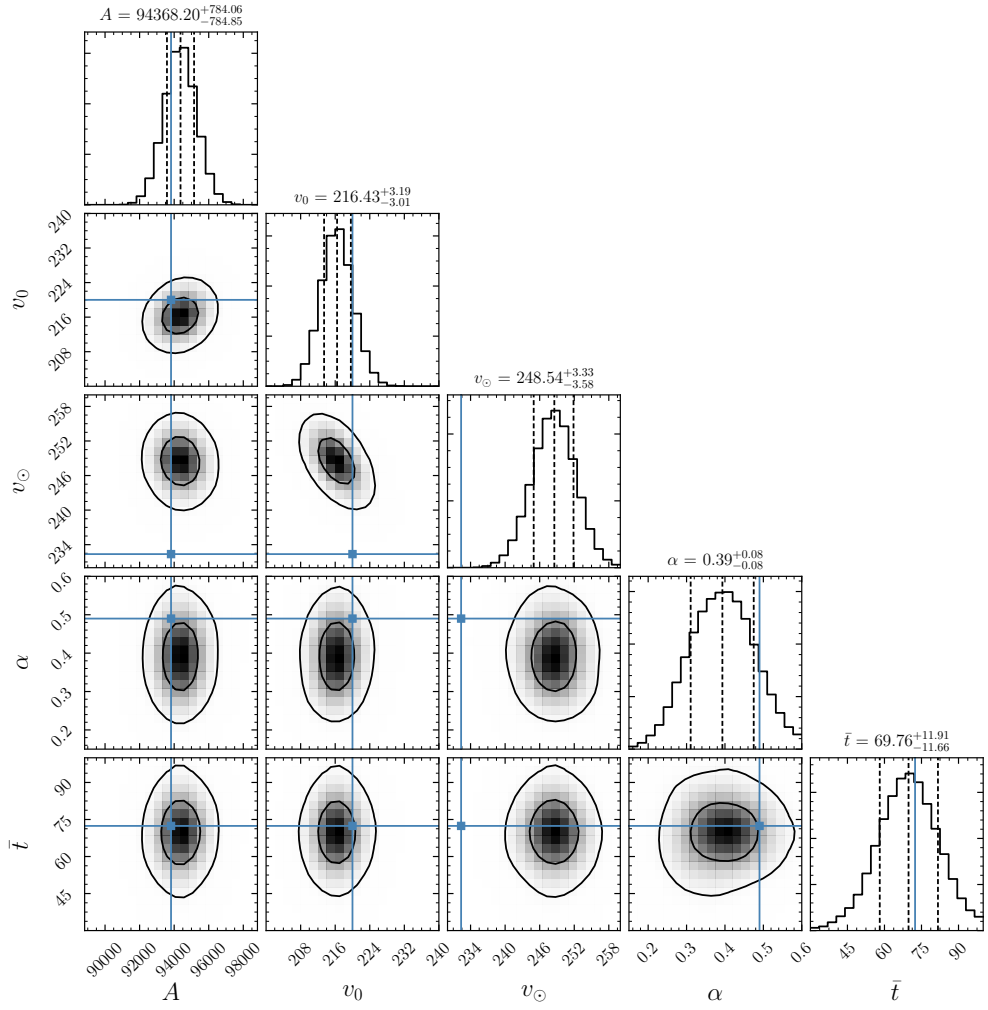



Figure 3.9: A Monte Carlo parameter estimation for the bulk halo parameters at the threshold of detection for annual modulation in the presence of a Sagittarius-like stream containing 5% of the DM and with a narrow velocity dispersion of 10 km/s. The accuracy of the parameter scan is worsened by the failure to account for the substructure in the analysis.

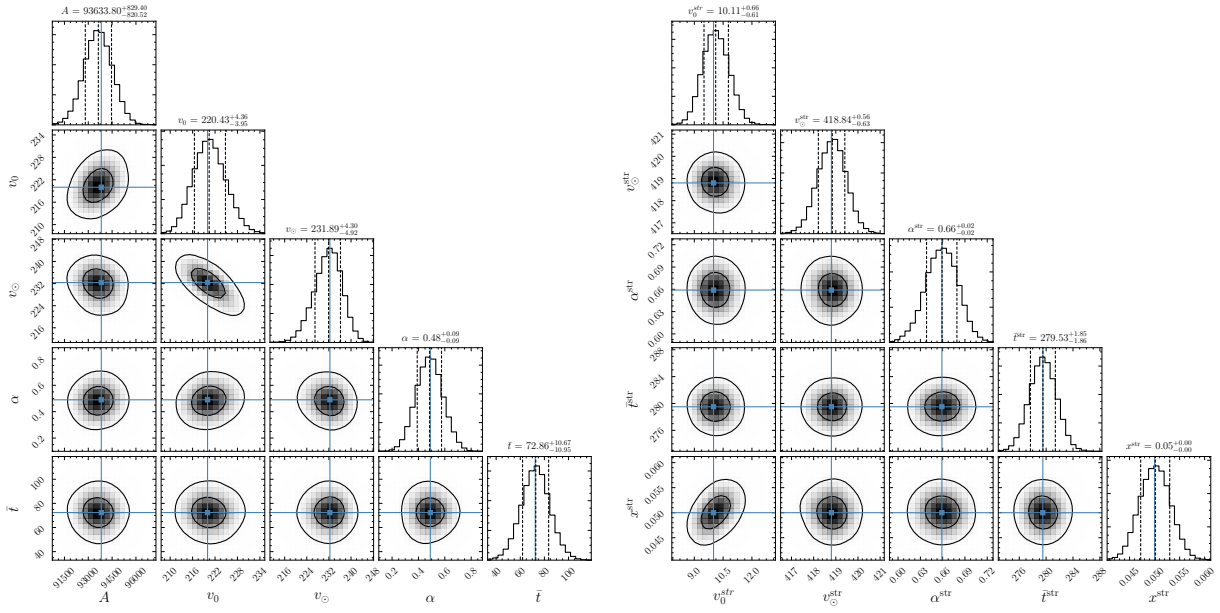


Figure 3.10: A simultaneous Monte Carlo parameter estimation for a signal containing a bulk halo and a Sagittarius-like stream with 5% of the DM using identical seed parameters as Fig. 3.9. Scanning for the bulk halo and substructure simultaneously allows us to accurately recover the signal parameters. Left, the bulk parameter scan results, right, the stream parameter scan results.

# CHAPTER 4

# The Statistics of Axion Direct Detection with Multiple Detectors

Cold, bosonic dark matter (DM) candidates with masses much smaller than the eV scale have macroscopic occupation numbers and may be described in the solar vicinity by classical fields. Two well-studied DM candidates in this category, which we broadly refer to as wave-like DM, are axions [38–46, 86, 87] and dark photons [215–217]. Wave-like DM candidates require distinctive experimental techniques for discovery that take advantage of their spatial and temporal coherence (see, *e.g.*, [187]). The spatial coherence length of the DM waves, $\lambda_c$, and the coherence time $\tau$ are given by[1] [80]

$$\lambda_c \sim \frac{1}{m_{\rm DM} v_0}\,, \qquad \tau \sim \frac{1}{m_{\rm DM} \bar{v}\, v_0}\,, \tag{4.1}$$

where $v_0$ parameterizes the DM velocity dispersion, $\bar{v}$ is the mean velocity, and $m_{\rm DM}$ is the DM mass. In the solar neighborhood we expect $\bar{v} \sim v_0 \sim 10^{-3}$ for the bulk of the DM, in natural units, such that the coherence length is around $10^3$ times the Compton wavelength, and the coherence time is around $10^6$ times the oscillation period for the DM wave. In this work we show that multiple phase-sensitive wave-like DM detectors separated by distances of order $\lambda_c$ may join their data – through a process we refer to as "DM interferometry" – to measure properties of the DM phase-space distribution that are inaccessible to single experiments operating in isolation.

Many axion and dark-photon detection strategies already leverage the axion coherence *time* as a "quality factor" that amplifies the DM signal in the experiment. For example, axion haloscopes [67, 68, 110, 138, 190, 218] use a resonant cavity with a strong static magnetic field to convert axion DM into electromagnetic cavity modes, which build up coherently over the DM coherence time; in this setup the DM Compton wavelength is of order the size of the experiment. Experiments operating in the quasistatic regime (where the DM Compton wavelength is much larger than the experiment) – including searches for the axion-photon coupling [71, 72, 74, 75, 82, 219], axion interactions with

[1] We briefly review both concepts in App. C.1.

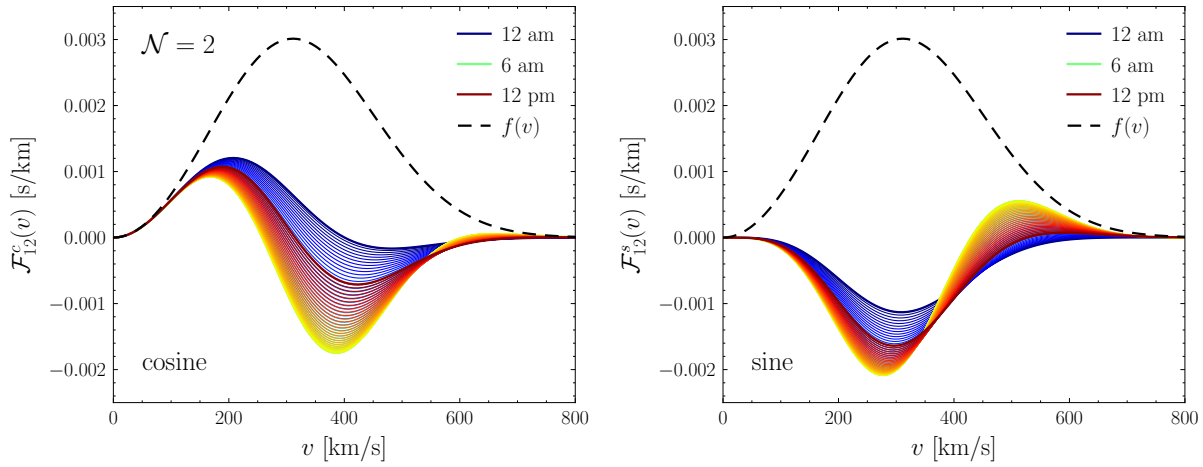



Figure 4.1: The imprint of DM interferometry. A single wave-like DM experiment is sensitive to the DM speed distribution $f(v)$. Two detectors separated by a vector $\mathbf{x}_{12}$, however, are sensitive to the speed distribution modulated by the $\mathbf{k}\cdot\mathbf{x}_{12}$ phase of the DM wave, replacing $f(v)$ with functions $\mathcal{F}^{c,s}_{12}(v)$ as defined in (4.3). As the figures demonstrate, the modified speed distributions exhibit daily modulation and carry additional information about the velocity distribution $f(\mathbf{v})$ that would be invisible to a single detector. For this example we take $m_{\rm DM} = 25.2\ \mu$eV [79], near the window where the HAYSTAC collaboration is searching for axion DM. Taking the Standard Halo Model ansatz for $f(\mathbf{v})$ in (4.50), we place one detector at a latitude and longitude of $(41^\circ\ \text{N},\ 73^\circ\ \text{W})$, and a second instrument $\sim 20$ m to the North, corresponding to $d \sim 2\lambda_c$. A curve is shown for every ten minutes starting from midnight on January 1st of 2020. Note that as $\mathcal{F}^{c,s}_{12}(v)$ are functions of $m_{\rm DM}d$, qualitatively similar effects exist for *e.g.* $m_{\rm DM} \sim 10^{-9}$ eV, in the mass range probed by ABRACADABRA and DM-Radio, for $d \sim 500$ km.

nuclear spins [58], or dark photons [111] – aim to detect a time-varying magnetic flux through a pickup loop, which can build up coherently in a lumped-element circuit [220, 221].

In this paper, we explore the phenomenology of spatial coherence for wave-like DM by exploiting spatially-separated detectors that probe the same DM field. It is straightforward to understand why multiple detectors offer unique insights for wave-like DM. Generically, the wave-like DM field may be written as $a(\mathbf{x}, t) = a_0 \cos(\omega t - \mathbf{k}\cdot\mathbf{x} + \phi)$, where $\omega$ is the oscillation frequency, $\mathbf{k}$ is the wave vector, $\phi$ is a random phase, and $a_0$ is the amplitude.[2] If the DM wave is traveling in the direction $\hat{\mathbf{k}}$ with speed $v \ll 1$, then $\omega \approx m_{\rm DM}(1 + v^2/2)$ and $\mathbf{k} \approx m_{\rm DM} v\hat{\mathbf{k}}$. For a single detector we may always choose coordinates such that $\mathbf{x} = 0$. This means that a single detector is only sensitive to the speed through $\omega$ and is not sensitive to the direction of the DM velocity.[3] By contrast, two

[2] Vector DM also has a polarization component with nontrivial coherence properties, but in this work we focus only on the amplitude, as appropriate for scalar or pseudoscalar DM.

[3] Exceptions would be experiments that make use of $\nabla a$, but such signals are suppressed by $v \sim 10^{-3}$ relative to experiments that are also sensitive to $\partial_t a$. Experiments only sensitive to the speed distribution may also detect annual modulation signals through shifts in the DM speed [183], though these are typically quite small because the Earth's speed relative to the Sun is small compared to the solar speed relative to the Galactic Center.

experiments located at positions $\mathbf{x}_1$ and $\mathbf{x}_2$ will be sensitive to phase factors $\mathbf{k}\cdot\mathbf{x}_1$ and $\mathbf{k}\cdot\mathbf{x}_2$. Only one of these can be removed by a coordinate choice, leaving a residual $\mathbf{k}\cdot\mathbf{x}_{12}$, with $\mathbf{x}_{12} = \mathbf{x}_1 - \mathbf{x}_2$, which manifestly probes the velocity rather than the speed. The interferometry proposed in this work is directly at the level of the DM field: the effect arises due to the phase difference wave-like DM exhibits between spatially separated points.[4] Indeed, due to the nonzero velocity dispersion $v_0$, DM waves are coherent up to distances of order $\lambda_c$; as we will show, phase-sensitive data combined from two experiments exhibits maximal modulation when $d \equiv |\mathbf{x}_{12}| \sim \lambda_c$, or

$$d \sim \frac{1}{m_{\rm DM} v_0} = 270 \text{ km} \left(\frac{220 \text{ km/s}}{v_0}\right)\left(\frac{10^{-9} \text{ eV}}{m_{\rm DM}}\right). \tag{4.2}$$

As we will demonstrate, this opens up striking new signatures, such as a unique daily modulation signal applicable only to wave-like DM with multiple detectors, because the direction of $\mathbf{x}_{12}$ rotates over a sidereal day with respect to the DM field.[5]

The main result of this paper is that interference effects between a pair of detectors separated by a distance $\mathbf{x}_{12}$ are fully characterized by the modified speed distributions

$$\begin{aligned} \mathcal{F}_{12}^c(v) &= \int d^3\mathbf{v} f(\mathbf{v}) \cos(m_{\rm DM}\mathbf{v}\cdot\mathbf{x}_{12})\delta[|\mathbf{v}| - v], \\ \mathcal{F}_{12}^s(v) &= \int d^3\mathbf{v} f(\mathbf{v}) \sin(m_{\rm DM}\mathbf{v}\cdot\mathbf{x}_{12})\delta[|\mathbf{v}| - v], \end{aligned} \tag{4.3}$$

with $f(\mathbf{v})$ the DM velocity distribution. Examples of these distributions at various times throughout the day are shown in Fig. 4.1 for optimally-separated detectors. If the goal is simply to enhance the total signal reach, we should maximize the constructive interference and take $d \ll \lambda_c$, in which case $\mathcal{F}_{12}^s(v) = 0$ and $\mathcal{F}_{12}^c(v) = f(v)$, with $f(v)$ the DM speed distribution. The observation that there is an enhanced sensitivity for an array of detectors located within the DM coherence length has been made previously in Ref. [228]; this is also the basis for the multiplexed cavity setup proposed by ADMX [107]. However, if the goal is to extract information about the full 3-dimensional DM phase space distribution, which encodes *e.g.* the boost of the Solar System with respect to the Galactic Center as well as possible DM substructure (including the Sagittarius stream [229] and the *Gaia* Sausage [230, 231]), we should take $d \sim \lambda_c$ as in (4.2). Ref. [228] points out the possibility of observing this daily modulation effect for experiments separated by distances of order $\lambda_c$; here we extend this analysis by focusing on constraining directional parameters in the phase space distribution. We will show that the sensitivity to the phase space information that may be

[4] This is conceptually distinct from the interferometry proposed in Refs. [222–224], where the interference results from a phase shift developed by electromagnetic fields as they propagate through axion DM.

[5] Several experimental proposals have noted or exploited sensitivity to the coherence length, see *e.g.* [70, 134, 165, 177, 225–227], but here we focus specifically on combining data between different experiments.

extracted from multiple detectors is comparable to the sensitivity to the initial discovery, since the interference effects have an $\mathcal{O}(1)$ effect on the data when $d \sim \lambda_c$. As such, in principle these unique signatures could be used to immediately verify a putative axion signal. More optimistically, DM interferometry would allow for the detailed mapping of the local DM phase space distribution after an initial detection.

For concreteness, we focus in this work on the case of axion DM coupled to electromagnetic signals, but our results would apply equally well to scalar and vector DM as long as the readout is proportional to the DM field. Similarly, for simplicity we will present most results for the case of two experiments, but our formalism holds for any number $\mathcal{N} \geq 2$ of experiments, and we will provide our key results for a general $\mathcal{N}$ also. Our results also apply equally well to resonant-type experiments and to broadband-type experiments (such as ABRACADABRA-10 cm [72, 82]), so long as the resonant experiments are able to preserve the phase of the data, as opposed to *e.g.* recording the power directly. One advantage of resonant experiments for wave-like DM, in addition to generically having enhanced sensitivity [220, 221], is that putative signal candidates may immediately yield detailed and high-significance studies, since the signal-to-noise ratio rapidly grows with measurement time when frequency-scanning is no longer necessary.

We organize the remaining discussion as follows. In Sec. 4.1, we sketch a derivation for the statistics of the correlated Fourier-transformed data from multiple experiments. A more extensive derivation and discussion is presented in App. C.2, with some useful orthogonality relations summarized in App. C.3. In Sec. 4.2, we construct a likelihood function for the axion signal as observed at $\mathcal{N}$ experiments, following the formalism of [80]; a practical data-stacking procedure is outlined in App. C.4. In Sec. 4.3 we use several simplified toy examples to illustrate analytic estimates of uncertainties on parameters of the velocity distribution using the Asimov data set – a technique where the asymptotic properties of the data are assumed in order to replace Monte Carlo simulations with analytic estimates (see Sec. 4.2) – and demonstrate that uncertainties on directional parameters in several simple examples with two detectors are minimized for $d \sim 2\lambda_c$. We also highlight the important distinction between the de Broglie wavelength and the coherence length for cold but boosted DM substructure. Furthermore, we show that there is a rotational symmetry of the likelihood which can lead to degenerate best-fit parameters for $\mathcal{N} = 2$ experiments. In Sec. 4.4, we extend the likelihood analysis to include daily modulation from the changing detector orientation throughout the day. We also perform analyses of simulated data sets to demonstrate how the likelihood may be implemented in practice to constrain the morphology of the DM phase-space distribution. Using the realistic examples of the Standard Halo Model (SHM) velocity distribution and the Sagittarius Stream, we show how daily modulation breaks the symmetry discussed in Sec. 4.3 and use this to perform parameter estimation using the effect. In App. C.1 we provide a brief review of the coherence length and time.

## 4.1 The Statistics of Multiple Detectors

In this section we describe the statistics of an axion DM signal collected by two or more spatially-separated detectors. In particular, while we expect background sources to be generally uncorrelated between detectors, the axion will induce non-trivial cross-correlations indicative of DM interferometry. These correlations will be the source of the additional information available to two or more experiments that we will extract using a likelihood formalism introduced in Sec. 4.2.

We imagine that a given detector, located at a position $\mathbf{x}$, is sensitive to the axion through a time-varying signal $\Phi$ proportional to the axion field,

$$\Phi(\mathbf{x}, t) = m_a \kappa_i \, a(\mathbf{x}, t). \tag{4.4}$$

The flux $\Phi$ is generated by the axion effective current, $\mathbf{J}_a \sim \partial_t a \sim m_a a$, which is the origin of $m_a$ in the expression. Accordingly, $\Phi \sim \kappa_i \mathbf{J}_a$, revealing $\kappa_i$ as characterizing the individual experimental response to the axion field. In the notation of Ref. [80], we take $\kappa_i = \sqrt{A_i/\rho_{\rm DM}}$. The dimensionful constant $A_i$ is characteristic of the individual experimental response to the axion field; for instance, in the case of ABRACADABRA [71, 72, 82], the magnetic flux induced in the pickup loop at the center of the detector is related to the axion field by $A = \rho_{\rm DM} g_{a\gamma\gamma}^2 B_0^2 V_B^2$, where $g_{a\gamma\gamma}$ is the axion-photon coupling, $B_0$ is the toroidal magnetic field strength, and $V_B$ is an effective magnetic field volume associated with the detector. In addition to the experimental factors, $A_i$ has been defined to include $\rho_{\rm DM} g_{a\gamma\gamma}^2$, which determines the mean power in the axion field. We assume for simplicity that the detector response $A_i$ is purely real and does not include phase delays. Similar expressions are available for other detectors [80]. For our discussion, all that is required is a measurement linear in the axion field, in order to ensure direct access to the axion phase. Measurements intrinsically proportional to $a^2$, such as the power in the cavity of an axion haloscope, cannot be directly ported to our formalism. Nevertheless, interferometry can still be performed by these resonant cavity experiments, as long as the phase information is extracted. This may be achieved for example by reading out electromagnetic signals with phase-sensitive amplifiers (*e.g.*, [68, 232]).

Ultimately, we envision a set of measurements $\Phi_i$ of the same axion field, made by $\mathcal{N}$ detectors at different spatial locations $\mathbf{x}_i$. The correlations between these data sets will arise due to the statistics of the underlying axion field, as we will describe in the following subsections, leaving the full derivation to App. C.2.

### 4.1.1 Construction of the Axion Field

It is useful to recall the underlying statistics in the axion field that result from its finite velocity dispersion and wave-like nature. In [80] it was shown that we may represent the axion field as seen by a single detector as

$$a(t) = \frac{\sqrt{\rho_{\rm DM}}}{m_a} \sum_j \alpha_j \sqrt{f(v_j)\Delta v} \cos\left[\omega_j t + \phi_j\right]. \tag{4.5}$$

Here, the sum over $j$ indicates a sum over subsets of particles with speeds in the interval $v$ to $v + \Delta v$. The phase is controlled by $\omega_j = m_a\left(1 + v_j^2/2\right)$ and a random contribution $\phi_j \in [0, 2\pi)$, and further $f(v)$ is the DM speed distribution in the laboratory frame. In Ref. [80] the continuum limit for speeds $\Delta v \to 0$ is taken; below (and in App. C.2) we will generalize to the continuum limit for velocities. In addition to the random phase, the random nature of the axion field is captured in the random variate $\alpha_j$ drawn from the Rayleigh distribution $p[\alpha] = \alpha\, e^{-\alpha^2/2}$.

While (4.5) represents the axion field constructed from the discretized frequency modes specified by the local DM velocity distribution, a more fundamental approach can be understood by considering the local DM field made up of $N_a$ axion particles (or wave packets), as detailed in [80]. The enormous occupation numbers characteristic of wave-like DM will then allow us to eventually convert this sum to an integral by taking the $N_a \to \infty$ limit; in detail, we should have $n_{\rm DM}\lambda_{\rm dB}^3 \gg 1$, where $n_{\rm DM}$ is the DM number density, which is satisfied locally for $m_a \ll 1$ eV. We note that the above construction also assumes DM is a non-interacting wave, which means that self-interactions should be negligible.

The axion field described in (4.5) is appropriate for a single detector, but to reveal the effects of DM interferometry we need to extend the description to include the spatial dependence of the DM wave. In particular, the phase will also include a contribution $\mathbf{k}\cdot\mathbf{x}$, with $\mathbf{k} = m_a\mathbf{v}$ for a non-relativistic wave. As $\mathbf{k}$ depends on the velocity, and not speed, we need to extend the above sum to three independent components, $\mathbf{v}_{abc} = v_a\hat{\mathbf{x}} + v_b\hat{\mathbf{y}} + v_c\hat{\mathbf{z}}$, where the indexes $a, b, c$ are integers. We may then write

$$a(\mathbf{x}, t) = \frac{\sqrt{\rho_{\rm DM}}}{m_a} \sum_{abc} \alpha_{abc} \sqrt{f(\mathbf{v}_{abc})(\Delta v)^3} \times \cos\left[\omega_{abc} t - \mathbf{k}_{abc}\cdot\mathbf{x} + \phi_{abc}\right], \tag{4.6}$$

where $\omega_{abc}$ depends on $v_{abc} = |\mathbf{v}_{abc}|$, and $\alpha_{abc}$ and $\phi_{abc}$ are Rayleigh and uniform random variables, respectively, as in (4.5). Here, $(\Delta v)^3$ is a discretization of the 3-dimensional velocity, generalizing $\Delta v$ for speeds; we will take the continuum limit in App. C.2.

In (4.6) we have written the axion field in a convenient form for revealing DM interferometry. To reiterate the point, if we measure the axion field at a single location, we can always choose our

coordinates such that $\mathbf{x} = 0$. In this case, the velocity information in $\mathbf{k}$ is lost, and we are now only sensitive to the speed $v = |\mathbf{v}|$ through $\omega$. This collapses $f(\mathbf{v}) \to f(v)$, and (4.6) to (4.5); information about the phase space is lost. However, if we measure the axion field at two locations, an irreducible $\mathbf{k}$ dependence remains, and the full velocity information is imprinted in the multi-detector covariance matrix.

We implicitly assume throughout this work that the non-interacting plane-wave superposition in (4.6) applies for all $\mathbf{x}$. Corrections to this picture should arise from *e.g.* the gravitational field of the Earth, which would slightly bend the DM trajectories between detectors. However, the DM velocities we consider in this work are much larger than the Earth's escape velocity, and also the detector separations are typically much smaller than the radius of the Earth, so we are justified in neglecting this effect.

### 4.1.2 The Multi-Detector Covariance Matrix

We will now outline how the statistics of the axion field, as described above, lead to a non-trivial covariance matrix in the data collected by $\mathcal{N}$ experiments. In this section we will simply state the key results, leaving a derivation to App. C.2.

We begin by considering the minimal case of $\mathcal{N} = 1$. Suppose that a single experiment takes a time-series of $N$ measurements $\{\Phi_n(\mathbf{x}) = \Phi(\mathbf{x}, n\Delta t)\}$, with $n = 0, 1, \ldots, N-1$, collected over a time $T$, so that $\Delta t = T/N$. In order to isolate a signal oscillating at a particular frequency, as we expect the axion to do, we calculate the discrete Fourier transform

$$\Phi_k(\mathbf{x}) = \sum_{n=0}^{N-1} \Phi_n(\mathbf{x})\, e^{-i2\pi kn/N}\,. \tag{4.7}$$

The transform is indexed by an integer $k = 0, 1, \ldots, N-1$, which is related to the angular frequency, $\omega = 2\pi k/T$. We will switch back and forth between talking about frequency $\omega$ and wave-number $k$ as convenient. It is convenient to partition the Fourier transform into appropriately normalized real and imaginary parts as follows,

$$\begin{aligned} R_k(\mathbf{x}) &= \frac{\Delta t}{\sqrt{T}} \mathrm{Re}\left[\Phi_k(\mathbf{x})\right], \\ I_k(\mathbf{x}) &= \frac{\Delta t}{\sqrt{T}} \mathrm{Im}\left[\Phi_k(\mathbf{x})\right]. \end{aligned} \tag{4.8}$$

We can then write the power spectral density (PSD) as,

$$S^k_{\Phi\Phi} = \frac{(\Delta t)^2}{T} |\Phi_k(\mathbf{x})|^2 = R_k^2(\mathbf{x}) + I_k^2(\mathbf{x}). \tag{4.9}$$

We will assume throughout that $T$ is long enough such that the signal is sufficiently well resolved, *i.e.* the bandwidth of the Fourier transform $2\pi/T$ is much smaller than the width of the signal in frequency space. When specifying Fourier components by frequency as opposed to wave-number we use notation as in $S^k_{\Phi\Phi}(\mathbf{x}) \to S_{\Phi\Phi}(\mathbf{x},\omega)$.

As shown in [80], both $R(\mathbf{x},\omega)$ and $I(\mathbf{x},\omega)$ are normally distributed with zero mean and variance given by

$$\langle R^2(\mathbf{x},\omega)\rangle = \langle I^2(\mathbf{x},\omega)\rangle = \frac{A}{2}\frac{\pi f(v_\omega)}{m_a v_\omega}\,, \tag{4.10}$$

where we have defined $v_\omega = \sqrt{2\omega/m_a - 2}$ as the axion velocity corresponding to a frequency $\omega$, and the speed distribution is defined as

$$f(v) = \int d^3\mathbf{v} f(\mathbf{v})\delta(v - |\mathbf{v}|) \tag{4.11}$$

This implies, for example, that $S_{\Phi\Phi}(\mathbf{x})$ is an exponentially distributed quantity, with mean

$$\langle S_{\Phi\Phi}(\omega)\rangle = A\frac{\pi f(v_\omega)}{m_a v_\omega}\,. \tag{4.12}$$

We can understand the velocity dependence by looking back to (4.5); the signal measured by a detector $\Phi$ is proportional to the time-dependent axion axion field, and since $a$ is proportional to $\sqrt{f(v)}$, we obtain a power spectrum $S_{\Phi\Phi}$ proportional to $f(v)$.

In any real experiment there will also be background. However, as long as the background is normally distributed in the time domain – as expected for, for example, thermal noise, SQUID flux noise, or Josephson parametric amplifier noise – then both $R(\mathbf{x},\omega)$ and $I(\mathbf{x},\omega)$ remain normally distributed but with variance

$$\langle R^2(\mathbf{x},\omega)\rangle = \langle I^2(\mathbf{x},\omega)\rangle = \frac{A}{2}\frac{\pi f(v_\omega)}{m_a v_\omega} + \frac{\lambda_B(\omega)}{2}\,, \tag{4.13}$$

where $\lambda_B(\omega)$ encapsulates the variance of the potentially frequency-dependent noise from the background sources only.

Note that $R_k(\mathbf{x})$ and $I_k(\mathbf{x})$ are uncorrelated; in particular, the $2\times 2$ covariance matrix for these two quantities is simply

$$\mathbf{\Sigma}_k = \left(\frac{A}{2}\frac{\pi f(v_\omega)}{m_a v_\omega} + \frac{\lambda_B(\omega)}{2}\right)\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}. \tag{4.14}$$

This implies that for a single detector, all information about the signal is contained in the PSD

$S_{\Phi\Phi}(\mathbf{x})$. Further, as shown in (4.14), the location $\mathbf{x}$ never enters for $\mathcal{N} = 1$. Even if we chose our coordinates such that $\mathbf{k} \cdot \mathbf{x} \neq 0$, the overall phase remains unphysical as it would vanish when computing the modulus squared in (4.9).

Now let us extend the discussion to the case of interest: data collected by $\mathcal{N}$ experiments at positions $\mathbf{x}_i$, with $i = 1, 2, \ldots, \mathcal{N}$. For each data set, we calculate the real and imaginary parts of the Fourier transform as above. The information collected by all detectors can then be organized into the following $2\mathcal{N}$ dimensional data vector,

$$\mathbf{d}_k = \left[R_k(\mathbf{x}_1), I_k(\mathbf{x}_1), \ldots, R_k(\mathbf{x}_\mathcal{N}), R_k(\mathbf{x}_\mathcal{N})\right]^T. \tag{4.15}$$

Correlations between the real and imaginary part for any given detector will be identical to the $\mathcal{N} = 1$ case discussed above. However, DM interferometry will reveal itself through non-trivial correlations amongst the different detectors.[6] Indeed, as justified in App. C.2, $\mathbf{d}_k$ will be a $2\mathcal{N}$-dimensional Gaussian random variable with zero mean and a symmetric $(2\mathcal{N} \times 2\mathcal{N})$-dimensional covariance matrix given by

$$\begin{aligned}
\boldsymbol{\Sigma}_k = &\begin{bmatrix}
\langle R_k(\mathbf{x}_1)R_k(\mathbf{x}_1)\rangle & \langle R_k(\mathbf{x}_1)I_k(\mathbf{x}_1)\rangle & \langle R_k(\mathbf{x}_1)R_k(\mathbf{x}_2)\rangle & \ldots & \langle R_k(\mathbf{x}_1)I_k(\mathbf{x}_\mathcal{N})\rangle \\
\langle I_k(\mathbf{x}_1)R_k(\mathbf{x}_1)\rangle & \langle I_k(\mathbf{x}_1)I_k(\mathbf{x}_1)\rangle & \langle I_k(\mathbf{x}_1)R_k(\mathbf{x}_2)\rangle & \ldots & \langle I_k(\mathbf{x}_1)I_k(\mathbf{x}_\mathcal{N})\rangle \\
\vdots & & \ddots & & \\
\langle I_k(\mathbf{x}_\mathcal{N})R_k(\mathbf{x}_1)\rangle & \langle I_k(\mathbf{x}_\mathcal{N})I_k(\mathbf{x}_1)\rangle & \langle I_k(\mathbf{x}_\mathcal{N})R_k(\mathbf{x}_2)\rangle & \ldots & \langle I_k(\mathbf{x}_\mathcal{N})I_k(\mathbf{x}_\mathcal{N})\rangle
\end{bmatrix}, \\
&\langle R_k(\mathbf{x}_i)R_k(\mathbf{x}_j)\rangle = \langle I_k(\mathbf{x}_i)I_k(\mathbf{x}_j)\rangle = \frac{1}{2}\left[c_{ij}(\omega) + \delta_{ij}\lambda_{B,i}(\omega)\right], \\
&\langle R_k(\mathbf{x}_i)I_k(\mathbf{x}_j)\rangle = -\langle I_k(\mathbf{x}_i)R_k(\mathbf{x}_j)\rangle = \frac{1}{2}s_{ij}(\omega).
\end{aligned} \tag{4.16}$$

Here $\lambda_{B,i}(\omega)$ is the background observed by the $i^{\text{th}}$ experiment, and its contribution is purely diagonal. The axion signal, however, induces off-diagonal correlations, which we quantify in terms of

$$\begin{aligned}
c_{ij}(\omega) &= \frac{\pi\sqrt{A_i A_j}}{m_a v_\omega}\mathcal{F}^c_{ij}(v_\omega), \\
s_{ij}(\omega) &= \frac{\pi\sqrt{A_i A_j}}{m_a v_\omega}\mathcal{F}^s_{ij}(v_\omega),
\end{aligned} \tag{4.17}$$

[6] Our analysis assumes that the experiments have identical timestamps on the data, or equivalently that the relative phase of the signal at each experiment is precisely known. Of course, this is not exactly true and in general there will be an additional contribution to the phase of $\Phi$ in (4.7) of the form $\omega\Delta\tau$, where $\Delta\tau$ is the timing error. As long as $\Delta\tau \ll |\mathbf{x}_{ij}|v$, this contribution can be safely neglected. For two detectors with $|\mathbf{x}_{12}| \sim 50$ m and $v \sim 200$ km/s, this implies $\Delta\tau \ll 10^{-10}$ s. Typical atomic clocks have timing error of $10^{-9}$ s/day, so the required $\Delta\tau$ can be achieved by synchronizing the two experiments to an atomic clock over data-taking intervals of about 2.5 hours, which is sufficient for the daily modulation analysis in Sec. 4.4.

with

$$\mathcal{F}^c_{ij}(v) = \int d^3\mathbf{v} f(\mathbf{v}) \cos(m_a \mathbf{v}\cdot\mathbf{x}_{ij})\delta[|\mathbf{v}| - v],$$
$$\mathcal{F}^s_{ij}(v) = \int d^3\mathbf{v} f(\mathbf{v}) \sin(m_a \mathbf{v}\cdot\mathbf{x}_{ij})\delta[|\mathbf{v}| - v]. \tag{4.18}$$

By translation invariance, the entries of the correlation matrix only depend on the relative distances $\mathbf{x}_{ij} \equiv \mathbf{x}_i - \mathbf{x}_j$. These expressions then simplify for the correlations amongst a single detector, as $\mathcal{F}^c_{ii}(v) = f(v)$ and $\mathcal{F}^s_{ii}(v) = 0$. But for $i \neq j$, the expressions in (4.18) contain a modulated version of the full velocity distribution, allowing us to extract non-trivial directional information about the velocity distribution $f(\mathbf{v})$ with multiple detectors separated by distances of order the de Broglie wavelength, where the integrand in (4.18) exhibits maximal variation. We note that the formalism we have developed assumes that the velocity distribution is stationary, or at least varies slowly on timescales compared to the axion coherence time. In Sec. 4.4 we will develop a formalism to take into account the daily modulation of $f(\mathbf{v})$ through a joint likelihood over multiple data-taking intervals.

## 4.2 A Likelihood for Multi-Detector Axion Direct Detection

Having understood the statistics underlying the data collected by multiple detectors, we now outline how to incorporate these lessons into an appropriate likelihood. The likelihood will be a simple generalization of the axion likelihood (generally applicable to wave-like DM) introduced in [80], and we will closely follow their approach. However, unlike in [80], we work explicitly with the data as represented in $R_k$ and $I_k$ rather than the PSD, as the former notation exposes the full set of multi-detector correlations, as captured by $\mathbf{\Sigma}$ in (4.16). We will then outline how we can extract information about the parameters of $f(\mathbf{v})$ using this likelihood, exploiting where possible the asymptotic Asimov procedure [172] to determine results analytically. In Sec. 4.3 we will then put the formalism to use in the context of several toy examples designed to highlight where interferometry opens up new avenues, and build intuition for the more realistic scenarios considered in Sec. 4.4.

### 4.2.1 The Multi-Detector Likelihood

As detailed in Sec. 4.1, we imagine we have a data set collected by $\mathcal{N}$ experiments, which each perform a time series of $N$ measurements collected at a frequency $f = 1/\Delta t$ of a quantity $\Phi \propto a$. The real and imaginary part of the discrete Fourier transform of each experiments data set is constructed according to (4.8), and then arranged into a single data set $d = \{\mathbf{d}_0, \mathbf{d}_1, \ldots, \mathbf{d}_{N-1}\}$, with

$\mathbf{d}_k$ as given in (4.15). We then define a model $\mathcal{M}$ with parameter vector $\boldsymbol{\theta}$ that has nuisance parameters $\boldsymbol{\theta}_{\rm nuis}$ describing the backgrounds in the individual experiments (encapsulated by $\lambda_{B,i}(\omega)$) and signal parameters $\boldsymbol{\theta}_{\rm sig}$ that characterize the axion contribution. For example, $\boldsymbol{\theta}_{\rm sig}$ includes $g_{a\gamma\gamma}$, $m_a$, and model parameters that describe the DM velocity distribution $f(\mathbf{v})$. Then, as the data set is distributed according to a multivariate Gaussian, the appropriate likelihood is given by,

$$\mathcal{L}(d|\mathcal{M},\boldsymbol{\theta}) = \prod_{k=0}^{N-1} \frac{\exp\left[-\frac{1}{2}\mathbf{d}_k^T\cdot\boldsymbol{\Sigma}_k^{-1}(\boldsymbol{\theta})\cdot\mathbf{d}_k\right]}{\sqrt{(2\pi)^{2\mathcal{N}}|\boldsymbol{\Sigma}_k(\boldsymbol{\theta})|}}\,, \tag{4.19}$$

where $|\boldsymbol{\Sigma}_k(\boldsymbol{\theta})|$ is the determinant of the covariance matrix.

The utility of the likelihood function is that it facilitates inferences regarding the signal parameters, $\boldsymbol{\theta}_{\rm sig}$, from the data. The ultimate goal of the axion DM program would be to infer a nonzero value of $A$, and hence the existence of a coupling between the Standard Model and DM, for example $g_{a\gamma\gamma}$. Taking a frequentist approach to that problem, it is useful to define the following test statistic (TS) from the profile likelihood:

$$\Theta(\boldsymbol{\theta}_{\rm sig}) = 2[\ln\mathcal{L}(d|\mathcal{M},\{\hat{\boldsymbol{\theta}}_{\rm nuis},\boldsymbol{\theta}_{\rm sig}\}) - \ln\mathcal{L}(d|\mathcal{M},\{\hat{\boldsymbol{\theta}}_{\rm nuis},\boldsymbol{\theta}_{\rm sig}=\mathbf{0}\})]. \tag{4.20}$$

In each likelihood, $\hat{\boldsymbol{\theta}}_{\rm nuis}$ denotes the value of the nuisance parameters that maximizes the likelihood for the given signal parameters. The TS is then a function of the signal model parameters. In particular, this means that in the first term in (4.20) the nuisance parameters are uniquely determined at each $\boldsymbol{\theta}_{\rm sig}$ point by the values which maximize the log likelihood. The second term in (4.20) is evaluated on the null model $\boldsymbol{\theta}_{\rm sig} = \mathbf{0}$, which can be achieved by setting the signal strength parameter $A$ to zero.

The TS in (4.20) is convenient for quantifying the significance of a putative signal, and we will use it throughout the following analysis. In App. C.4 we describe a data-stacking procedure which reduces the data storage requirements for practical applications of our formalism.

### 4.2.2 Asimov Test Statistic

In order to build intuition for the information accessible to multiple detectors, we will use the Asimov data set [172] to study the asymptotic TS analytically. More precisely, the Asimov analogue of the TS in (4.20) is the average value taken over data realization,

$$\widetilde{\Theta}(\boldsymbol{\theta}_{\rm sig}) = \langle\Theta(\boldsymbol{\theta}_{\rm sig})\rangle\,, \tag{4.21}$$

where the expectation value is taken on the data. In order to evaluate the asymptotic TS, it is convenient to separate the model prediction, which enters through $\boldsymbol{\Sigma}$, into background and signal contributions:

$$\boldsymbol{\Sigma} = \mathbf{B} + \mathbf{S}\,. \tag{4.22}$$

Referring back to (4.16), recall that the background is purely diagonal:

$$\mathbf{B} = \frac{1}{2}\,\mathrm{diag}(\lambda_{B,1},\, \lambda_{B,1},\, \ldots,\, \lambda_{B,\mathcal{N}},\, \lambda_{B,\mathcal{N}}). \tag{4.23}$$

Using this partitioning of the model prediction, we can then express the TS as follows

$$\Theta = \sum_{k=1}^{N-1} \left( \mathbf{d}_k^T \left[ \mathbf{B}_k^{-1} - \boldsymbol{\Sigma}_k^{-1} \right] \mathbf{d}_k - \ln \left[ \frac{|\boldsymbol{\Sigma}_k|}{|\mathbf{B}_k|} \right] \right). \tag{4.24}$$

Note that the values of $\mathbf{B}$ appearing in this expression are understood as being set to the value required by the profile likelihood technique. In order to evaluate the Asimov form of this expression, we only need to evaluate the average on the first term, as the average is taken over the data. This can be evaluated as follows,

$$\left\langle \mathbf{d}_k^T \left[ \mathbf{B}_k^{-1} - \boldsymbol{\Sigma}_k^{-1} \right] \mathbf{d}_k \right\rangle = \mathrm{Tr}\left( \left\langle \mathbf{d}_k \mathbf{d}_k^T \right\rangle \left[ \mathbf{B}_k^{-1} - \boldsymbol{\Sigma}_k^{-1} \right] \right), \tag{4.25}$$

and then as the data has mean zero, we know the above expected value is simply given by the true covariance matrix,

$$\left\langle \mathbf{d}_k \mathbf{d}_k^T \right\rangle = \boldsymbol{\Sigma}_k(\boldsymbol{\theta} = \boldsymbol{\theta}_{\mathrm{truth}}) = \boldsymbol{\Sigma}_k^t\,. \tag{4.26}$$

Here the truth parameters can be considered as, for example, the parameters one would use when generating Monte Carlo to simulate expected experimental results. For instance, to estimate the expected limit, the truth parameters would commonly have $A = 0$, whereas if we are estimating our sensitivity to features in $f(\mathbf{v})$, we will take $A \neq 0$ in the Asimov data. In this work we are interested in the latter case, and therefore we will further assume the background has been fixed to the true value as a result of the profile likelihood technique,

$$\boldsymbol{\Sigma}_k^t = \mathbf{S}_k^t + \mathbf{B}_k\,, \tag{4.27}$$

where $\mathbf{S}^t$ is the true signal model and the same $\mathbf{B}$ appears in both the Asimov and model predictions. We further assume that the signal is always parametrically smaller than the background, which is the regime we will be in for any realistic experimental setup.[7] Implementing these assumptions, the Asimov form of (4.24) is[8]

$$\widetilde{\Theta}(\boldsymbol{\theta}_{\rm sig}) \approx \sum_{k=1}^{N-1} \mathrm{Tr}\left[\left(\mathbf{S}_k^t - \frac{1}{2}\mathbf{S}_k\right)\mathbf{B}_k^{-1}\mathbf{S}_k\mathbf{B}_k^{-1}\right]. \tag{4.28}$$

In (4.28) we have a convenient form of the expected TS that is amenable to analytic study. In the present work, our particular interest is the information contained in $f(\mathbf{v})$ that we can only access as a result of DM interferometry. As such, it is convenient to evaluate a form of the Asimov TS, where all parameters except for those that control $f(\mathbf{v})$, as encoded in $\mathcal{F}_{ij}^c(v)$ and $\mathcal{F}_{ij}^s(v)$ in (4.18), are set to their true values in the presence of a non-zero signal. If we further assume that our frequency resolution is sufficiently fine with respect to the scales over which the signal and background vary, then we can approximate the sum over Fourier components $k$ with an integral over frequencies $\omega$, or equivalently speeds $v = \sqrt{2\omega/m_a - 2}$. Under these assumptions, the TS becomes

$$\widetilde{\Theta} = \frac{T\pi}{m_a}\int\frac{dv}{v}\sum_{i,j=1}^{\mathcal{N}}\frac{A_iA_j}{\lambda_{B,i}\lambda_{B,j}}\left[\mathcal{F}_{ij}^c(v)\left(\mathcal{F}_{ij}^{c,t}(v) - \frac{1}{2}\mathcal{F}_{ij}^c(v)\right) + \mathcal{F}_{ij}^s(v)\left(\mathcal{F}_{ij}^{s,t}(v) - \frac{1}{2}\mathcal{F}_{ij}^s(v)\right)\right]. \tag{4.29}$$

Much of the remainder of this work is devoted to studying the implications of this result.

### 4.2.3 Limiting Cases of Zero and Infinite Separation

We can use (4.29) to confirm basic asymptotic scalings expected for an analysis performed with DM interferometry. To begin with, the Asimov TS for a single detector with response $A$ and background $\lambda_B$, recalling $\mathcal{F}_{ii}^s(v) = 0$ and $\mathcal{F}_{ii}^c(v) = f(v)$, is given by

$$\widetilde{\Theta}_{\mathcal{N}=1} = \frac{A^2T\pi}{m_a}\int\frac{dv}{v}\frac{f(v)}{\lambda_B^2}\left(f^t(v) - \frac{1}{2}f(v)\right). \tag{4.30}$$

This expression agrees with the result in [80], which was derived for a single detector when analyzing the PSD. Importantly, we emphasize once more that (4.30) is only dependent on the speed

[7] This assumption also ensures the validity of the fixed background being the same in *e.g.* (4.27) and (4.24); if the signal is comparable to the background, then varying $A$ will generically alter the background determined by the profile likelihood technique.

[8] To derive this result, the following identity is useful: $\ln|\mathbf{M}| = \mathrm{Tr}\ln\mathbf{M}$, for a matrix $\mathbf{M}$.

distribution, so directional parameters which affect the velocity distribution but not the DM speed distribution are inaccessible. Before moving on to multiple detectors, we note that the expected discovery significance, which we denote TS, is given by the Asimov $\Theta$ evaluated at the model parameters that maximize the likelihood, which are the truth parameters. Setting $f(v) = f^t(v)$ in (4.30) above gives

$$\mathrm{TS}_{\mathcal{N}=1} \approx \frac{A^2 T\pi}{2m_a} \int \frac{dv}{v} \frac{f(v)^2}{\lambda_B^2}. \tag{4.31}$$

In order to extract directional parameters we need at least two detectors. To that end, consider our expression in (4.29) for $\mathcal{N} = 2$. For simplicity, we take $A_1 = A_2 = A$ and $\lambda_{B,1} = \lambda_{B,2} = \lambda_B$, in which case the Asimov TS becomes

$$\begin{aligned} \widetilde{\Theta}_{\mathcal{N}=2} = \frac{2A^2 T\pi}{m_a} \int \frac{dv}{v\,\lambda_B^2} \Bigg[ & f(v) \left( f^t(v) - \frac{1}{2} f(v) \right) \\ & + \mathcal{F}_{12}^c(v) \left( \mathcal{F}_{12}^{t,c}(v) - \frac{1}{2} \mathcal{F}_{12}^c(v) \right) \\ & + \mathcal{F}_{12}^s(v) \left( \mathcal{F}_{12}^{t,s}(v) - \frac{1}{2} \mathcal{F}_{12}^s(v) \right) \Bigg]. \end{aligned} \tag{4.32}$$

In particular, the discovery TS is given by

$$\mathrm{TS}_{\mathcal{N}=2} = \frac{A^2 T\pi}{m_a} \int \frac{dv}{v} \frac{f(v)^2 + \mathcal{F}_{12}^c(v)^2 + \mathcal{F}_{12}^s(v)^2}{\lambda_B^2}. \tag{4.33}$$

Through $\mathcal{F}_{12}^c$ and $\mathcal{F}_{12}^s$, the discovery TS depends on the spatial separation of the two experiments $d = |\mathbf{x}_{12}|$. In the limit where the experiments are close with respect to the DM coherence length, *i.e.* $d \ll \lambda_c$, then the two experiments see the same phase of the DM wave ($\mathbf{k} \cdot \mathbf{x}$ does not vary appreciably between them). In this case, we would expect a coherent enhancement in the signal. Defining for future use

$$\mathrm{TS}_0 = \lim_{d \to 0} \mathrm{TS} = \frac{2A^2 T\pi}{m_a} \int \frac{dv}{v} \frac{f(v)^2}{\lambda_B^2}, \tag{4.34}$$

we see from (4.18) that for $\mathbf{x}_{12} = 0$ we have $\mathcal{F}_{12}^c(v) = f(v)$ and $\mathcal{F}_{12}^s(v) = 0$, so $\mathrm{TS}_0 = 4\mathrm{TS}_{\mathcal{N}=1}$. The $\mathcal{N}^2 = 4$ enhancement of the TS represents a coherent enhancement, a point emphasized in [228]. This configuration provides a benchmark for the largest TS we can achieve for a general $\mathcal{N} = 2$ configuration, and therefore will provide a convenient benchmark in the studies that follows. On the other hand, for widely separated detectors with $d \gg \lambda_c$, the DM fields will add incoherently. For the problem at hand, again returning to (4.18), we see that the sine and cosine

factors will oscillate rapidly, driving the integrals to zero. What remains is,

$$\lim_{d\to\infty} \mathrm{TS} = \frac{A^2 T\pi}{m_a} \int \frac{dv}{v} \frac{f(v)^2}{\lambda_B^2} = 2\mathrm{TS}_{\mathcal{N}=1}\,, \tag{4.35}$$

so that the TS now only scales as $\mathcal{N}$, an incoherent enhancement.

The above argument can be readily generalized to $\mathcal{N}$ detectors. Typically the signal strength $A$ is proportional to $g_{a\gamma\gamma}^2$, so for $\mathcal{N}$ experiments all with pairwise separations $d \ll \lambda_c$, we expect our sensitivity to $g_{a\gamma\gamma}$ should scale coherently as $\mathcal{N}^{1/2}$. If instead all experiments have $d \gg \lambda_c$, the scaling is reduced to $\mathcal{N}^{1/4}$, and for scenarios outside these two extremes the scaling will be somewhere in between. However, it is precisely this intermediate regime, where neither $\mathcal{F}$ reduces to the speed distribution nor vanishes, where we expect to be able to extract additional information about $f(\mathbf{v})$. We turn to the problem of estimating parameters of $f(\mathbf{v})$ in the context of the Asimov data set in the next section.

## 4.3 Asimov Parameter Estimation

In this section we will use (4.32) to perform frequentist parameter estimation and show explicitly that additional information about $f(\mathbf{v})$ can be extracted via DM interferometry. For the purpose of simplifying the discussion, we will restrict our attention to the case of two detectors with equal background and detector responses as given in (4.32). However, the entire discussion can be readily generalized to $\mathcal{N}$ arbitrary detectors by using the asymptotic TS expression in (4.29).

To be specific, imagine we are interested estimating a set of signal parameters $\boldsymbol{\alpha}$, which are a subset of the full set of signal parameters $\boldsymbol{\alpha} \subset \boldsymbol{\theta}_{\rm sig}$ related to $f(\mathbf{v})$, and which have true values $\boldsymbol{\alpha}^t$. The Asimov procedure allows us to study our ability to infer these parameters. For example, it is straightforward to confirm that $\widetilde{\Theta}(\boldsymbol{\alpha})$ is maximized for $\boldsymbol{\alpha} = \boldsymbol{\alpha}^t$.[9] Beyond the best fit values, we are interested in determining the associated expected uncertainties and correlations between the various parameters, which are encompassed in the covariance matrix between the parameters, which we denote $\mathbf{C}$. An estimator for this covariance matrix is given by the inverse Fisher information evaluated at the maximum likelihood, $\mathbf{C}^{-1} = \mathbf{I}(\hat{\boldsymbol{\alpha}})$ (again $\hat{\boldsymbol{\alpha}}$ are the parameters that maximize the likelihood), where

$$I_{ij}(\boldsymbol{\alpha}) = -\frac{\partial^2 \ln \mathcal{L}(\boldsymbol{\alpha})}{\partial\alpha_i \partial\alpha_j} = -\frac{1}{2}\frac{\partial^2 \Theta(\boldsymbol{\alpha})}{\partial\alpha_i \partial\alpha_j}\,, \tag{4.36}$$

where we use (4.20). Given this relation, asymptotically our estimate for the covariance matrix is

[9] We emphasize that there is no guarantee that other parameters besides $\boldsymbol{\alpha}^t$ cannot also maximize the likelihood. Indeed we will see exactly this possibility realized in a number of examples considered below.

given by

$$
\begin{aligned}
[\widetilde{\mathbf{C}}^{-1}]_{ij} =& -\frac{1}{2}\left.\frac{\partial^2\widetilde{\Theta}(\boldsymbol{\alpha})}{\partial\alpha_i\partial\alpha_j}\right|_{\boldsymbol{\alpha}=\boldsymbol{\alpha}^t} \\
=&\frac{A^2T\pi}{m_a}\int\frac{dv}{v\,\lambda_B^2}\left[(\partial_i f(v))(\partial_j f(v))\right. \\
&\left.+\,(\partial_i\mathcal{F}_{12}^c(v))(\partial_j\mathcal{F}_{12}^c(v))+(\partial_i\mathcal{F}_{12}^s(v))(\partial_j\mathcal{F}_{12}^s(v))\right].
\end{aligned}
\tag{4.37}
$$

This expression involves the following shorthand for derivatives of functions then evaluated at their truth values, $\partial_i = \partial/\partial\alpha_i|_{\alpha_i=\alpha_i^t}$. The expression in the first line of this result lays bare a simple fact: if $\widetilde{\Theta}$ has no dependence on a particular parameter, for example the incident direction of a DM stream, or orientation of the Sun's motion through the DM halo, then the associated entries of the inverse covariance matrix vanish along with our ability to estimate that parameter. For the case of a single parameter $\alpha$, we can readily invert the covariance matrix, and the above expression simplifies to

$$
\sigma_\alpha^{-2} =\frac{A^2T\pi}{m_a}\int\frac{dv}{v\,\lambda_B^2}\left[(\partial_\alpha f(v))^2(\partial_\alpha\mathcal{F}_{12}^c(v))^2+(\partial_\alpha\mathcal{F}_{12}^s(v))^2\right], \tag{4.38}
$$

where again all parameters are evaluated at their truth values after derivatives. We can already calibrate our basic expectation for parameter estimation from this result. Optimal estimation of $\alpha$ amounts to maximizing the right hand side of the expression; indeed, as expected, increasing the signal strength, $A$, or the integration time, $T$, both achieve this. If a parameter can be estimated from the speed distribution $f(v)$ (in other words, $\partial_\alpha f(v)\neq 0$), then that parameter may be estimated by a detector configuration with $d\ll\lambda_c$. However, the true power in the multi detector setup arises for parameters invisible to a single detector, defined by $\partial_\alpha f(v)=0$, but where $\partial_\alpha\mathcal{F}_{12}^{c,s}(v)\neq 0$. In generic cases, such parameters are optimally estimated for $d\sim\lambda_c$.

Continuing, let us assume that $\lambda_B$ is independent of frequency, in which case (4.38) becomes

$$
\sigma_\alpha^2 = \frac{2}{\mathrm{TS}_0}\left[\int\frac{dv}{v}(f^t(v))^2\right]\left\{\int\frac{dv}{v}\left[(\partial_\alpha f(v))^2+\ (\partial_\alpha\mathcal{F}_{12}^c(v))^2+(\partial_\alpha\mathcal{F}_{12}^s(v))^2\right]\right\}^{-1}, \tag{4.39}
$$

expressed in terms of $\mathrm{TS}_0$ as introduced in (4.34). In particular, this result demonstrates the expected scaling of $\sigma_\alpha\sim(\mathrm{TS}_0)^{-1/2}$; the exact details will require a specific $f(\mathbf{v})$ and experimental configuration. In the following subsections we will continue this line of thinking, demonstrating in several toy examples that a second detector can lift degeneracies from the single detector likelihood.

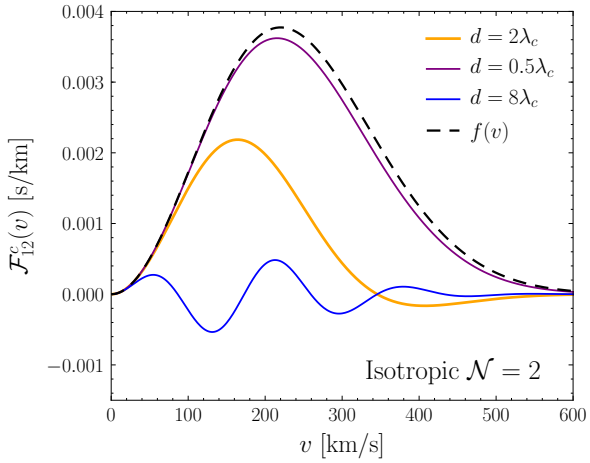



Figure 4.2: The modified speed distribution, $\mathcal{F}_{12}^c(v)$, that carries the imprint of DM interferometry. Here we show the particularly simple example of an isotropic SHM for $\mathcal{N} = 2$ detectors, in which case the expression is given in (4.40). The result is shown for various choices of the two detector separation $d$ as compared to the axion coherence length $\lambda_c = (m_a v_0)^{-1}$, with $v_0 = 220$ km/s. The limiting cases of $\mathcal{F}_{12}^c(v) \to f(v)$ for $d \ll \lambda_c$ and $\mathcal{F}_{12}^c(v) \to 0$ for $d \gg \lambda_c$ are apparent. For $d \sim \lambda_c$, however, the profile is modulated with the interference inherent in the cross-spectrum. In this simple case, there is no additional information about the velocity distribution that may be extracted by having multiple detectors.

### 4.3.1 The Minimal $\mathcal{N} = 2$ Example

We begin our exploration of the above parameter estimation formalism with a simple scenario: $\mathcal{N} = 2$ detectors measuring DM drawn from an isotropic velocity in the laboratory frame, $4\pi v^2 f(\mathbf{v}) \equiv f(v)$. This example is obviously idealized; in reality, the finite boost velocity of the Sun about the Galactic Center implies that even an isotropic velocity distribution in the Galactic frame will become anisotropic in the laboratory frame. Nonetheless, this example will provide basic intuition for the impact of interferometry.

Invoking isotropy to perform the angular integrals, $\mathcal{F}_{12}^{c,s}(v)$ can be computed as

$$\mathcal{F}_{12}^c(v) = f(v)\frac{\sin(m_a v d)}{m_a v d}\,, \qquad \mathcal{F}_{12}^s(v) = 0\,, \tag{4.40}$$

where again $d$ is the distance between the two detectors. Thus for this example, we see explicitly that for $d \to 0$, we have $\mathcal{F}_{12}^c(v) \to f(v)$, whereas for $d \to \infty$, instead $\mathcal{F}_{12}^c(v) \to 0$. As we will see in the examples below, it is the dispersion $v_0$ rather than the average speed $\bar{v}$ which determines the crossover between small and large $d$.

To progress further, we assume a concrete form for $f(v)$: the Maxwell-Boltzmann distribution,

$$f(v) = \frac{4v^2}{\sqrt{\pi}v_0^3} e^{-v^2/v_0^2}\,, \tag{4.41}$$

where $v_0$ is the velocity dispersion. Taking $v_0 \approx 220$ km/s, this velocity distribution is an approximation to the SHM that is expected to describe the bulk of the local DM, neglecting the finite velocity boost of the Sun relative to the Galactic Center, which breaks the isotropy in the laboratory frame. We will utilize the Maxwellian ansatz repeatedly in this work as an illustrative example. In Fig. 4.2 we show $\mathcal{F}_{12}^c(v)$ for various choices of $d/\lambda_c$; there is a clear deviation from $f(v)$ when $d \sim \lambda_c$, which is a manifestation of the nontrivial correlations in the multi-detector spectrum. Note that we have defined, for the Maxwell-Boltzmann distribution, $\lambda_c = (m_a v_0)^{-1}$, where $v_0$ is the velocity dispersion parameter that enters into (4.41); this is a particular realization of (4.1). Anticipating the more general scenario where the velocity distribution is not isotropic, it is precisely the deviation from $f(v)$ that we will use to extract information about the full velocity profile.

As we have chosen an isotropic $f(\mathbf{v})$, there is no additional information to extract about the velocity distribution in this case. Indeed, the distribution in (4.41) is defined by a single parameter, $v_0$, which we can envision estimating. Evaluating (4.39) analytically in this case, we find

$$\frac{\sigma_{v_0}^2\,\mathrm{TS}_0}{v_0^2} = \frac{8\xi}{9\xi - \xi^3 + \sqrt{2}(15 + 2\xi^2 + \xi^4)F[\xi/\sqrt{2}]}\,, \tag{4.42}$$

written in terms of a dimensionless distance scale $\xi = m_a v_0 d = d/\lambda_c$, and Dawson's integral $F$. We find explicitly that $\sigma_{v_0}$ is minimized for $\xi \to 0$, *i.e.* $d \ll \lambda_c$, since $\partial_{v_0} f(v) \neq 0$.

### 4.3.2 The Infinitely-Cold Stream

We now consider our first example of an anisotropic velocity distribution, a DM stream, and show that we can infer the direction of this stream using DM interferometry. In addition to the bulk SHM, it is expected that the local DM velocity distribution could contain non-virialized substructure, such as cold tidal streams [80, 175, 176, 183, 233–241]. Streams are characterized by low velocity dispersions but large velocity boosts in the solar frame. Let us suppose that in the laboratory frame the stream is boosted at velocity $\mathbf{v}_{\rm str}$ and has velocity dispersion $v_0 \ll |\mathbf{v}_{\rm str}|$. In the limit $v_0 \to 0$, the velocity distribution approaches a delta function,

$$f(\mathbf{v}) = \delta^3(\mathbf{v} - \mathbf{v}_{\rm str}), \tag{4.43}$$

which has an infinite coherence length but a finite de Broglie wavelength. This is clearly an artificial example – it is the maximally anisotropic velocity distribution – but it is one we can evaluate fully analytically. Further, a number of the conclusions that we will reach for the infinitely-cold stream will hold also in more realistic cases. Note that for this example $f(v) = \delta(v - v_{\rm str})$, which has no dependence on the direction of the stream, and therefore a single detector cannot infer the direction.

As claimed, for this simple scenario, we can compute the exact global TS using (4.32), and find

$$\widetilde{\Theta}(\theta_{\rm str}, \phi_{\rm str}) = {\rm TS}_0 \cos\left[m_a d v_{\rm str}(\hat{\mathbf{v}}_{\rm str} - \hat{\mathbf{v}}^t_{\rm str}) \cdot \hat{\mathbf{x}}_{12}\right] . \tag{4.44}$$

We consider the TS as a function of the spherical coordinates of our test stream direction, $\boldsymbol{\alpha} = \{\theta_{\rm str}, \phi_{\rm str}\}$, with the aim being to use the TS to infer the true direction of the stream, given by $\boldsymbol{\alpha}^t = \{\theta^t_{\rm str}, \phi^t_{\rm str}\}$. In this case we can also compute ${\rm TS}_0$, as defined in (4.34), and we obtain[10]

$${\rm TS}_0 = \frac{2A^2 T\pi}{m_a \lambda_B^2} \frac{\delta(v_{\rm str} - v_{\rm str})}{v_{\rm str}} . \tag{4.45}$$

Now consider the angle-dependent factor in (4.44). Without loss of generality, we take $\hat{\mathbf{x}}_{12} = \hat{\mathbf{z}}$ and define spherical coordinates with respect to $\hat{\mathbf{x}}_{12}$, so that the argument of the cosine in (4.44) simplifies to

$$(\hat{\mathbf{v}}_{\rm str} - \hat{\mathbf{v}}^t_{\rm str}) \cdot \hat{\mathbf{x}}_{12} = \cos\theta_{\rm str} - \cos\theta^t_{\rm str} , \tag{4.46}$$

where $\theta_{\rm str}$ and $\theta^t_{\rm str}$ are the usual polar angles in spherical coordinates. Neither azimuthal coordinate $\phi$ appears in this expression, and hence the azimuthal angles are also absent in the TS. This implies we cannot infer one of the angular coordinates of $\mathbf{v}^t_{\rm str}$ from the data. For our particular choice of coordinates, we can infer the parameter $\theta^t_{\rm str}$, as we will describe below, but the likelihood has a flat direction in $\phi_{\rm str}$, so that we cannot infer the associated truth value. The degeneracy is physical. In our coordinates the symmetry of the likelihood is represented by an invariance under changes in $\phi$, but more generally the TS is unchanged by rotations about the detector separation axis, $\hat{\mathbf{x}}_{12}$. This can be seen from the dependence of the TS on $(\hat{\mathbf{v}}_{\rm str} - \hat{\mathbf{v}}^t_{\rm str}) \cdot \hat{\mathbf{x}}_{12}$: any change in the test or true $\hat{\mathbf{v}}_{\rm str}$ that is perpendicular to $\hat{\mathbf{x}}_{12}$ has no impact.

This symmetry of the TS under rotations around $\hat{\mathbf{x}}_{12}$ is in fact not a relic of our idealized example. Our ability to infer the direction of a velocity parameter vector that defines a given $f(\mathbf{v})$ enters through the $\mathbf{v} \cdot \mathbf{x}_{12}$ in $\mathcal{F}^{c,s}_{12}$. But as $\mathbf{v} \cdot \mathbf{x}_{12}$ is itself invariant to rotations of the velocity

[10]Note the fact that ${\rm TS}_0$ formally diverges, ${\rm TS}_0 \propto \delta(0)$, is an artifact of the stream having a delta-function speed distribution. The divergence is regulated by the finite dispersion of the stream, as we discuss below.

about the $\hat{\mathbf{x}}_{12}$ axis, one can show that this flat direction in the likelihood exists generally – indeed we will see it in more realistic cases (a direct analogue is apparent in the symmetry observed in Fig. 4.4, related to the SHM example discussed below). This symmetry will be broken by a dependence in the likelihood on multiple detector axes that are not parallel, provided either by a third detector or alternatively by daily modulation, where the single $\hat{\mathbf{x}}_{12}(t)$ will vary throughout the day at different times $t$. We will explore this latter example in detail in Sec. 4.4 – indeed the optimal detector configuration will be determined by maximally violating this symmetry – but until then the symmetry will represent a basic feature of the physics.

Returning to our specific coordinate system where $\hat{\mathbf{x}}_{12} = \hat{\mathbf{z}}$, we may perform parameter estimation on the angle between the stream and detector. From (4.39), we have

$$\sigma^2_{\theta_{\rm str}} = \frac{2}{\mathrm{TS}_0} \frac{1}{(m_a v_{\rm str} d)^2} \frac{1}{\sin^2 \theta^t_{\rm str}} \,. \tag{4.47}$$

Note that the uncertainty on the parameter $\theta_{\rm str}$ is minimized for $\theta^t_{\rm str} = \pi/2$, *i.e.* when the stream is perpendicular to the two-detector axis. On the other hand, if the two vectors are parallel, $\theta^t_{\rm str} = 0$ or $\pi$, then we see $\sigma_{\theta_{\rm str}}$ diverges. Yet we can still infer $\theta^t_{\rm str}$ in this case. Indeed, looking to (4.46) we see that the asymptotic TS depends on $\theta_{\rm str}$; the likelihood is not globally flat, and we can estimate the angle from contours around the maximum likelihood. Instead, in this case there is a breakdown of the quadratic approximation around the maximum likelihood. If we were to incorporate higher derivatives than in (4.37), we would confirm that the likelihood is not truly flat at these points. This of course should be contrasted with the true flat direction in the likelihood associated with $\phi_{\rm str}$. Note, however, that as $\theta^t_{\rm str}$ approaches either $0$ or $\pi$, becoming parallel to $\mathbf{x}_{12}$, the undetermined parameter $\phi_{\rm str}$ is less relevant. In the limit where the two vectors are parallel, we can infer the true direction of the stream, in spite of this degeneracy.

There is another interesting feature in (4.47): the result suggests that we can take $d \to \infty$ to constrain this one direction of the stream to arbitrary precision. This is a manifestation of our assumption that the stream has no velocity dispersion: it remains coherent over arbitrary large distances, allowing for an improved baseline over which we can measure the stream direction. To study this feature further, imagine making this example slightly more realistic by introducing a finite velocity dispersion $v_0$, with $v_0 \ll v_{\rm str}$, such that $f(\mathbf{v})$ has support in a small volume of radius $\sim v_0$ around $\mathbf{v}_{\rm str}$. For small enough $v_0$ we would expect the results of the $\delta$-function stream to hold. Yet there is an important conceptual difference: the coherence length is no longer infinite because the different waves that constitute the local DM field now have speeds that vary by $\mathcal{O}(v_0)$. Parametrically, the argument of the interferometric terms scale as $m_a|\mathbf{v}||\mathbf{x}_{12}| \sim m_a d(v_{\rm str} + \mathcal{O}(v_0))$, but with the $\mathcal{O}(v_0)$ term varying between states. If we now take $d \gg (m_a v_0)^{-1}$, then the different waves will add incoherently, suppressing the power. But if we choose $d \sim (m_a v_0)^{-1}$, a degree of

coherence can be maintained, along with the interference pattern carrying the information we seek to extract (see also the orange curve $d = 2\lambda_c$ in Fig. 4.2). Accordingly, for the optimal separation, the scaling of the sensitivities in (4.47) is (taking $\sin^2\theta^t_{\rm str} \sim 1/2$ for definiteness)

$$\sigma_\theta \sim \frac{2}{\sqrt{\rm TS_0}}\frac{v_0}{v_{\rm str}} = \frac{2}{\sqrt{\rm TS_0}}\frac{\lambda_{\rm dB}}{\lambda_c}\,. \tag{4.48}$$

In the more realistic examples we will confirm the conclusion that $d \sim \lambda_c$ provides the maximum sensitivity. There is another consequence of this choice. Taking $d = (m_a v_0)^{-1}$ in (4.44), the prefactor of the dot product in (4.46) is $v_{\rm str}/v_0 \gg 1$, by definition of this being a cold stream. Small variations in $(\hat{\mathbf{v}}_{\rm str} - \hat{\mathbf{v}}^t_{\rm str})\cdot\hat{\mathbf{x}}_{12}$ will induce large variations in the argument of the cosine, implying that the global structure of the TS is highly nontrivial. Although the maximum TS will be attained at the true $\theta$, there will be a pattern of local maxima with comparable TS (this result is depicted in Fig. 4.5, and persists even with daily modulation as shown in Fig. 4.7).

### 4.3.3 The (boosted) Standard Halo Model

The bulk DM halo of the Milky Way is expected to be Maxwell-Boltzmann distributed as in (4.41) in the Galactic frame, except for a possible cut-off around the escape velocity ∼500 km/s [242]. On the other hand, the Sun is boosted with respect to the Galactic frame by [243]

$$\mathbf{v}_\odot \approx (11,\, 232,\, 7)\,{\rm km/s}\,, \tag{4.49}$$

in Galactic coordinates, where $\hat{\mathbf{x}}$ points towards the Galactic Center, $\hat{\mathbf{y}}$ points in the direction of the local rotation of the disk, and $\hat{\mathbf{z}}$ points towards the Galactic north pole. Thus in the laboratory frame (neglecting the Earth's motion), the velocity distribution becomes that of the SHM,

$$f(\mathbf{v}) = \frac{1}{\pi^{3/2}v_0^3}e^{-(\mathbf{v}+\mathbf{v}_\odot)^2/v_0^2}\,, \tag{4.50}$$

with a velocity dispersion $v_0 \approx 220$ km/s [185, 193]. Note in particular that $v_0 \sim |\mathbf{v}_\odot| \equiv v_\odot$, so for the SHM $\lambda_c \sim \lambda_{\rm dB}$. The associated speed distribution is

$$f(v) = \frac{v}{\sqrt{\pi}v_0 v_\odot}e^{-(v+v_\odot)^2/v_0^2}\left(e^{4vv_\odot/v_0^2} - 1\right). \tag{4.51}$$

As we have emphasized many times already, single detectors are only sensitive to the speed distribution, which only depends on $v_\odot$ but not the orientation of the solar velocity $\hat{\mathbf{v}}_\odot$. Thus, a single detector may constrain the model parameters $v_0$ and $v_\odot$ (as shown in [80]), but determining the

orientation requires multiple detectors.[11]

To determine the expected sensitivity to the direction $\hat{\mathbf{v}}_\odot$ we need to compute the derivatives of $\mathcal{F}_{12}^{c,s}(v)$ that appear in (4.39):

$$\partial_{\theta_\odot}\mathcal{F}_{12}^{c}(v)\big|_{\theta_\odot=\theta_\odot^t} = \frac{4v^3 v_\odot e^{-(v^2+v_\odot^2)/v_0^2}}{\sqrt{\pi}v_0^5}\int d\theta \sin\theta \, \exp\left[-\frac{2vv_\odot}{v_0^2}\cos\theta\cos\theta_\odot^t\right]\cos(m_a v d\cos\theta) \times\left[I_0\left[\frac{2vv_\odot}{v_0^2}\sin\theta\sin\theta_\odot^t\right]\cos\theta\sin\theta_\odot^t + I_1\left[\frac{2vv_\odot}{v_0^2}\sin\theta\sin\theta_\odot^t\right]\sin\theta\cos\theta_\odot^t\right], \tag{4.52}$$

where $I_{0,1}$ are both modified Bessel functions (an analogous expression holds for $\mathcal{F}_{12}^{s}$). In computing this result we have again chosen coordinates $\hat{\mathbf{x}}_{12} = \hat{\mathbf{z}}$, but left the direction of $\hat{\mathbf{v}}_\odot$ arbitrary, defined by $(\theta_\odot^t, \phi_\odot^t)$. The most important feature of this result is that it exhibits no dependence upon $\phi_\odot^t$: again, there is a symmetry in the likelihood for rotations around $\hat{\mathbf{x}}_{12}$. Beyond this, we can also see that the derivative vanishes when $\mathbf{v}_\odot$ is parallel to the detector separation ($\theta_\odot^t = 0$). Accordingly, in this case we will find $\sigma_{\theta_\odot}$ diverges, as we did for the stream. But again this is not a global flat direction in this case; the likelihood is just sufficiently flat at the maximum that the first three derivatives vanish.

To proceed beyond these analytic insights, we will compute the remaining results numerically. We define the angle between $\mathbf{v}_\odot$ and $\mathbf{x}_{12}$ as $\theta_\odot$. To begin with, we take a generic value of $\theta_\odot^t = \pi/4$ and consider how well we can infer this angle as a function of detector separation. The results are shown in Fig. 4.3. Unlike for the $\delta$-function stream, there is now a minimum at a finite value of $d$, and as argued on general grounds this occurs when $d \sim \lambda_c = (m_a v_0)^{-1}$. That the uncertainty diverges for $d \to 0$ is consistent with the fact that a single detector cannot infer this direction. In more detail we find the minimum occurs at $d \sim 2\lambda_c$, where we obtain $\sigma_{\theta_d} \approx 2/\sqrt{\mathrm{TS}_0}$. For example, if $\mathrm{TS}_0 = 25$, corresponding to a $5\sigma$ local significance detection with $d = 0$, then at the distance $d \sim 2\lambda_c$ corresponding to minimum uncertainty, the solar velocity direction with respect to the detector axis could be localized to $0.4\ \mathrm{rad} \sim 20^\circ$ on the sky. We can understand the magnitude of $\sigma_{\theta_d}$ at its minimum from (4.48): the SHM has the form of a stream where $v_0 \sim v_{\mathrm{str}} = v_\odot$, and therefore we would expect $\sigma_{\theta_d} \times \sqrt{\mathrm{TS}_0} \sim 2$, exactly as observed.

However, as we have emphasized already, it is important to keep in mind that our estimate of $\sigma_{\theta_\odot}$ is a measure of the expected curvature of the likelihood in the vicinity of the true value and does not capture the global structure of the expected likelihood function. To illustrate these features we fix $d = 2\lambda_c$ for definiteness and illustrate the global map $\widetilde{\Theta}(\theta_\odot, \phi_\odot)/\mathrm{TS}_0$ in Fig. 4.4, for three different values of $\theta_\odot^t$. Note that we have divided out the overall significance $\mathrm{TS}_0$, so that exactly how well we can localize the direction will depend on how significantly the DM signal

[11] In principle annual modulation may be used by a single detector to infer $\hat{\mathbf{v}}_\odot$, as discussed in [80, 183].

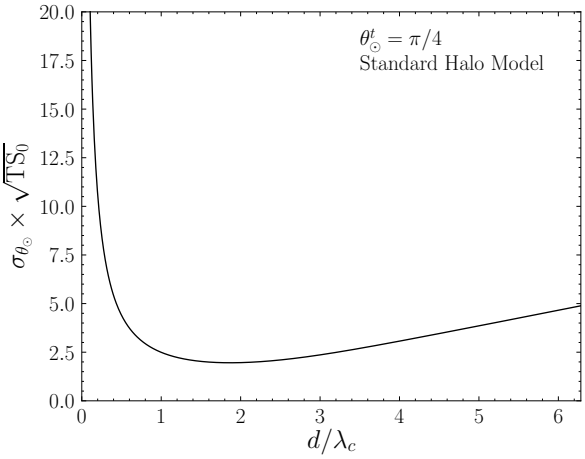



Figure 4.3: The expected uncertainty on the angle between the detector axis and solar velocity, $\theta_\odot = \arccos(\hat{\mathbf{v}}_\odot \cdot \hat{\mathbf{x}}_{12})$, as a function of $d/\lambda_c = d \times m_a v_0$. In this example we have set the true orientation to $\theta^t_\odot = \pi/4$. With this configuration, we find that the maximum precision is obtained for $d \approx 2\lambda_c$.

has been measured. However, the expected global structure of the TS will be a rescaled version of these maps. In each case the true $\hat{\mathbf{v}}_\odot$ that we are seeking to infer is located in the center of the Mollweide projection maps. The left panel illustrates the scenario with $\hat{\mathbf{x}}_{12} = \hat{\mathbf{v}}_\odot$ ($\theta_\odot = 0$), the center has $\theta_\odot = \pi/4$, while in the right panel two directions are perpendicular and $\hat{\mathbf{x}}_{12}$ points between the poles of the map ($\theta_\odot = \pi/2$). In all cases the symmetry of the TS around the $\hat{\mathbf{x}}_{12}$ axis is apparent. The only case where this flat direction in the maximum TS is not an obstruction to determining the true direction of $\hat{\mathbf{v}}_\odot$ is when $\theta_d = 0$. In that case we are still able to localize the true direction, although we note the likelihood is relatively flat around the maximum (consistent with the second derivative vanishing). In Sec. 4.4 we illustrate how daily modulation generically allow us to fully determine both of the angles associated with the direction of $\mathbf{v}_\odot$.

### 4.3.4 The Sagittarius Stream

As a final example working with a single static $\hat{\mathbf{x}}_{12}$, we return to the case of the cold stream with non-vanishing velocity dispersion. We expect many of the conclusions reached in Sec. 4.3.3 to hold in this case. In particular, the symmetry around the $\hat{\mathbf{x}}_{12}$ axis will remain, but we will see explicitly in this case the non-trivial structure induced in the global likelihood by the ratio of $v_0/v_{\rm str} \sim \lambda_{\rm dB}/\lambda_c \gg 1$. To make the example concrete, the DM component of the Sagittarius stream may extend to the Sun's location, and estimates [194, 244] suggest that it could make up $\sim$5% of the local DM density. However, the DM associated with the stream would be highly collimated in phase space; we follow [80] and model the Sagittarius stream DM velocity distribution by a boosted Maxwellian as in (4.50), but with $v_0 = 10$ km/s and $\mathbf{v}_\odot$ replaced

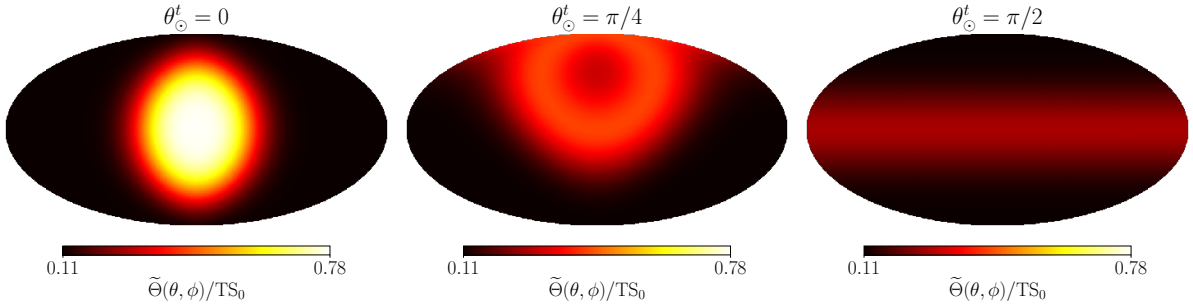



Figure 4.4: (*Left*) A Mollweide projection of the Asimov test statistic $\widetilde{\Theta}(\theta, \phi)$ for the SHM divided by the co-located detection significance $\mathrm{TS}_0$. The detectors are configured so that the displacement vector between them is parallel to the SHM boost velocity, and the Mollweide plot is rotated so that it is centered on the maximum test statistic. (*Center*) As on the left, but for a detector configuration where the displacement vector is at a $45^\circ$ angle to the North ($\theta^t_\odot = \pi/4$) with respect to the SHM boost velocity. (*Right*) As on the left, but for a detector configuration where the displacement vector is perpendicular to the SHM boost velocity ($\theta^t_\odot = \pi/2$). In this configuration the location of the boost velocity can only be localized to a great circle on the celestial sphere.

by $\mathbf{v}_{\rm str} = (0, 93.2, -388)$ km/s [175, 176]. We consider the stream in isolation, as opposed to in conjunction with the bulk SHM DM phase-space distribution, because even though the stream component is sub-dominant in terms of DM density, it still dominates in the narrow region of phase space where the Sagittarius stream has compact support. To simplify the discussion we will simply take $v_{\rm str} = 400$ km/s, with a direction that we will again specify by its angle with respect to the detector axis (given the degeneracy in rotations about that axis). Note that this example could apply equally well to other putative DM streams, such as the newly discovered S1 stream [239–241].

To begin with, in the left panel of Fig. 4.5 we show the expected uncertainty on the recovered angle between the stream and detector, $\theta_{\rm str}$, as a function of the distance in units of $\lambda_c = (m_a v_0)^{-1}$, for a true value $\theta^t_{\rm str} = \pi/4$. This figure is the stream analogue of what we showed for the SHM in Fig. 4.3. Once more, following the general discussion in Sec. 4.3.2, the optimal sensitivity is achieved for $d \sim \lambda_c$, and from (4.48), we expect $\sigma_{\theta_d} \times \sqrt{\mathrm{TS}_0} \sim 2v_0/v_{\rm str} \sim 0.05$ at the minimum-uncertainty distance, compatible with what we see in Fig. 4.5.

However, just like in the case of the SHM it is important to also examine the global properties of the TS in addition to the curvature of the expected TS at the true parameter values. Towards that end, on the right of Fig. 4.5 we show the expected TS $\widetilde{\Theta}$, normalized to $\mathrm{TS}_0$, as a function of the reconstructed angle between the stream and detector, $\theta_{\rm str}$. For this figure we have fixed the true orientation at $\theta^t_{\rm str} = \pi/4$ along with the separation $d = 2\lambda_c$. We see that $\widetilde{\Theta}$ drops off quickly around the true value of $\theta_{\rm str} = \pi/4$ (vertical dashed), but that there is non-trivial structure with local maxima at larger and smaller $\theta_{\rm str}$ values. This is a direct manifestation of the non-trivial interference patterns discussed in Sec. 4.3.2 for cold streams: the large ratio $v_{\rm str}/v_0$ enters into the

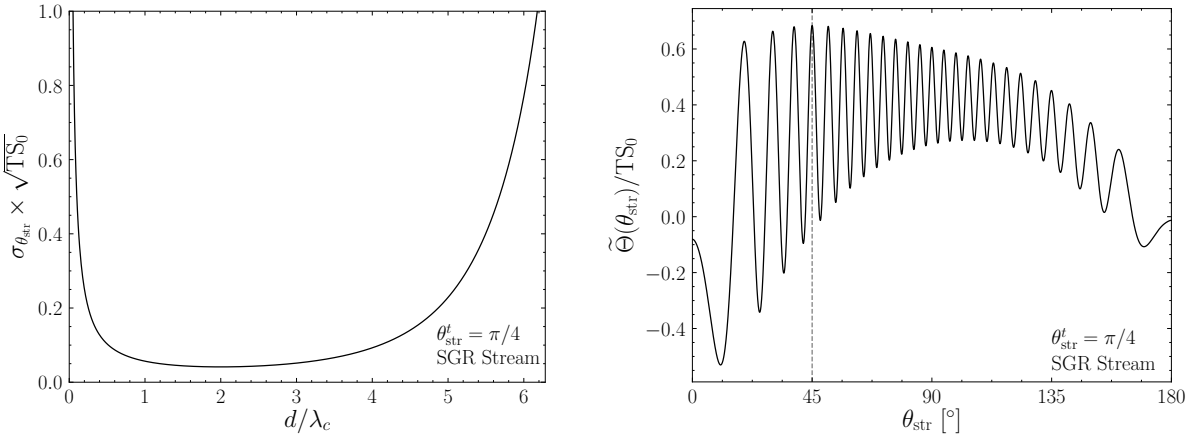



Figure 4.5: (*Left*) As in Fig. 4.3, but for the Sagittarius (SGR) stream rather than for the SHM. As before, the maximum precision for the inferred value of $\theta_{\rm str}$ is achieved at $m_a v_0 d \approx 2$, although the overall dependence is somewhat softened outside of the extremes at $m_a v_0 d = 0$ and $m_a v_0 d = 2\pi$. The values of $\sigma_{\theta_{\rm str}} \times \sqrt{\rm TS_0}$ are also considerably smaller than those found in the SHM example, indicating that the angle $\theta_{\rm str}$ can be reconstructed with much greater precision for the SGR stream as compared to the SHM. (*Right*) The Asimov TS $\widetilde{\Theta}(\theta_{\rm str})$ for the SGR stream rescaled by the co-located detection significance ${\rm TS_0}$ as a function of $\theta_{\rm str}$ for a detector configuration where the true stream direction is $\theta^t_{\rm str} = \pi/4$ (dashed vertical line). We have fixed $m_a v_0 d = 2$. The TS $\widetilde{\Theta}(\theta_{\rm str})$ is maximized at the true value of $\theta_{\rm str}$, but there is considerable nontrivial global structure with a large number of local minima and maxima in $\widetilde{\Theta}$.

argument of the trigonometric functions in $\mathcal{F}_{12}^{c,s}(v)$.

## 4.4 Daily Modulation

One of the most dramatic signatures of DM interferometry is the unique daily modulation signal available to multiple detectors. This effect, which we describe in the current section, would be a smoking gun signature that an emerging excess has a DM origin, and it also allows two detectors to better determine geometric parameters describing the velocity distribution. The basic idea is simply that for two detectors fixed at generic locations on the surface of the Earth, the separation vector $\mathbf{x}_{12}$ is rotating in the inertial Galactic frame throughout the day. This is in contrast to the angular parameters entering in the DM velocity distribution, such as the Solar direction $\hat{\mathbf{v}}_\odot$, which should always point in the same Galactic direction, regardless of the orientations of the detectors at any point in time on Earth. The rotation of $\mathbf{x}_{12}$ with respect to the fixed $\hat{\mathbf{v}}_\odot$ implies that we will sample a variety of angles between the two vectors, and therefore vary the modulation of the speed distributions in $\mathcal{F}_{12}^{c,s}(v)$, as already depicted in Fig. 4.1. Critically this will lift the flat direction in the maximum likelihood associated with rotations around $\mathbf{x}_{12}$ that we observed repeatedly in Sec. 4.3: as the likelihood will now depend on a collection of different vectors $\mathbf{x}_{12}(t)$

, the symmetry that exists around any one of them will not be preserved in the full TS.

In the rest of this section we divide the discussion into three parts. Firstly, we describe how to construct the likelihood for the generic case of $\mathcal{N}$ detectors incorporating daily modulation and describe how it is straightforward to generalize our full formalism to this case. We then focus on the specific case of $\mathcal{N} = 2$ and show, within the Asimov formalism, how the examples of the SHM and Sagittarius stream discussed above are modified in the presence of daily modulation. Finally, we turn to a Monte Carlo simulation of a realistic example and demonstrate how, within a day, a resonant experiment could constrain the direction of the solar velocity vector, $\mathbf{v}_\odot$, that controls the SHM to sub-degree accuracy.

### 4.4.1 A Likelihood with Daily Modulation

So far in this work, we have envisioned a set of $\mathcal{N}$ experiments collecting measurements of the signal-plus-background frequency spectra for a duration of time $T$ while the detector separations were fixed with respect to the boost velocity of the DM component under consideration. However, this framework cannot be extended to the case of daily modulation, as the signal prediction will fundamentally be varying over a 24-hour period. In order to properly account for this effect, the data must be collected in time intervals of duration $T \ll 24$ hours and analyzed with a joint likelihood over all the collected intervals. In detail, if we imagine that we collect $M$ such time intervals, indexed by $r = 0, 1, \ldots, M-1$, then for each of these we will have a data set $d_r = \{\mathbf{d}_{k,r}\}$, where again $k$ labels the Fourier mode. For each data set $d_r$, we can compute the likelihood as in (4.19), and the full joint likelihood is the product of these over $r$. Explicitly, we have

$$\mathcal{L}(d|\mathcal{M}, \boldsymbol{\theta}) = \prod_{r=0}^{M-1} \mathcal{L}(d_r|\mathcal{M}, \boldsymbol{\theta}) = \prod_{r=0}^{M-1} \prod_{k=0}^{N-1} \frac{\exp\left[-\frac{1}{2}\mathbf{d}_{k,r}^T \cdot \boldsymbol{\Sigma}_{k,r}^{-1}(\boldsymbol{\theta}) \cdot \mathbf{d}_{k,r}\right]}{\sqrt{(2\pi)^{2\mathcal{N}}|\boldsymbol{\Sigma}_{k,r}(\boldsymbol{\theta})|}}. \tag{4.53}$$

Importantly, note that we have also attached an index $r$ to the signal prediction $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, as we need to account for the variation of the detector separations $\mathbf{x}_{ij}$ throughout the day.

In a similar fashion the full formalism of Secs. 4.2 and 4.3 can be extended to include the varied detector orientation: within a given sub-interval we simply adjust $\mathbf{x}_{ij}$ as appropriate, and then we form joint quantities by combining these as in the likelihood above. To provide just a single illustrative example, the Fisher information computed in (4.36), would become

$$I_{ij}(\boldsymbol{\alpha}) = -\frac{1}{2}\sum_{r=0}^{M-1} \frac{\partial^2 \Theta_r(\boldsymbol{\alpha})}{\partial\alpha_i \partial\alpha_j}\,, \tag{4.54}$$

with other expressions similarly generalized.

### 4.4.2 Asimov Examples with Daily Modulation

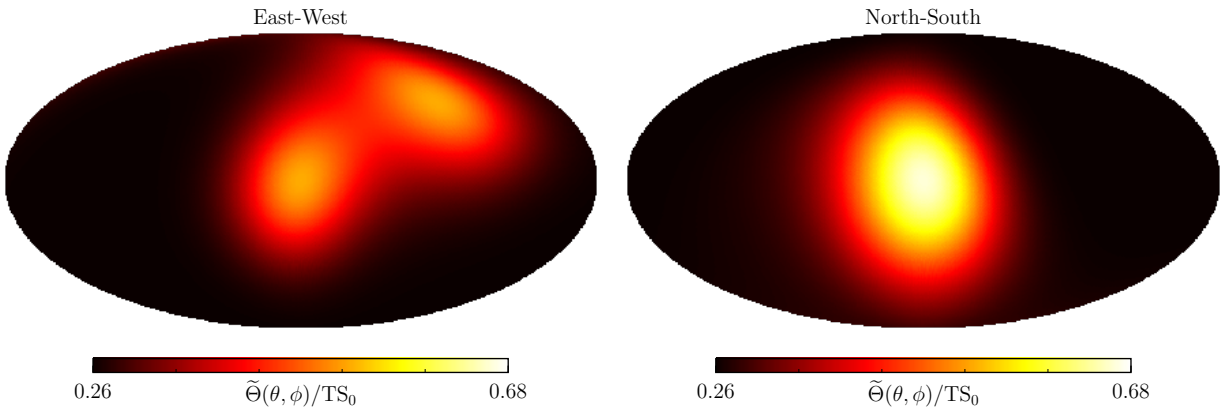



Figure 4.6: As in Fig. 4.4, we construct Mollweide projections of the Asimov test statistic $\widetilde{\Theta}(\theta, \phi)$ for the SHM rescaled by the co-located detection significance $\mathrm{TS}_0$. However, we now perform a joint likelihood over data collected over a 24-hour period so that the daily modulation of the detector displacement vector produces a time-varying signal, which helps break degeneracies in the reconstructed directional parameters. The Mollweide projection for a configuration in which the detectors are oriented along an East-West (North-South) orientation is shown on the left (right). While the results of obtained in an East-West configuration do not depend on the latitude of the detectors, the North-South configuration results do, so for definiteness, we have taken the detectors to be located in New Haven, CT, the site of the HAYSTAC detector. In both configurations, the SHM boost velocity direction can be localized effectively, although there remains a non-trivial degeneracy in the East-West map between two points on the sphere.

While the alteration to our formalism imposed by daily modulation is minimal – as exhibited in (4.53) – the impact on the results can be dramatic. We will demonstrate this with several examples in this section, all within the Asimov formalism. To begin with, we consider using $\mathcal{N} = 2$ detectors in order to determine the direction of $\mathbf{v}_\odot$ in the SHM. This is the same problem we considered in Sec. 4.3.3, which produced the results shown in Fig. 4.4, where there is a clear degeneracy associated with rotations around $\mathbf{x}_{12}$. We will now see explicitly that daily modulation helps lift this degeneracy. To do so, let us suppose that the DM velocity distribution follows the SHM in (4.50), with $v_0 = 220$ km/s and $v_\odot = 232$ km/s. Our goal, as previously, will be to infer the direction of $\hat{\mathbf{v}}_\odot$. We consider two detectors separated by $d = 2\lambda_c = 2/(m_a v_0)$, and for definiteness we place one detector at a latitude $41.3°$ N and longitude $72.9°$ W. In Fig. 4.6 we show results where a second detector is placed a distance $d$ to the East (left) or North (right) of this detector with data stacked at two-hour intervals over the 24-hour period.[12] For the North-South configuration, we

[12]Note that since the Earth's rotation is aligned with the East-West direction, results obtained for the East-West configuration are independent of the exact experimental locations, so long as the detector separation is much smaller than the Earth's radius of curvature. For any other configuration, however, the result will generically depend on latitude.

see that the direction can be well-localized: a high significance axion detection in this case would lead to a precise estimation of the direction of $\mathbf{v}_\odot$, as we show explicitly in Sec. 4.4.3 below. This configuration clearly outperforms an East-West configuration, where there remains a degeneracy that has not been fully lifted by the daily modulation. Additionally, the maximum test statistic realized in the North-South configuration would be approximately 10% larger than one realized in an East-West configuration for otherwise identical data collections.

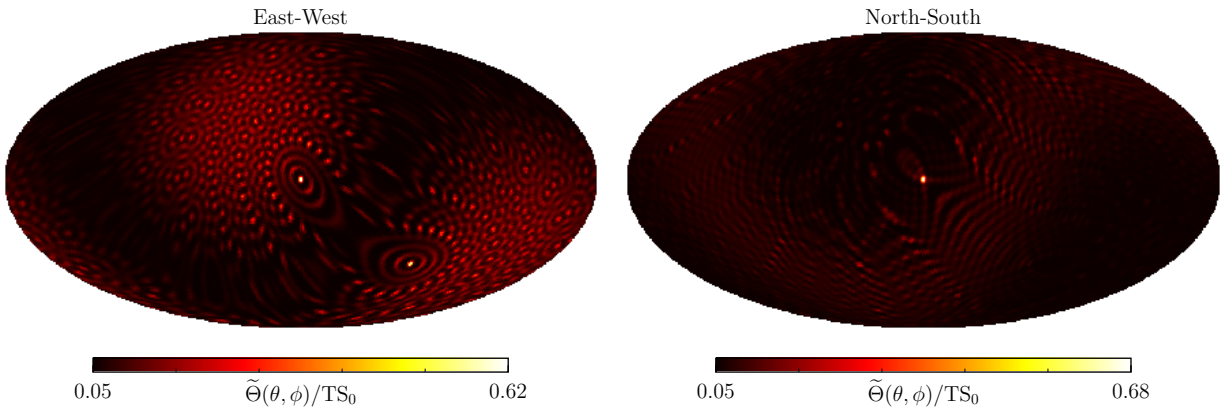



Figure 4.7: As in Fig. 4.6, but for the Sagittarius stream example. For a fixed axion mass, the physical detector separation $d = 2\lambda_c$ is a factor of 20 larger than in Fig. 4.6 because of the larger coherence length of the stream. While there are many local maxima in both configurations, the North-South orientation produces only a single global maximum, at the true detector localization, while the East-West orientation leads to two degenerate global maxima (one at the true detector location and the other displaced). An animated version of these figures, showing how the localization improves throughout the day as more orientations of $\mathbf{x}_{12}$ are sampled, can be found at github.com/joshwfoster/DM_Interferometry.

Using the same experimental design, we can also revisit the example of the Sagittarius stream discussed in Sec. 4.3.2. In Fig. 4.7 we construct the analogue of Fig. 4.6, but now for the much colder stream. Note that since $v_0$ for the stream is a factor of $\sim 20$ smaller than for the SHM, the optimal detector distance $d = 2\lambda_c$ is a factor of 20 larger than in Fig. 4.6. Although in both configurations the TS is maximized at the expected location on the sphere, nontrivial structure due to the presence of many local maxima are apparent in both the North-South and East-West configurations. We note that, as in the SHM example, there is only one global maximum for the North-South configuration, located at the true direction of the stream. However, there remains a degeneracy in the East-West configuration.

The degeneracy represented in the Mollweide maps for the SHM and the Sagittarius stream in the East-West configuration is exact. It has its origin in the dimensionality of the space swept out by the detector separation vector $\mathbf{x}_{12}$ over the course of the day. As studied in Sec. 4.3, for data taken at fixed $\mathbf{x}_{12}$, the test statistic $\widetilde{\Theta}(\hat{\mathbf{v}})$ evaluated as a function of the orientation of the boost velocity

depends only on the angle between $\hat{\mathbf{v}}$ and $\mathbf{x}_{12}$. As a result, $\widetilde{\Theta}(\hat{\mathbf{v}}^t) = \widetilde{\Theta}(\hat{\mathbf{v}}')$ where $\hat{\mathbf{v}}^t$ is the true boost direction and $\hat{\mathbf{v}}'$ is a velocity obtained by reflecting $\hat{\mathbf{v}}^t$ across any plane which contains $\mathbf{x}_{12}$. For detectors in an East-West configuration, the Earth's rotation produces a daily modulation of $\mathbf{x}_{12}$ that is confined to the plane orthogonal to the Earth's rotational velocity vector. As a result, the TS measured at each point in the day, and therefore the sum of such TSs, will be exactly preserved under reflections of the boost velocity across that plane. This means that accounting for daily modulation in the East-West configuration the directional parameters can only be determined up to a reflection across the plane perpendicular to the Earth's rotation axis. By contrast, for detectors in the North-South configuration, the set of detector separation vectors throughout the day will generically not be co-planar, and thus there is no analogous degeneracy.[13]

### 4.4.3 Monte Carlo Example with Daily Modulation

As a realistic demonstration of our ability to perform parameter estimation using the daily modulation effect, we generate a Monte Carlo realization of data in the North-South SHM scenario (as depicted in the right panel of Fig. 4.6) using $A = 38.25$ and $\lambda_B = 1$, both of which we take to be dimensionless without loss of generality. The values of $A$ and $m_a$ was chosen to generate a signal of expected $5\sigma$ significance during a 100-second collection in a single detector to mimic a realistic resonant scanning strategy in which one of two independently operated detectors detects an excess and both are then used for a 24-hour observation of the excess candidate. We constructed 24 hours of Monte Carlo data for this signal, taking a detector separation of $d = 2\lambda_c$; with these parameters, the excess would be expected to appear at $\mathrm{TS}_0 \approx 60,000$ after 24 hours. While large, this $\mathrm{TS}$ is consistent with the power of a resonant strategy once the axion mass is known.

Using uniform priors on $A$ between $[33, 43]$, on $v_\odot$ between $[212.5, 252.5]$ km/s, on $v_0$ between $[200, 240]$ km/s, on $\lambda_B$ between $[.999, 1.001]$ and a uniform prior on the sphere for $(\theta_\odot, \phi_\odot)$, we construct a Bayesian posterior distribution for the model parameters. The results of an analysis performed using Multinest [202, 245–247] with 2,000 live points are shown in Fig. 4.8. In particular, we see that the true location of the stream has been located to degree precision. This precision can be understood from (4.48), which gives the expectation $\sigma_{\theta_\odot} \sim 0.5^\circ$, consistent with what is shown in the figure. Let us suppose that the Sagittarius stream, as modeled in this work, comprises 10% of the local DM. In the example above, we would expect that after 24 hours the location of the stream could be localized to $\sim 10'$; interestingly, this represents greater accuracy for stream localization than localization of the bulk SHM even though the stream is a sub-dominant component

[13] An exception occurs if the two detectors have the same longitude and equal and opposite latitudes (*i.e.*, opposite sides of the equator on the same line of longitude). An extreme example would be having one detector at each pole. Then, $\mathbf{x}_{12}$ is parallel to the rotation axis of the Earth and does not change direction throughout the day. Consequently, daily modulation provides no additional information, and the full degeneracy that was present throughout Sec. 4.3 returns.

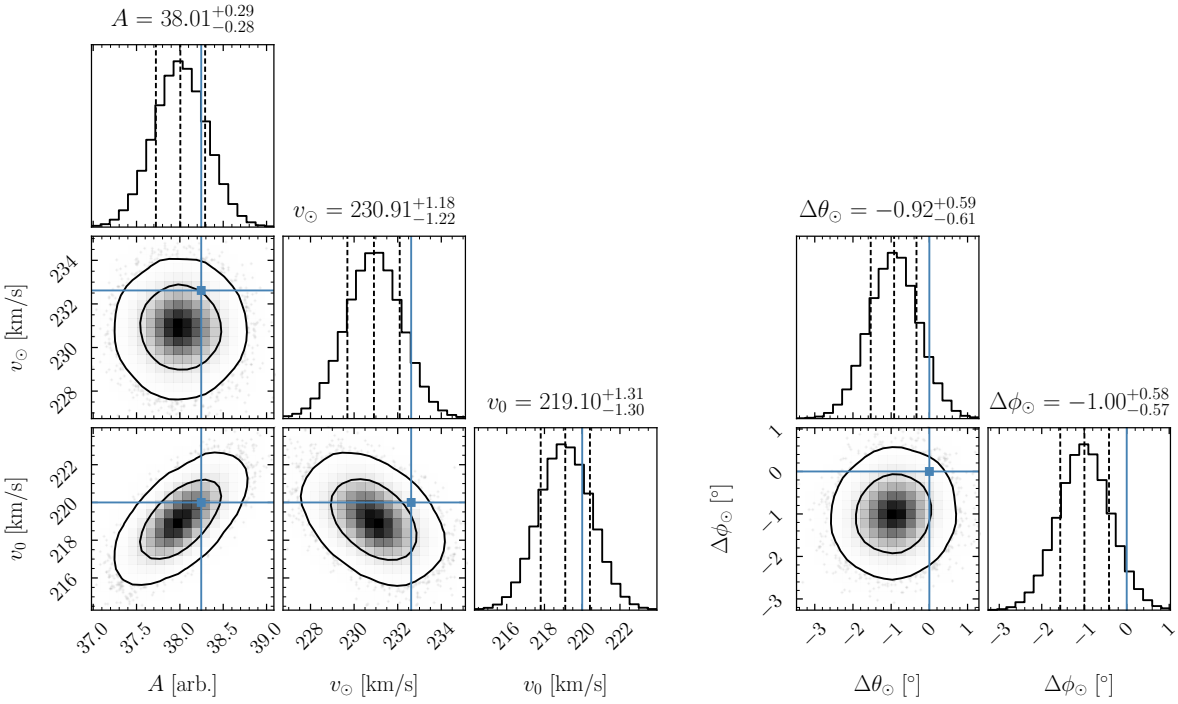



Figure 4.8: The posterior distribution for a model with daily modulation where the signal strength is at the threshold of an expected $5\sigma$ detection for a 100 second observation with a single detector. Monte Carlo data is generated for 24 hours of data collection with two detectors separated along the North-South direction by a distance with $2 \times (m_a v_0)^{-1}$. The true parameters are indicated in blue, with the $1\sigma$ confidence intervals on the parameter estimations are indicated by the dashed black lines in the single-parameter posteriors. The two parameter posteriors show the $1\sigma$ and $2\sigma$ contours. On the left, we display the posterior distributions for the overall signal strength, the boost speed of the SHM, and the velocity dispersion of the SHM, all of which are parameters accessible in a single detector configuration. On the right, we display the posterior distributions for the angles $\Delta\theta_\odot = \theta_\odot - \theta_\odot^t$ and $\Delta\phi_\odot = \phi_\odot - \phi_\odot^t$ which specify the orientation of SHM boost velocity and are only accessible in a multiple-detector configuration. Both $\theta_\odot$ and $\phi_\odot$ are determined at degree precision in this scenario.

of the DM.

## 4.5 Conclusion

In this work we have demonstrated the power of DM interferometry for wave-like DM. The spatial coherence of the DM field imprints phase correlations on the signals observed at spatially-separated detectors, and these phase correlations are sensitive to parameters in the full 3-dimensional velocity distribution $f(\mathbf{v})$, whereas a single detector is blind to all effects beyond the speed distribution $f(v)$. As a result, the advantages of DM interferometry go beyond a simple coherent enhancement of the signal strength as the number of detectors is increased. By taking advantage of the fact

that the correlation matrix of the Fourier-transformed signals at multiple detectors depends on modified speed distributions which contain modulated forms of $f(\mathbf{v})$, we have demonstrated that parameters such as the solar velocity vector may be reliably extracted from two detectors separated by a distance $d \sim \lambda_c$. Furthermore, directional parameters of coherent substructure such as DM streams may be estimated at even higher significance, though in that case the optimal separation $\lambda_c$ is parametrically different from the DM de Broglie wavelength $\lambda_{\rm dB}$.

Our formalism has immediate practical applications for new and upcoming axion DM experiments. The sensitivity to $g_{a\gamma\gamma}$ for resonant-cavity axion experiments which use external magnetic fields, like ADMX and HAYSTAC, is typically $BV^{1/2}$, where $B$ is the peak magnetic field strength and $V$ is the magnetic field volume. In order to achieve resonant enhancement, the volume of an individual cavity is fixed to be of order $1/m_a^3$, so to achieve greater sensitivity, one must either increase the $B$-field or construct a multiplexed readout with multiple cavities. Assuming the latter strategy is chosen, our results motivate placing at least one of the cavities at a distance $\lambda_c$: if a signal is detected, the loss of coherent enhancement of the signal is more than compensated by the ability to localize the boost direction of the DM velocity distribution to within 1 degree with just 24 hours of data.

While there are many challenges to the construction of additional instruments, we emphasize that all of the important phenomenology is captured by a two-detector array. This smoking-gun signature of DM is invisible to a multiplexed setup where all cavities lie inside a single coherence length. A similar analysis applies to experiments in the quasistatic regime like ABRACADABRA and DM-Radio, where the physical volume of the experiment is decoupled from $m_a$. For both types of experiments, our formalism may be applied to determine the optimal detector orientation for localizing the solar velocity to the desired precision (with North-South orientations generally being preferred to East-West). The optimal detector separation corresponds to physically reasonable distances for well-motivated axion masses – $\mathcal{O}(10)$ m for the $10^{-5}$ eV mass range of ADMX and HAYSTAC and $\mathcal{O}(1000)$ km for the $10^{-9}$ eV targeted by ABRACADABRA/DM-Radio – and as such the coherence length and detector orientation can form an important design parameter for future experiments, in much the same way as $L/E$ determines the design of neutrino oscillation experiments.

The future of axion detection involves readout beyond the standard quantum limit, using tools such as Josephson parametric amplifiers and squeezed states. In this regime, it is important to note that our variables $R_k$ and $I_k$ are canonically conjugate, and thus cannot be simultaneously measured to arbitrary precision. In future work, we plan to investigate how our formalism must be modified for quantum-limited readouts. As the number of new axion experiments proliferates, this work motivates careful consideration of the spatial configuration of multiplexed detectors.

# CHAPTER 5

# First Results from ABRACADABRA-10 cm

The particle nature of dark matter (DM) in the Universe remains one of the greatest mysteries of contemporary physics. Axions are an especially promising candidate as they can simultaneously explain both the particle nature of DM and resolve the strong-CP problem of quantum chromodynamics (QCD) [38, 43–46, 86]. axion-like particle (ALP) are generically expected to have a coupling to electromagnetism of the form [190]

$$\mathcal{L} \supset -\frac{1}{4} g_{a\gamma\gamma} a \widetilde{F}_{\mu\nu} F^{\mu\nu} = g_{a\gamma\gamma} a \, \mathbf{E} \cdot \mathbf{B}, \tag{5.1}$$

where $g_{a\gamma\gamma}$ is the axion-photon coupling. The QCD axion is predicted to have a narrow range of couplings proportional to the axion mass, while a general ALP may have any $g_{a\gamma\gamma}$. In this work, "axion" refers to a general ALP. axion dark datter (ADM) with mass $m_a \ll 1\,\text{eV}$ behaves today as a classical field oscillating at a frequency $f = m_a/(2\pi)$ [46, 86]. The Lagrangian (5.1) implies that a time-dependent background density of ADM modifies Maxwell's equations. In particular, in the presence of a static magnetic field $\mathbf{B}_0$, ADM generates an oscillating magnetic field, $\mathbf{B}_a$, as if sourced by an effective AC current density parallel to $\mathbf{B}_0$ [59],

$$\mathbf{J}_{\text{eff}} = g_{a\gamma\gamma} \sqrt{2\rho_{\text{DM}}} \mathbf{B}_0 \cos(m_a t). \tag{5.2}$$

Here $\rho_{\text{DM}}$ is the local DM density, which we take to be $0.4\,\text{GeV/cm}^3$ [248, 249]. The ABRACADABRA experiment, as first proposed in [71], is designed to search for the axion-induced field, $\mathbf{B}_a$, generated by a toroidal magnetic field (see also [147] for a proposal using a solenoidal field). ABRACADABRA searches for an AC magnetic flux through a superconducting pickup loop in the center of a toroidal magnet, which should host no AC flux in the absence of ADM. The time-averaged magnitude of the flux through the pickup loop due to $\mathbf{B}_a$ can be written as

$$|\Phi_a|^2 = g_{a\gamma\gamma}^2 \rho_{\text{DM}} V^2 \mathcal{G}^2 B_{\text{max}}^2 \equiv A, \tag{5.3}$$

where $V$ is the volume of the toroid, $\mathcal{G}$ is a geometric factor calculated for our toroid to be 0.027 [82], and $B_{\max}$ is the maximum $B$-field in the toroid. The pickup loop is read out using a SQUID current sensor, where an axion signal would appear as a small-amplitude, narrow ($\Delta f/f \sim 10^{-6}$) peak in the power spectral densitie (PSD) of the SQUID output at a frequency given by the axion mass. The present design uses a simplified broadband readout, but the same approach can be significantly enhanced using resonant amplification and recent developments in powerful quantum sensors [221, 250], which is the subject of future work.

In this Letter, we present first results from ABRACADABRA-10 cm, probing the axion-photon coupling $g_{a\gamma\gamma}$ for ADM in the frequency range $f \in [75\,\mathrm{kHz},\ 2\,\mathrm{MHz}]$, corresponding to axion masses $m_a \in [3.1 \times 10^{-10},\ 8.3 \times 10^{-9}]$ eV. This mass range is highly motivated for QCD axions, where the axion decay constant lies near the GUT scale and is easily compatible with pre-inflationary Peccei-Quinn (PQ) breaking in a variety of models, including grand unified theories [251] or string compactifications [51, 52], and such low-mass axions may be favored anthropically [252]. Additionally, such light ALPs may explain the previously-observed transparency anomaly of the Universe to TeV gamma-rays [253–256], though in this case the ALP is not required to be DM. Recently, this mass range has gathered significant experimental interest [58, 71, 76, 147, 222, 224, 257] to name a few, or see [56] for a comprehensive review. Furthermore, this mass range is highly complementary to that probed by the ADMX experiment [67, 258, 259], HAYSTAC [260–262], which probe $m_a \sim 10^{-6} - 10^{-5}$ eV. Our result represents the most sensitive laboratory search for ADM below 1 $\mu$eV, is competitive with leading astrophysical constraints from CAST [263], and probes a region where low-mass ALPs which can accommodate all the DM of the universe without overclosure [264–268], as well as particular models of QCD axions with enhanced photon couplings [169, 269]. Aside from the ALP models currently being probed, this result is a crucial first step towards a larger-scale version of ABRACADABRA sensitive to the smaller values of $g_{a\gamma\gamma}$ relevant for the typical QCD axion models in the mass range where axions can probe GUT-scale physics.

## 5.1 Magnet and Cryogenic Setup

ABRACADABRA-10 cm consists of a superconducting persistent toroidal magnet produced by Superconducting Systems Inc. [270] with a minimum inner radius of 3 cm, a maximum outer radius of 6 cm, and a maximum height of 12 cm. The toroidal magnet is counter-wound to cancel azimuthal currents; see [82] for details. We operate the magnet in a persistent field mode with a current of 121 A, producing a maximum field of 1 T at the inner radius. We confirmed this field with a Hall sensor to a precision of $\sim 1\,\%$. Due to the toroidal geometry of the magnet, the field in the center should be close to zero (in the absence of an axion signal).

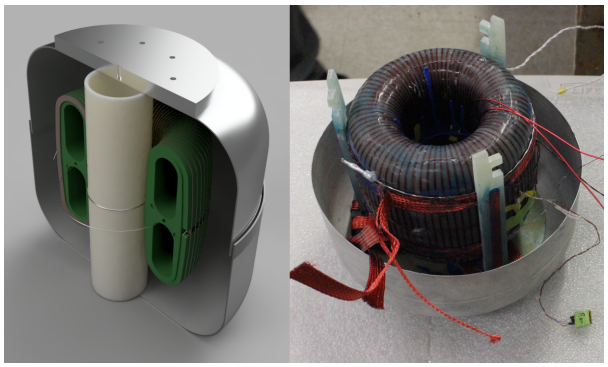


Figure 5.1: *Left:* Rendering of the ABRACADABRA-10 cm setup. The primary magnetic field is driven by 1,280 superconducting windings around a POM support frame (*green*). The axion-induced field is measured by a superconducting pickup loop mounted on a PTFE support (white). A second superconducting loop runs through the volume of the magnet to produce a calibration signal. All of this is mounted inside a superconducting shield. *Right:* Picture of the exposed toroid during assembly.

To reduce AC magnetic field noise, we use both magnetic shielding and vibrational isolation. The toroid is mounted in a G10 support inside a tin-coated copper shell which acts as a magnetic shield below 3.7 K, when the tin coating becomes superconducting. The toroid/shield assembly is thermalized to the coldest stage of an Oxford Instruments Triton 400 dilution refrigerator and cooled to an operating temperature of $\sim 1.2$ K. The weight of the shield and magnet is supported by a Kevlar string which runs $\sim$2 m to a spring attached to the top of the cryostat. This reduces the mechanical coupling and vibration between the detector and cryostat.

We measure AC magnetic flux in the center of the toroid with a solid NbTi superconducting pickup loop of radius 2.0 cm and wire diameter 1 mm. The induced current on this pickup loop is carried away from the magnet through $\sim 50$ cm of 75 $\mu$m solid NbTi twisted pair readout wire up to a Magnicon two-stage SQUID current sensor. The 75 $\mu$m wire is shielded by superconducting lead produced according to [271]. The majority of the 1 mm wire is inside the superconducting shielding of the magnet, but about 15 cm is only shielded by stainless steel mesh sleeve outside the shield.

The two-stage Magnicon SQUID current sensor is optimized for operation at $< 1$ K; we operate it at 870 mK. The input inductance of the SQUID is $L_{\rm in} \approx 150$ nH and the inductance of the pickup loop is $L_p \approx 100$ nH. The SQUID is operated with a flux-lock feedback loop (FLL) to linearize the output, which limits the signal bandwidth to $\approx 6$ MHz. We read out the signal with an AlazarTech ATS9870 8-bit digitizer, covering a voltage range of $\pm 40$ mV. The digitizer is clocked to a Stanford Research Systems FS725 Rb frequency standard. In order to fit the signal into the range of our digitizer, we filter the signal through a 10 kHz high-pass filter and a 1.9 MHz anti-aliasing filter before sending it to the digitizer.

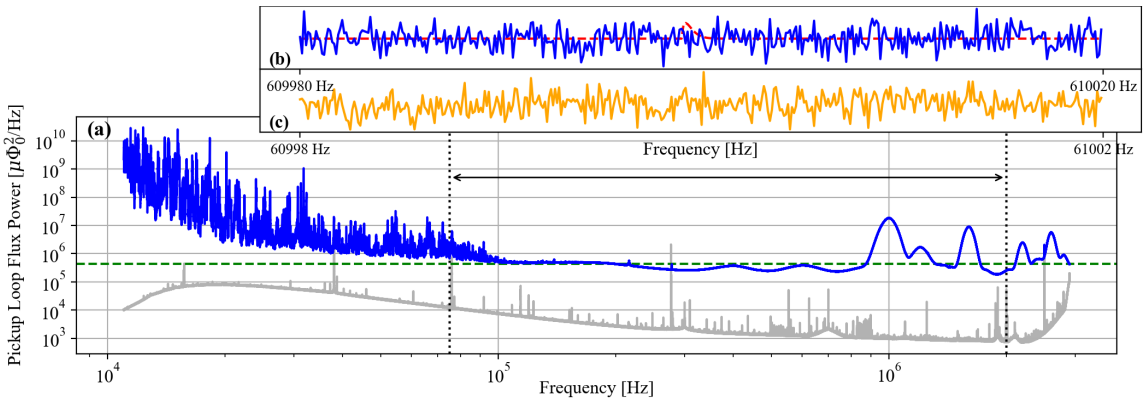



Figure 5.2: Flux spectrum averaged over the the data used in this analysis. (*a*) The spectrum over the frequency range $11\,\mathrm{kHz} < f < 3\,\mathrm{MHz}$, corrected for the pre-digitizer filters (*blue*). For comparison, we also show the digitizer noise floor, corrected for pre-digitizer filters (*gray*) and the characteristic SQUID flux floor (*green dashed*). The axion search range is between the dotted black lines. (*b*) A zoomed view of the 10 MS/s spectrum (*blue*) with $\Delta f = 100\,\mathrm{mHz}$ and and an example axion signal at the 95% upper limit (*red dashed*). (*c*) A zoomed view of the 1 MS/s spectrum with $\Delta f = 10\,\mathrm{mHz}$. Note that the digitizer data was collected at a different time from the SQUID data, and shows a few transient peaks that are not present in the SQUID data.

To calibrate the detector, we run a superconducting wire through the volume of the toroid at a radius of 4.5 cm into which we can inject an AC current to generate a field in the pickup loop, similar to what we expect from an axion signal. The coupling between the calibration and pickup loop can be calculated from geometry to be ≈50 nH. We perform a calibration scan to calculate the end-to-end gain of our readout system. Our calibration measurements indicate that our pickup-loop flux-to-current gain is lower than expected by a factor of $\sim 6$. We determined this to be likely due to parasitic impedances in the circuit, and we will address this issue in future designs.

## 5.2 Data Collection

We collected data from July 16, 2018 to August 14, 2018, for a total integration time of $T_{\mathrm{int}} = 2.45 \times 10^6\,\mathrm{s}$. The data stream was continuously sampled at a sampling frequency of 10 MS/s for the duration of the data-taking period. After completing the magnet-on data run, we collected two weeks of data with the magnet off, but otherwise in the same configuration.

During data taking, the data follow two paths. First, we take the discrete Fourier transform (DFT) of individual sequential 10 s buffers of $10^8$ samples each to produce a series of PSDs. These are accumulated together to produce an average PSD, called $\bar{\mathcal{F}}_{10\mathrm{M}}$, with a Nyquist frequency of 5 MHz and a frequency resolution of $\Delta f = (10\,\mathrm{s})^{-1} = 100\,\mathrm{mHz}$. In the second path, the streamed data are decimated by a factor of 10, to a sampling frequency of 1 MS/s, collected into a 100 s buffer

of $10^8$ samples, then transformed and compiled into a similar running average PSD, $\bar{\mathcal{F}}_{1\mathrm{M}}$, with Nyquist frequency of 500 kHz and $\Delta f = (100\,\mathrm{s})^{-1} = 10\,\mathrm{mHz}$. The 1 MS/s data stream is further decimated in real time to a 100 kS/s stream and written directly to disk. This can be transformed offline to produce $\mathcal{F}_{100\mathrm{k}}$, with a Nyquist frequency of 50 kHz and $\Delta f = 1/T_{\mathrm{int}} \approx 408\,\mathrm{nHz}$. We do not use $\mathcal{F}_{100\mathrm{k}}$ for the present search. All DFT transforms are taken with the FFTW3 library [272].

The $\bar{\mathcal{F}}_{10\mathrm{M}}$ spectra are written to disk and reset after every 80 averages; each stored spectrum thus covers a period of 800 s. This allows us to separate time-dependent noise signals from a constant axion signal. Similarly, the $\bar{\mathcal{F}}_{1\mathrm{M}}$ spectra are written to disk and reset every 16 averages, and cover a period of 1600 s. Figure 5.2 shows the full $\bar{\mathcal{F}}_{10\mathrm{M}}$ spectrum as well as close-ups of the $\bar{\mathcal{F}}_{1\mathrm{M}}$ spectra, converted to pickup loop flux spectral density using the calibration measurements.

Each of the $\bar{\mathcal{F}}_{10\mathrm{M}}$, $\bar{\mathcal{F}}_{1\mathrm{M}}$ and $\mathcal{F}_{100\mathrm{k}}$ spectra have a usable range limited by the Nyquist frequency on the high end, and the frequency resolution required to resolve a potential axion signal on the low end. With our sampling frequency and integration times, we could perform a search over the range from 440 mHz – 5 MHz with enough resolution that a potential signal would span 5 – 50 frequency bins (assuming a typical ADM velocity of $\sim 220$ km/s), though in practice our search range is limited by the signal filters.

We observed large $1/f$-type behavior below ∼20 kHz, with broad noise peaks extending up to ∼100 kHz. This noise is strongly correlated with vibration on the top plate of the cryostat up to the highest frequency measured by our accelerometer, $\sim 10$ kHz [82]. We believe that the tail of this noise continues up to higher frequencies before becoming sub-dominant to the flux noise of the SQUID above 100 kHz. This noise degrades our sensitivity at lower frequencies and we restrict our search range to $75\,\mathrm{kHz} < f < 2\,\mathrm{MHz}$.

For ∼1 week after starting the data collection, we observed very narrow and variable noise peaks in the PSD above $\sim 1.2$ MHz. We are investigating the source of these peaks. After about a week, these peaks died away slowly and did not return until we re-entered the lab to refill an $LN_2$ dewar, then died away again after a few days. The affected time periods were removed and account for a ∼30% decrease in our exposure. We hope that in the future, a more detailed analysis will allow us to recover a significant fraction of this lost exposure.

## 5.3 Data Analysis Approach

Our data analysis procedure closely follows the method introduced in [80]. Our expected signal is a narrow peak in the pickup loop PSD, with a width $\Delta f/f \sim 10^{-6}$ arising from the DM velocity dispersion. When averaged over $N_{\mathrm{avg}}$ independent PSDs, the signal in each frequency bin $k$ ($f_k$)

will follow an Erlang distribution with shape parameter $N_{\rm avg}$ and mean

$$s_k = \begin{cases} A \left.\frac{\pi f(v)}{m_a v}\right|_{v=\sqrt{4\pi f_k/m_a - 2}} & f_k > m_a/2\pi\,, \\ 0 & f_k \leq m_a/2\pi\,, \end{cases} \tag{5.4}$$

where $A$ is defined in Eq. (5.3). We assume $f(v)$ is given by the Standard Halo Model, with velocity dispersion $v_0 = 220\,\text{km/s}/c$, and $v_{\rm obs} = 232\,\text{km/s}/c$ the DM velocity in the Earth frame [273], with $c$ the speed of light. With the DM density and velocity distribution specified, the only free parameter in the predicted signal rate is $g_{a\gamma\gamma}$.

We expect our background noise sources to be normally distributed in the time domain, such that when combined with an axion spectrum, the resulting PSD data is still Erlang-distributed. Accordingly, our combined signal-plus-background model prediction in each frequency bin is an Erlang distribution $P(\bar{\mathcal{F}}_k; N_{\rm avg}, \mu_k)$ with shape parameter $N_{\rm avg}$ and mean $\mu_k = s_k + b$ (see [80] for details). Although the background PSD varies slowly with frequency, the axion signal for a given mass is narrow enough that we restrict to a small frequency range and parameterize the background as a constant $b$ across the window. We verified that the results of our analysis were not sensitive to the size of the window chosen.

We performed our analysis on the $\bar{\mathcal{F}}_{\rm 10M}$ and $\bar{\mathcal{F}}_{\rm 1M}$ spectra over frequency ranges 500 kHz to 2 MHz and 75 kHz to 500 kHz, respectively. We chose the frequency at which we transition from one set of spectra to the other so that the axion signal window is sufficiently resolved everywhere, though we have seen that the exact choice has little effect on the final result. We rebin the $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) spectra in time into 53 (24) spectra that cover 32,000 s (64,000 s) each. This was done to speed up processing time, though it is not necessary for our analysis approach.

We test for an axion signal at mass $m_a$ and coupling strength $A$ by constructing a joint likelihood of Erlang distributions over the 53 (24) $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) given the observed PSD data [80, 82]. For each axion mass, we assign a unique background nuisance parameter to each of the rebinned $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) spectra and profile over the joint likelihood to construct the profile likelihood for $A$ at that mass. This accounts for the possibility that the background level might change on timescales of hours to days.

To detect an axion signal, we place a $5\sigma$ threshold on a discovery test statistic (TS). To evaluate this we first calculate the profile likelihood ratio $\lambda(m_a, A)$, at fixed $m_a$, as the ratio of the background-profiled likelihood function as a function of $A$ to the likelihood function evaluated at the best-fit value $\hat{A}$. From here, we define $\text{TS}(m_a) = -2\log\lambda(m_a, 0)$ for $\hat{A} > 0$ and zero otherwise. This quantifies the level at which we can reject the null hypothesis of $A = 0$. The $5\sigma$

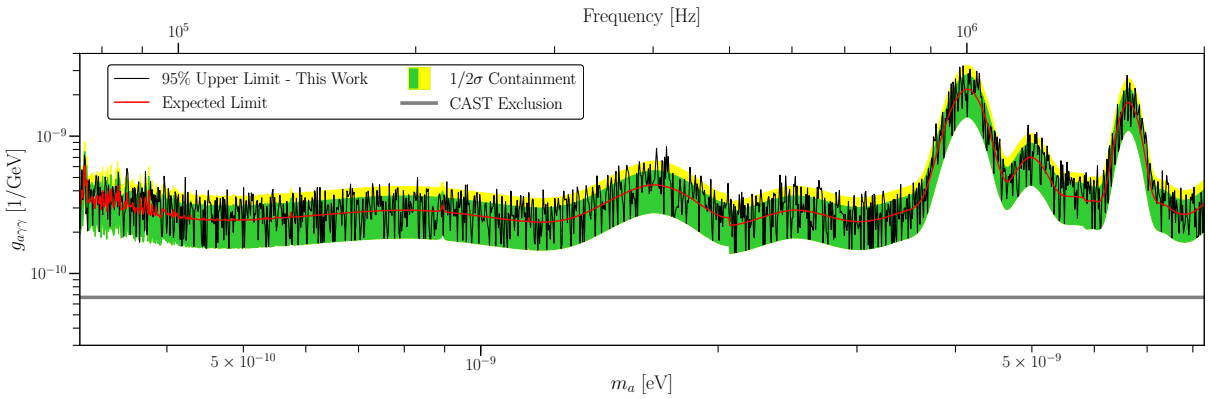



Figure 5.3: The limit on the axion-photon coupling $g_{a\gamma\gamma}$ constructed from ABRACADABRA-10cm data described in this work. We compare the observed limit, which has been down-sampled in the number of mass points by a factor of $10^4$ for clarity of presentation, to the expectation for the power-constrained limit under the null hypothesis. This down-sampling excludes the 87 isolated mass points vetoed in the discovery analysis; further details will be presented in [82]. Additionally, we show the astrophysical constraint on $g_{a\gamma\gamma}$ in this mass range from the CAST helioscope experiment [263]; the region above the grey line is excluded.

condition for discovery at a given $m_a$ is $\mathrm{TS}(m_a) > \mathrm{TS}_{\mathrm{thresh}}$, where [80]

$$\mathrm{TS}_{\mathrm{thresh}} = \left[\Phi^{-1}\left(1 - \frac{2.87 \times 10^{-7}}{N_{m_a}}\right)\right]^2 \tag{5.5}$$

accounts for the local significance as well as the look-elsewhere effect (LEE) for the $N_{m_a}$ independent masses in the analysis (here $\Phi$ is the cumulative distribution function for the normal distribution with zero mean and unit variance). For this analysis, $N_{m_a} \approx 8.1 \times 10^6$ between 75 kHz and 2 MHz, and $\mathrm{TS}_{\mathrm{thresh}} = 56.1$.

Where we have no detection, we set a 95% C.L. limit, $A_{95\%}$, again with the profile likelihood ratio. To do so, we use the statistic $t(m_a, A) = -2\log\lambda(m_a, A)$, with $A > \hat{A}$, by $t(m_a, A_{95\%}) = 2.71$. We implement one-sided power-constrained limits [195], which in practice means that we do not allow ourselves to set a limit stronger than the $1\sigma$ lower level of the expected sensitivity band. We compute the expected sensitivity bands using the null-hypothesis model and following the procedure outlined in [80].

We had to exclude a few specific mass points from our discovery analysis due to narrow background lines that were also observed when the magnet was off. To veto these mass points as potential discoveries, we analyze data collected while the magnet was off (where no axion signal is expected) using the same analysis framework. If in this analysis we find a mass point with LEE-corrected significance greater than $5\sigma$, we exclude that mass point from our axion discovery analysis. In total, this procedure ensures a signal efficiency of $\gtrsim 99.8\%$. Our axion search

yielded 83(0) excesses with significance $\geq 5\sigma$ in the frequency range 500 kHz to 2 MHz (75 kHz to 500 kHz), however all of these points are vetoed by the magnet-off data. We do not exclude these points from our upper-limit analysis, though the observed limits at these isolated points are weaker (they do not appear in Fig. 5.3 because of down-sampling for clarity).

We verified our analysis framework by injecting a simulated software axion signal into our real data and confirmed that the data-quality cuts and analysis framework described above are able to correctly detect or exclude the presence of an axion signal. In the future we hope to build this into a hardware-based option, using the calibration loop to inject "blinded" signals similar to the approach used by ADMX [67]. Further details of the analysis and statistical tests we have performed, as well as an extended discussion of the noise in the excluded exposure, will be further described in a future publication [82].

## 5.4 Results and Discussion

We observe no evidence of an axion signal in the mass range $3.1 \times 10^{-10}$ eV – $8.3 \times 10^{-9}$ eV and place upper limits on the axion-photon coupling $g_{a\gamma\gamma}$ of at least $3.3 \times 10^{-9}\,\text{GeV}^{-1}$ over the full mass range and down to $1.4 \times 10^{-10}\,\text{GeV}^{-1}$ at the strongest point. Our full exclusion limits are shown in Fig. 5.3. This result represents the first search for ADM with $m_a < 1\,\mu$eV, and with one month of data is already competitive with the strongest present astrophysical limits from the much larger CAST helioscope [263] in the range of overlap.

We note that for a significant range in frequency, we achieved the SQUID noise-limit. However, constraints on the detector configuration introduced parasitic impedances into the readout circuit, which lead to a loss in the ultimate axion coupling sensitivity [82]. This will be addressed in future efforts and could yield up to a factor of $\sim$6 improvement in sensitivity with a similar exposure.

As ABRACADABRA-10 cm is a prototype detector, there are many potential directions for future improvement. Our focus in this work has been on demonstrating the feasibility and power of this new approach. Future upgrade paths for the ABRACADABRA program will include improvements to shielding and mechanical vibration isolation, reduction of parasitic inductances, improvements to the readout configuration, expanded frequency range, and construction of a larger toroid.

# CHAPTER 6

# Second Results from ABRACADABRA-10 cm

The axion is a well-motivated candidate to explain the particle nature of DM [45, 46, 86]. This pseudoscalar particle is naturally realized as a pseudo-Goldstone boson of the PQ symmetry, which is broken at a high scale $f_a$; the axion would be exactly massless but for its low-energy interactions with quantum chromodynamics (QCD) [38, 39, 43, 44]. The axion mass is tied to the scale $f_a$ by $m_a \approx 5.7(10^{15}\ \mathrm{GeV}/f_a)$ neV [123]. The range of scales $f_a \approx 10^{15} - 10^{16}$ GeV is particularly compelling because of connections to String Theory [51] and Grand Unification [251, 268], and in the corresponding mass range of $m_a \sim 1 - 10$ neV the axion may naturally explain the observed DM abundance [252, 268]. In this Article we provide the most sensitive probe of ADM in this mass range to date.

ADM that couples to photons modifies Ampère's law such that in current-free regions

$$\vec{\nabla} \times \mathbf{B} = \frac{\partial \mathbf{E}}{\partial t} - g_{a\gamma\gamma} \left( \mathbf{E} \times \vec{\nabla} a - \frac{\partial a}{\partial t} \mathbf{B} \right) , \tag{6.1}$$

with $\mathbf{E}$ and $\mathbf{B}$ the electric and magnetic fields, respectively, $a(\mathbf{x}, t)$ the ADM field, and $g_{a\gamma\gamma}$ the axion-electromagnetic coupling constant. In the presence of a static external magnetic field ADM behaves like an effective current density $\mathbf{J}_{\text{eff}} = g_{a\gamma\gamma}(\partial_t a)\mathbf{B}$. If the axion makes up all of the observed DM then, to leading order in the DM velocity, $\partial_t a = \sqrt{2\rho_{\mathrm{DM}}}\cos(m_a t)$, with $\rho_{\mathrm{DM}} \approx 0.4\,\mathrm{GeV/cm^3}$ the local DM density [28]. It was pointed out in [71, 147] that the effective current induces an oscillating secondary magnetic field which may be detectable in the laboratory without the aid of a resonant cavity for sufficiently small $m_a$. The oscillation frequency is given by $f = m_a/(2\pi)$, with bandwidth $\delta f/f \approx 10^{-6}$ arising from the finite axion velocity dispersion [190]. In this work we leverage this theoretical principle to search for axions in the laboratory.

The most common detection strategy for ADM is through the electromagnetic coupling $g_{a\gamma\gamma}$, which for the QCD axion is directly proportional to the mass $m_a$. Until recently, experiments have focused on searching for axions in the mass range $1 \lesssim m_a \lesssim 40\,\mu\mathrm{eV}$, which is well-suited to microwave cavity searches [67, 68, 110, 258, 260]. In the low-mass regime targeted in this work,

the Compton wavelength of the axion $\lambda_C \sim \mathrm{km}$ is much larger than the experimental apparatus, and so the sensitivity of the experiment improves with volume as $V^{5/6}$, roughly independent of $m_a$ until the size of the experiment approaches $\lambda_C$ [71]. This scaling is important because the expected coupling $g_{a\gamma\gamma}$ is smaller at lower masses, requiring ever-more-sensitive experiments to achieve a detection. ABRACADABRA is an experimental program designed to detect axions at the Grand Unification scale using a strong toroidal magnetic field [71]. ABRACADABRA is part of a suite of ADM experiments which together aim to probe the full QCD axion parameter space [67, 68, 70, 76, 110, 138, 218, 274–276]. The experiment we report on here, ABRACADABRA-10 cm, is a prototype for a larger ADM detector that would be sensitive to the QCD axion. This Article presents data collected in 2020 that is up to an order of magnitude more sensitive than our previous results [72] and places strong limits on ADM in the $0.41 - 8.27$ neV range of axion masses.

## 6.1 ABRACADABRA-10 cm detector

The ABRACADABRA-10 cm detector is built around a 12 cm diameter, 12 cm tall, 1 T toroidal magnet fabricated by Superconducting Systems Inc [270]. The axion interactions with the toroidal magnetic field $\mathbf{B}_0$ drive the effective current, $\mathbf{J}_{\mathrm{eff}}$, which oscillates parallel to $\mathbf{B}_0$ and sources a real oscillating magnetic field through the toroid's center. The oscillating magnetic flux is read out with a two-stage DC-SQUID via a superconducting pickup in the central bore. Unlike other axion detector designs, this novel geometry situates the readout pickup in a nominally field-free region unless axions are present [71]. The detector can be calibrated by injecting fake axion signals (*i.e.*, AC currents) through a wire calibration loop that runs through the body of the magnet. The detector, illustrated schematically in Fig. 6.1, is located on MIT's campus in Cambridge, MA.

In 2019, we performed several detector upgrades from the Run 1 configuration in order to improve our sensitivity [72, 82]. In this Article we report the results of the subsequent data campaign (Run 3), collected after the detector upgrade. Run 3 data consists of ∼430 hours of data collected from June 5 to June 29, 2020.

Before the upgrades were complete, we took additional, uncalibrated data (Run 2), which is not presented here. A subset of that data was instead used to develop our data analysis procedure in order to run a blind analysis on the Run 3 data, as described in detail below.

The total expected axion power, $A$, coupled into our readout pickup is related to the axion-induced flux $\Phi_a$ as

$$A \equiv \langle |\Phi_a|^2 \rangle = g_{a\gamma\gamma}^2 \rho_{\mathrm{DM}} \mathcal{G}^2 V^2 B_{\mathrm{max}}^2, \tag{6.2}$$

where $\mathcal{G}$ is a geometric coupling, $V$ is the magnetic field volume, $B_{\mathrm{max}}$ is the maximum value of $|\mathbf{B}_0|$, and the angle brackets denote the time average [71, 80]. Run 1 utilized a 4.02 cm diameter

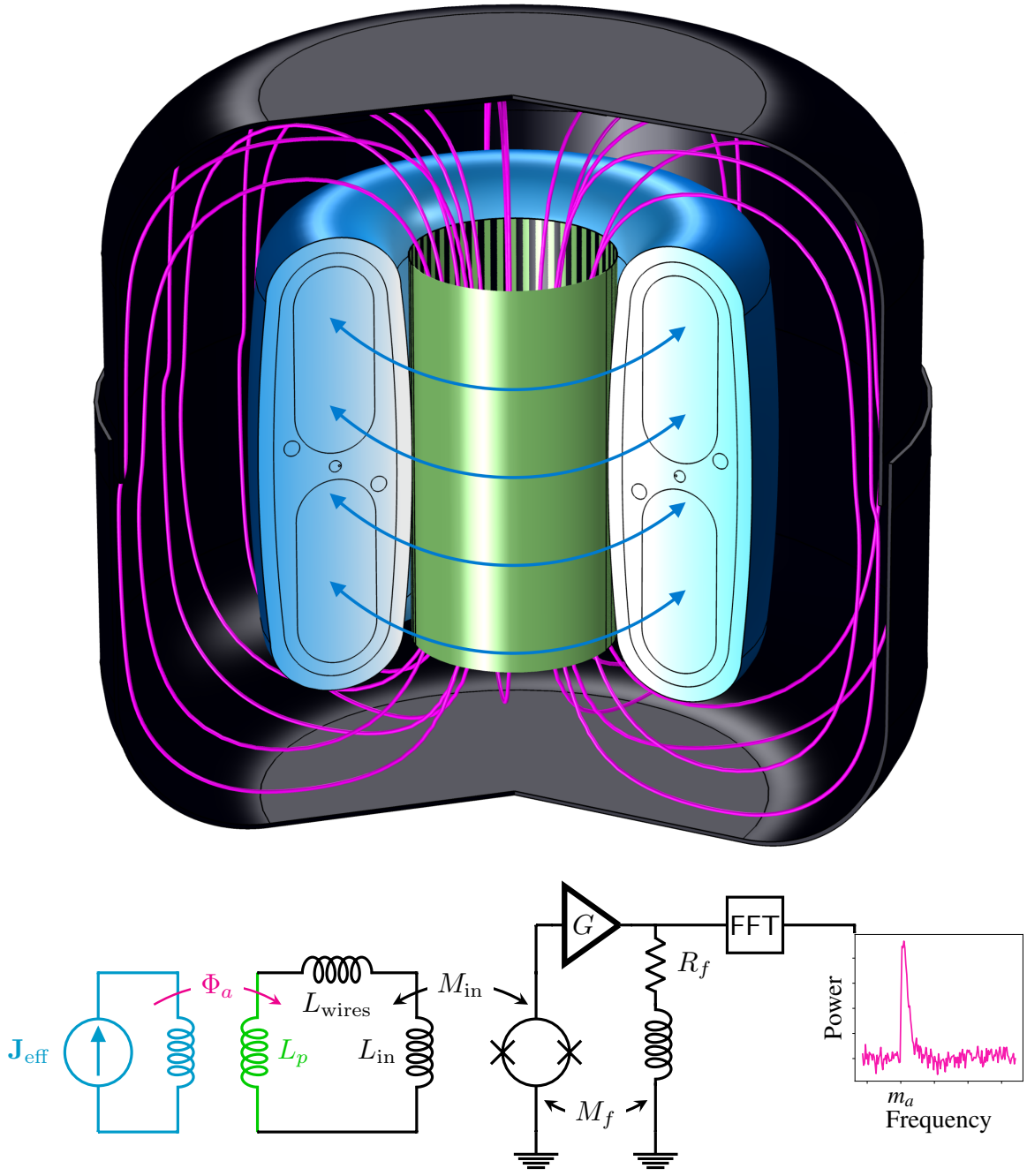



Figure 6.1: *Top*: Schematic of ABRACADABRA-10 cm showing the effective axion-induced current (blue), sourced by the toroidal magnetic field, generating a magnetic flux (magenta) through the pickup cylinder (green) in the toroid bore. *Bottom*: Simplified schematic of the ABRACADABRA-10 cm readout (full circuit diagram in Fig. E.1). The pickup cylinder $L_p$ is inductively coupled to the axion effective current $\mathbf{J}_{\rm eff}$. The power spectrum of the induced current is read out through a DC SQUID inductively coupled to the circuit through $L_{\rm in}$. An axion signal would appear as excess power above the noise floor at a frequency corresponding to the axion mass.

pickup loop made from a 1 mm diameter wire, giving $\mathcal{G} \approx 0.027$. In 2019, we replaced this readout with a 10 cm tall, 5.1 cm diameter superconducting cylinder pickup centered in the toroid bore. This consisted of a 150 $\mu$m-thick Nb sheet wrapped around a PTFE cylinder. This design yields a stronger geometric coupling to $\mathbf{J}_{\text{eff}}$ of $\mathcal{G} \approx 0.031$ and decreases the inductance of the pickup [71]. We compute $\mathcal{G}$ using electromagnetic simulations in the COMSOL Multiphysics package [82, 277].

To amplify our signal, $\Phi_a$ is coupled into the readout SQUID through the pickup circuit (see Fig. 6.1) yielding a transformer gain $M_{\text{in}}/L_T$, where $M_{\text{in}}$ is the input coupling to the SQUID, and $L_T \equiv L_p + L_{\text{in}} + L_{\text{wires}}$ is the total inductance of the pickup circuit, with $L_p$ the pickup cylinder inductance, $L_{\text{in}}$ the input inductance of the SQUID package, and $L_{\text{wires}}$ the parasitic inductance, dominated by the twisted pair wiring. The SQUID, manufactured by Magnicon [278], is read out using Magnicon's XXF-1 SQUID electronics operating in closed feedback loop mode. The Run 1 sensitivity was limited by parasitic inductance in the NbTi wiring of this circuit that placed a lower limit on $L_T \gtrsim 1.6\,\mu$H. During the upgrade, we replaced this wiring, moving the SQUIDs closer to the detector to reduce the wire length. Based on calibration data, we found that the total impedance in the circuit is $\sim 800$ nH. Finally, the SQUID was operated at a higher flux-to-voltage gain setting of 4.3 V/$\Phi_0$ in Run 3, compared to the previous Run 1 which we ran at 1.29 V/$\Phi_0$ due to higher levels of environmental noise. This change does not directly improve the signal gain, but does reduce system noise. We also improved our noise floor by reducing the operating temperature of the SQUID package from ∼870 mK to ∼450 mK. All together, the upgrade campaign increased the expected power coupled into our readout and reduced the total system noise.

The improved sensitivity of the upgraded readout circuit also amplified the low-frequency vibrational backgrounds seen in Run 1, which caused the SQUID amplifier to rail when the magnet was on. In order to correct this, we implemented an active feedback stabilization (AFS) system to reduce the low-frequency noise, which is discussed further in the appendices.

As in Run 1, the magnet and pickup were placed inside a superconducting tin-copper shield and hung from a passive vibration isolation system, consisting of a string pendulum and spring, within an Oxford Instruments Triton 400 dilution refrigerator [82]. The magnet and pickup were operated at $\lesssim$1 K and the SQUIDs were at ∼400 mK, which kept the readout circuit superconducting over the course of the run and kept thermal noise subdominant to SQUID flux noise. Following the procedure of Run 1, the output of the SQUID was run into an 8-bit AlazarTech AT9870 digitizer via a 70 kHz-5 MHz bandpass filter. The digitizer was locked to a Stanford Research Systems FS725 Rubidium frequency standard in order to maintain clock accuracy over the coherence time of the axion signal, $\sim 1$ s for signals at 1 MHz, throughout the data and calibration runs.

We performed *in situ* magnet-on and magnet-off calibrations in the data-taking configuration by attaching a harmonic signal generator to the calibration circuit and scanning across frequencies and

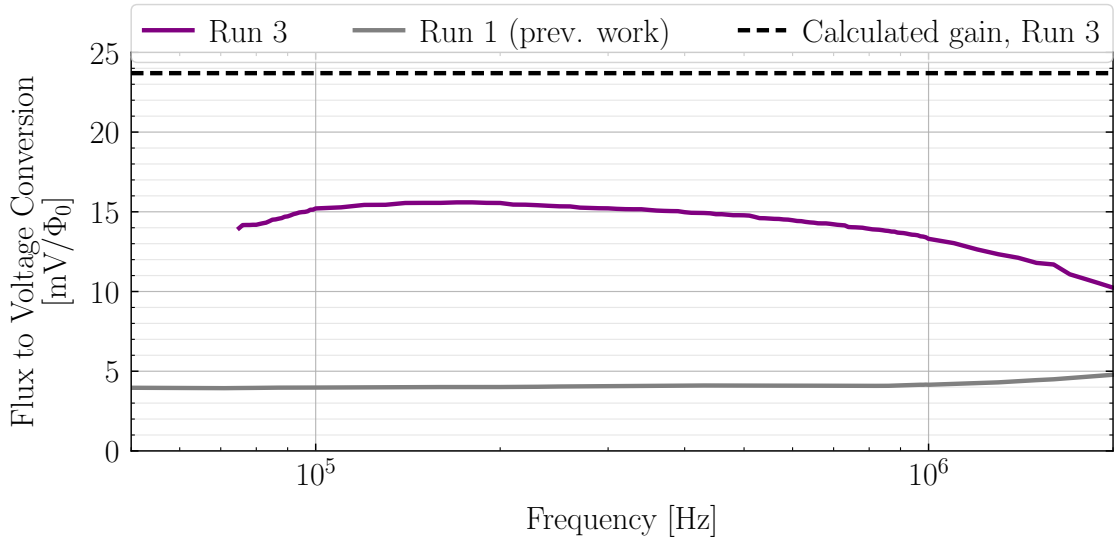



Figure 6.2: The gain shown here is defined as the change in amplifier output voltage over a corresponding change in input flux amplitude on the pickup cylinder ($\partial V_{\text{out}}/\partial \Phi_a$). Both transfer functions roll off at high frequencies due to the amplifier bandwidth, which we estimate to have a cutoff frequency of approximately 1 MHz. We believe the difference in calculated and measured gain is due to inconsistency in the total inductance of the pickup circuit.

amplitudes. The calibration signal was attenuated and fed into the calibration loop, mimicking the axion effective current signal $\mathbf{J}_{\text{eff}}$ up to geometric factors. The geometry is modeled in COMSOL Multiphysics [277], from which we extract the coupling between both the calibration loop and axion effective current signal to the pickup cylinder. By combining the results of the calibration scans and geometric modeling, we can determine the effective gain, $\partial V_{\text{out}}/\partial \Phi_a$, of the SQUID amplifier output voltage as a function of flux on the pickup cylinder (see Fig. 6.2). This procedure is analogous to that used in Run 1 [82].

The gain measured by the calibrations for Run 3 differs from the calculated gain by a factor of $\sim$1.8. By individually calibrating various components of the end-to-end circuit, we determined that this is likely due to a misestimation of the calculated total inductance of the pickup circuit. The calibrated SQUID noise floors, which set the lower limits of our sensitivity, are shown in Fig. E.2

## 6.2 Data Collection

The axion search data was collected using an identical procedure as in Run 1 [82]. The SQUID amplifier output voltage was sampled at a frequency of 10 MS/s, with a $\pm 40$ mV voltage window. The data were stored as a series of PSDs, which were computed on-the-fly: $\bar{\mathcal{F}}_{\text{10M}}$ with a Nyquist frequency of 5 MHz and frequency resolution of $\Delta f = 100$ mHz, $\bar{\mathcal{F}}_{\text{1M}}$ with a Nyquist frequency

of 500 kHz and frequency resolution of $\Delta f = 10\,\text{mHz}$, and a continuous data stream sampled at 100 kHz that can be analyzed offline. $\bar{\mathcal{F}}_{10\text{M}}$ ($\bar{\mathcal{F}}_{1\text{M}}$) is averaged over 800 s (1600 s) before being written to disk. In this work, we used the $\bar{\mathcal{F}}_{10\text{M}}$ to search the frequency range from $500\,\text{kHz} - 2\,\text{MHz}$, and the $\bar{\mathcal{F}}_{1\text{M}}$ spectra to search from $50 - 500\,\text{kHz}$.

## 6.3 Data Analysis and Results

An axion signal is expected to manifest as a narrow peak in the PSD data, as illustrated in Fig. 6.1. The width and overall shape of the signal are set by the local DM velocity distribution, which we take to be the Standard Halo Model with a velocity dispersion of $v_0 = 220\,\text{km/s}$ and a boost from the halo to the solar rest frame of $v_\odot = 232\,\text{km/s}$ [185]. With the speed distribution and local DM density fixed, the two free signal parameters are the axion mass, $m_a$, which determines the minimum frequency of the signal, and the coupling $g_{a\gamma\gamma}$, which determines its amplitude through Eq. (6.2). Our analysis procedure closely follows the approach used in the Run 1 search [72, 82] based on [80], which constrains the allowable values of $g_{a\gamma\gamma}$ at each possible value of $m_a$.

The search is performed with a frequentist log-likelihood ratio TS; exact expressions are provided in the SI (see also [82]). Our broadband search procedure probes $\sim 11.1$ million mass points between $0.41 - 8.27\,\text{neV}$ ($100\,\text{kHz} - 2\,\text{MHz}$) in Run 3. As we expect only one axion signal in our search (or at most a small number), the majority of the TS values are probing the distribution of the null hypothesis. Assuming only Gaussian noise, we expect this null distribution to be a one-sided $\chi^2$-distribution [80], which was indeed the case in Run 1 [72, 82]. However, the increased sensitivity from the detector upgrades introduced non-Gaussian noise sources that required us to modify our Run 1 analysis procedure. We developed and validated our new procedure on a randomly-selected sample of 10% of Run 2's $\sim$ 13.7 million mass points, after which we unblinded the Run 3 data with the procedure fixed.

In Run 1, we searched for an axion signal as a feature appearing above a flat white noise background. For each $m_a$, the search was performed in a narrow window around that mass with the background level allowed to vary independently in each window. For the Run 2 and Run 3 analyses we allow the mean background level of the noise to vary linearly with frequency uniquely in each sliding window. We use sliding windows of relative width $\delta f/f \approx 5.5 \times 10^{-6}$, starting at $f = (1 - 10^{-6}) \times m_a/(2\pi)$.

As in Run 1, we use the magnet-off data to veto frequency ranges that also display statistically significant TS values when $|\mathbf{B}_0| = 0$ and thus the axion power should vanish. However, we observed narrow single-bin 'spikes' that appear to drift in frequency on the timescale of our data collection (see E.6 for an example). If interpreted in isolation, these spikes sometimes correspond to statistically-significant excesses. Nevertheless, they are inconsistent with axion signals and are

most likely due to unknown environmental noise sources near the detector, persisting throughout Runs 2 and 3; indeed, many of the peaks are distributed at multiples of 50 Hz. To remove these artifacts, we leverage the fact that the PSDs are saved periodically to disk yielding a time evolution of the environmental backgrounds; we veto single-bin spikes that move in frequency. We place a 1.0 Hz veto window around these single-bin spikes. These cuts remove 3.8% of the axion mass points from our search in the Run 3 data. The magnet-off vetoing procedure removes an additional $0.07\%$ of mass points.

After implementing the vetoes, we found the distribution of TS values in the 10% Run 2 validation sample deviated from the expected $\chi^2$ distribution; for example, there were 27 mass points with $\mathrm{TS} > 25$ whereas from the $\chi^2$ distribution we would have expected less than one. To account for the deviation in the TS distribution from the $\chi^2$ distribution in a data-driven fashion, we follow the prescription developed and implemented in [279–281] for searches for DM-induced lines in astrophysical gamma-ray data sets. At each mass point, we introduce and profile over a systematic nuisance parameter that would be degenerate with the signal parameter but for a prior that is determined by forcing the TS distribution to follow the $\chi^2$ distribution. Specifically, we force the TS distribution to match the null hypothesis distribution at $4\sigma$ local significance. This is described further in the SI.

After the nuisance parameter and vetoing procedures, we construct a likelihood as a function of $g_{a\gamma\gamma}$ at each mass point. The final distribution of TS values computed from the likelihoods is shown in Fig. 6.3; no TS values were found in excess of the $5\sigma$ look-elsewhere effect-corrected discovery threshold. In the calibration of our analysis procedure, we found one signal candidate in the Run 2 data at over $5\sigma$ global statistical significance (see Fig. E.6, where a corresponding feature can be seen in the magnet-off data), but that mass point is not significant in the Run 3 analysis.

In the absence of an excess consistent with an ADM origin, we can determine 95% one-sided upper limits on $g_{a\gamma\gamma}$ as a function of the mass, $m_a$. The average 95% upper limits from the Run 3 analysis along with their $1\sigma$ and $2\sigma$ expectations under the null hypothesis are indicated in Fig. 6.4. In that figure we compare our upper limits to those found from the ADM experiment SHAFT [276] along with results from the solar axion experiment CAST [263] and astrophysical $X$-ray searches (SSC) [84], both of which do not require the axion to comprise the DM. The fraction of vetoed mass points is illustrated in a sliding window inFig. E.4, which also shows the distribution of data fractions included in the analyses. In Fig. E.5 we illustrate the magnitude of the systematic nuisance parameter $g_{a\gamma\gamma}^{\mathrm{nuis.}}$, while in Fig. E.7 we show what the limit would be without the nuisance parameter tuning. Fig. E.8 shows that the 95% upper limit and discovery TS behave as expected when synthetic axion signals are injected into the real data.

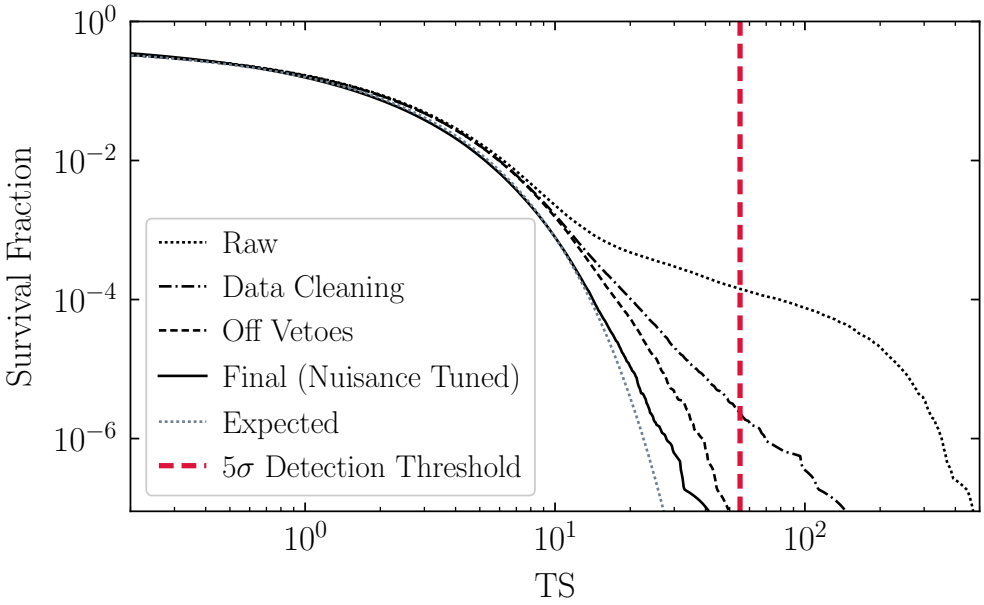



Figure 6.3: The survival function of TS values from the likelihood analysis of the Run 3 results. The $y$-axis indicates the fraction of mass points tested with a discovery TS at or above the value on the $x$-axis. Under the null hypothesis, the distribution should follow the survival function of the one-sided $\chi^2$ distribution with one degree of freedom ("Expected," dotted gray). This is indeed the case after data cleaning for *e.g.* single-channel excesses in time slices, magnet-off vetoes, and the inclusion of a systematic nuisance parameter, which is tuned in a sliding window at $4\sigma$ local significance to give the correct number of excesses at or above that significance, masking the signal of interest. No excesses are found beyond our indicated $5\sigma$ LEE-corrected discovery threshold.

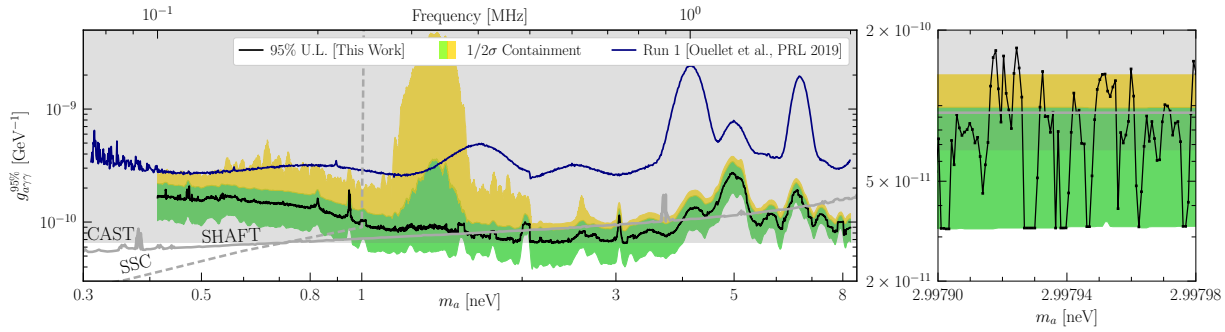



Figure 6.4: (*Left*) The one-sided 95% upper limit (U.L.) on $g_{a\gamma\gamma}$ from this work excludes previously unexplored regions of ADM parameter space. The $1\sigma$ and $2\sigma$ containment regions are constructed by taking the appropriate percentiles of the distributions of the limits over narrow mass ranges; note that this means that $\sim$16% of the upper limits lie at the bottom of the green band. Around 11.1 million mass points are analyzed, so the plotted data is smoothed for clarity. Our limits surpass those from a number of indicated astrophysical and laboratory searches in this mass range, including CAST (solid grey region), super star cluster constraints (dashed grey line), and SHAFT (solid grey line). See text for details. (*Right*) The un-smoothed limit in a narrow mass range between 2.99790 and 2.99798 neV. This provides a detailed view of variations in the limit at each axion mass that arise from statistical fluctuations across the collected data that are not visible in the smoothed data shown in the left plot. This range also depicts the location where our maximum sensitivity is achieved, with our strongest limit at $g_{a\gamma\gamma} \lesssim 3.2 \times 10^{-11}$ GeV$^{-1}$.

## 6.4 Conclusion

In this work we present the results from ABRACADABRA-10 cm's second physics campaign, searching for ADM in the mass range 0.41-8.27 neV. We find no evidence for ADM and constrain the axion-photon coupling down to the world-leading level $g_{a\gamma\gamma} \lesssim 3.2 \times 10^{-11}$ GeV$^{-1}$ at 95% confidence. Our work motivates key elements of the design of future larger-scale experiments. These include the mitigation of stray fields from the magnet and vibrations induced by a modern pulse-tube-based cryogenic system, which limits our current low-frequency reach. The ABRACADABRA-10 cm results presented in this Article demonstrate the power of mature simulations for optimizing the design of the detector and for modeling the calibration response. An advanced and novel analysis framework was used to identify noise sources and account for systematic uncertainties in a data-driven fashion.

Our work identifies three areas that can be addressed in the next physics campaign: (i) moderate improvements (up to a factor $\sim$0.4 in $g_{a\gamma\gamma}$) could be achieved by further reducing the wire and SQUID inductances, (ii) better shielding from environmental noise could increase the sensitivity to $g_{a\gamma\gamma}$ by an order of magnitude at low frequencies, so long as (iii) the fringe fields are reduced or better vibrationally isolated (see Fig. E.2). To significantly increase the sensitivity of the experiment, larger magnets with higher fields are needed since the sensitivity to $g_{a\gamma\gamma}$ scales

with the detector volume $V$ and field $B_0$ as $g_{a\gamma\gamma}^{-1} \sim B_0 V^{5/6}$ [71]. The addition of a resonant readout circuit could enhance the reach in $g_{a\gamma\gamma}$ by an additional $\sim$2 orders of magnitude depending on the scanning strategy, with a high frequency readout permitting sensitivity to masses up to 800 neV [71, 221]. ABRACADABRA is merging with the DMRadio program to realize a series of experiments that chart a path toward discovering the QCD axion in the parameter space corresponding to new physics at the Grand Unification scale [282–286].

# CHAPTER 7

# Radio Searches for Axion Dark Matter

Recently, it was proposed that radio telescope observations of neutron stars (NSs) can probe axion dark matter (DM) [101–104, 287]. In the magnetosphere surrounding a NS, axion DM may resonantly convert into radio-frequency photons at locations where the plasma frequency of the magnetosphere equals the axion mass, with conversion probabilities determined in part by the strength of the magnetic field surrounding the NS. The central frequency of the hypothetical radio signal from an individual NS is set by the mass of the axion, red-shifted by the line-of-sight velocity of the NS. The predicted axion-induced radio signal would appear as a nearly monochromatic peak in the otherwise smoothly-varying radio spectrum from the NS and its nearby environment. The frequency of this peak is universal for all sources and is determined by the currently unknown mass of the axion particle.

In [103, 104, 287, 288] it was shown that high-frequency-resolution observations with radio telescopes such as the Robert C. Byrd Green Bank Telescope (GBT) and the Effelsberg 100-m Telescope towards nearby isolated NSs (INSs) and towards regions of high NS and DM density, such as the Galactic Center (GC) of the Milky Way, would be sensitive to vast regions of previously unexplored axion parameter space. In this work, we perform such searches with the GBT and the Effelsberg radio telescope.

The quantum chromodynamics (QCD) axion is a well-motivated DM candidate because in addition to explaining the observed abundance of DM [45, 46, 86] it may also resolve the strong *CP* problem of the neutron electron dipole moment [38,39,43,44] (see [56] for a detailed review). The QCD axion may make up the observed abundance of DM over a wide range of masses [289], but a natural mass range is 5–25 $\mu$eV. In this work we target axion masses in the range $m_a \in (4.5, 10.5)$ $\mu$eV, corresponding to radio frequencies $f = m_a/(2\pi) \in (1.1, 2.7)$ GHz.

The conversion of axion DM to radio photons arises from the Lagrangian $\mathcal{L} = g_{a\gamma\gamma}\, a\, \mathbf{E}\cdot\mathbf{B}$, where $\mathbf{E}$ ($\mathbf{B}$) are electric (magnetic) fields, $a$ is the axion field, and $g_{a\gamma\gamma}$ is a coupling constant with units of inverse energy. For the QCD axion, $g_{a\gamma\gamma}$ is proportional to $m_a$, but models of more general axion-like particles can have $g_{a\gamma\gamma}$ and $m_a$ as independent parameters. The mass range that we

target here with radio telescope searches is also the subject of significant longstanding laboratory search efforts for the coupling $g_{a\gamma\gamma}$. The Rochester-Brookhaven-Fermilab (RBF) [261, 290] and University of Florida (UF) [260] axion haloscope experiments set competitive constraints on axion DM in the mass range covered by this analysis, though our results exclude new parameter space beyond what was probed by those experiments. More recently, the ADMX experiment has reached sensitivity to the QCD axion at ∼2–3.5 $\mu$eV [67, 131, 291], and the HAYSTAC experiment has set strong constraints on axion DM in the mass range $m_a \sim 23 - 24$ $\mu$eV [110].

## 7.1 Parametrics of Indirect Detection of Axion Conversion

As this Chapter represents a departure from our previous focus on axion direct detection, we begin with a brief discussion of why astrophysical observations of NSs represent promising opportunities for axion discovery. As discussed at various points in this thesis, axion direct detection attempts to provide a background electromagnetic field with which axions may interact and produce observable signals. In the laboratory, one typically seeks to provide as strong as possible a magnetic field. NSs realize magnetic fields much larger than can be generated in the laboratory, and, as a result, can generate bright signals in the presence of axions, even accounting for their large distance from Earth.

NSs generate axion-induced signals by the conversion of ambient DM axions to photons in their magnetospheres, enabling a direct comparison with the CAST experiment, which seeks to convert relativistic axions produced by the sun to detectable photons [263]. The probability of axion-to-photon conversion in a background magnetic field is given by

$$p_{a\gamma} \approx g_{a\gamma\gamma}^2 B^2 L^2 \tag{7.1}$$

where $g_{a\gamma\gamma}$ is the axion photon coupling, $B$ is the strength of the background magnetic field, and $L$ is the length over which the conversion process occurs. For CAST, $B = 9$ T and $L = 9$ m whereas for NS conversion, we can have $B$ as large as $10^{11}$ T and $L \approx 100$ m. As is then clear, the conversion probability for axions at NSs is considerably larger.

However, conversion of axions at NSs is somewhat more limited in terms of mass-range than CAST. Conversion of axions at NSs only occurs efficiently if the mass of the photon matches the mass of the axion. While this is never true in vacuum, the nonzero charge density in NS magnetospheres provides the photon a plasma mass, which varies between 100 neV and 50 $\mu$eV at distances where the NS magnetic field is large. In fact, the $L \approx 100$ m length for conversion at NSs is determined by the typical length over which the mass-matching is sufficiently good. Given this mass range of 100 neV to 50 $\mu$eV, conservation of energy tells us that converted axions will become

radio photons, motivating searches for axion conversion with radio telescopes as is considered in this Chapter.

Axion conversion will produce a monochromatic excess of flux density as observed by a radio telescope at a frequency determined by the axion mass. For an axion that produces a signal at a frequency of 1 GHz with a coupling of $g_{a\gamma\gamma} = 10^{-11}$ GeV$^{-1}$ by converting at NS a distance of 100 pc from Earth with typical period and magnetic field strength, we would expect to see a total flux density excess of $6.7 \times 10^{-3}$ Jy with a width of $5$ kHz. With three hours of observing time, single-dish radio telescopes can achieve statistical sensitivity to such a signal, enabling us to constrain $g_{a\gamma\gamma} \lesssim 10^{-11}$ GeV. Limits along these lines, which exceed CAST in sensitivity, are realized in this work through observations of two isolated NSs. Even greater sensitivity is achieved through observations of NS populations containing stars with larger magnetic fields in the presence of enhanced DM densities.

## 7.2 Data Acquisition

We collected data in the L-band ($1.15 - 1.73$ GHz) with the GBT and in the L-Band ($1.27 - 1.45$ GHz) and S-band ($2.4 - 2.7$ GHz) using the Effelsberg radio telescope to search for axion DM signatures from a variety of different sources. We describe the data taking procedures from the two telescopes in turn.

### 7.2.1 GBT Observations

The GBT observations were performed with the VErsatile GBT Astronomical Spectrometer (VEGAS) backend [292] on March 10 and 29, 2019 with a notch filter applied from 1.2 to 1.34 GHz, so these frequencies are not included in our analysis (project AGBT19A_362, PI: Safdi). The nearby INS targets observed by the GBT are summarized in Tab. 7.1. Note that we also observed the GC, M31, and M54 with the GBT, but the resulting axion limits are less robust than those from the INSs and from the Effelsberg GC observations and so are presented in the appendices. (The GBT GC observations lead to weaker limits than the Effelsberg GC observations because the GBT observations were taken with lower frequency resolution.) All observations used the "Spectral Line" observing type and with one beam covering an area on the sky $\sim \pi(\mathrm{FWHM}/2)^2$, where FWHM is the full width at half maximum of the telescope response, which is $8.4'$ at 1.5 GHz for the GBT.

The INS observations used five VEGAS spectrometers in mode 9 across the L-band, leading to the frequency resolution $\delta f_{\rm obs}$ reported in Tab. 7.1. For our fiducial analyses the data is further down-binned to resolutions $\delta f_{\rm fid}$ given in Tab. 7.1. Data were collected in both polarizations, though in this analysis we only analyze the polarization-averaged flux. (See [103] for possible

polarization signatures.) The observations performed position switching so that for a given observational target, half the data collection time was on-source ("ON") and half was spent observing blank-sky locations at similar elevations ("OFF") in order to establish a reference baseline for the analysis. The ON exposure times $t_{\rm exp}$ are listed in Tab. 7.1. The OFF locations were chosen to be $1.25^\circ$ away from the target of interest. The position-switching was carried out at five-minute intervals for each of the targets, leading to four separate observations of ON and OFF positions.

Over the observing period, data were saved in independent short exposures for ON and OFF observations of RX J0720.4−3125 and RX J0806.4−4123. In each successive exposure, a calibration noise diode was alternated between on and off with a switching period of 0.2097 seconds. The timing resolution allows for the identification of transient effects and data filtering, which is discussed further below and in the appendices. The calibration source 3C48 was observed for approximately two minutes to flux-calibrate the INS observations. Additionally, we observed the star-forming region W3(OH) for approximately five minutes to verify that our analysis framework is able to successfully identify the OH maser lines.

### 7.2.2 Effelsberg Observations

We also carried out L-Band and S-Band observations with the Effelsberg 100-m radio telescope towards the GC (project 77-17, 64-18, PI: Desvignes). The observations were taken with the PSRIX backend [293] − performing baseband sampling − in mid-June 2018 and early-February 2019 using the prime (secondary) focus receiver P217mm (S110mm) for the L- and S-band, respectively. In both cases we recorded orthogonal polarizations, which were later averaged offline for further analysis. Note that the FWHM of the Effelsberg beam is $9.78'$ ($4.58'$) at 1.408 GHz (2.64 GHz). Observations were carried out towards the magnetar SGR J1745−2900, which is $\sim 2.4''$ away from the GC, and the planetary nebula $\rm NGC\,7027$ for subsequent use in the flux calibration procedure. For the measurements towards the GC we used a position switching mode, with ON-source integration times of 61.9 min and 40.0 min for S-band and L-band, respectively, and respective OFF-source integration times of 22.8 min and 37.0 min (see Tab. 7.1). The ON observation was performed first, followed by a single OFF observation taken $16.4^\circ$ away from the GC.

## 7.3 Analysis

We reduced and calibrated the GBT data following a modified implementation of the `GBTIDL` data reduction pipeline [294], extended to include a time-series data filtering performed independently at each channel and a channel-dependent system temperature calibration. The full procedure results in measurements of flux densities $\{d_i\}$ at frequencies $\{f_i\}$, with $i$ labeling the frequency channel.

| Target | $t_{\rm exp}$ [min] | $\delta f_{\rm obs}$ [kHz] | $\delta f_{\rm fid}$ [kHz] | type |
|---|---|---|---|---|
| RX J0806.4−4123 | 20.0 | 0.8 | 8.4 | INS |
| RX J0720.4−3125 | 20.0 | 0.8 | 8.4 | INS |
| GC (Eff., S-Band) | 61.9 | 3.81 | 11.44 | pop. |
| GC (Eff., L-Band) | 40.0 | 2.44 | 7.32 | pop. |

Table 7.1: The targets observed by the GBT and Effelsberg for evidence of axion DM. "Pop." refers to populations of NSs, while "INS" refers to a single isolated NS. The bin widths $\delta f_{\rm obs}$ correspond to those of the original observation, but we down-bin the data before performing the axion line search to the resolution given by $\delta f_{\rm fid}$ to account for the finite width of the signal. The INS (GC) observations were performed with the GBT (Effelsberg radio telescope). The GBT INS observations cover the frequency range 1.15 to 1.73 GHz, with a gap from 1.2 to 1.35 GHz, and the L-band (S-band) Effelsberg observation covers 1.27 to 1.45 GHz (2.4 to 2.7 GHz). Note that the $t_{\rm exp}$ are the ON exposure times.

Because the stacked, calibrated data has been constructed by averaging many ($> 10^3$) independent antenna measurements together, the $\{d_i\}$ are approximately normally distributed.

For Effelsberg, high-resolution frequency spectra (131072 spectral channels) were generated from the raw 'baseband' data using the `DSPSR`[1] software tools [295]. We used the full-integrated spectra in our analysis, with a calibration procedure described in the appendices. Before analyzing the data we first down-bin in frequency space to bins of width $\sim$8 kHz (see Tab. 7.1) to account for the finite width of the signal, such that the majority of the signal should appear in a single frequency bin. As discussed further in the appendices and first suggested in [287], reflection and refraction of the outgoing electromagnetic waves in the rotating plasma induces a frequency broadening at the level $\delta f/f \sim 5\times 10^{-6}$ or less from the INSs. We note that even though the Effelsberg observations are searching for emission from a population of NSs, the data are at sufficiently high frequency resolution that we may search simply for the brightest converting NS from that population.

To inspect the data for excess flux at frequency channel $i$, we construct the likelihood

$$\mathcal{L}_i(\vec{d}|A,\mathbf{a}) = \prod_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[-\frac{(d_k - \mu(f_k|\mathbf{a}) - A\delta_{ik})^2}{2\sigma_k^2}\right], \tag{7.2}$$

where $A$ is the excess flux density in the central frequency channel. Note that the index $k$ labels the analysis-level frequency channel, and the product runs over the frequency bins included in the analysis window. We model the background in the narrow sliding frequency window with a frequency-dependent mean flux density $\mu(f|\mathbf{a})$ and a single variance parameter $\sigma^2$, such that the variance in each frequency channel is given by $\sigma_i^2 = \sigma^2/\alpha_i$ for an acceptance fraction $\alpha_i$ of data at

[1] http://dspsr.sourceforge.net

frequency channel $i$ after the data filtering. Note that $\alpha_i = 1$ for all Effelsberg frequency channels as we do not apply the time-filtering procedure to that data. The nuisance parameter vector $\mathbf{a}$ characterizes the frequency dependence of the mean; in practice we take $\mu$ to be a quadratic function of $f$ so that $\mathbf{a}$ has three independent parameters, though our final results are not sensitive to this choice (see the appendices).

In our fiducial analysis we include within the sliding analysis window the $10$ frequency bins to the left and to the right of the central frequency channel, excluding the two bins on either side of the signal bin in case of signal leakage into those bins, if *e.g.* the axion mass does not line up with the bin center. Note that to account for this possibility we also perform the analyses with all frequency bins shifted by approximately half a bin spacing. The variance parameter $\sigma^2$ is fixed by fitting the background-only model to the frequency sidebands with the central frequency channel masked out. We construct the profile likelihood $\mathcal{L}_i(\vec{d}|A)$ by maximizing $\mathcal{L}_i(\vec{d}|A, \mathbf{a})$ over the nuisance parameters $\mathbf{a}$ at each fixed value of $A$, and we use the profile likelihood to construct the one-sided 95% upper limit on the flux density as shown in Fig. 7.1 (see, *e.g.*, [172]). In particular, we consider positive and negative values of $A$ and we take the 95% upper limit to be the value of $A > \hat{A}$ such that $2[\ln \mathcal{L}_i(\vec{d}|A) - \ln \mathcal{L}_i(\vec{d}|\hat{A})] \approx -2.71$, where $\hat{A}$ is the signal parameter that maximizes the profile likelihood. We then further power-constrain our limits to avoid setting limits that are stronger than expected due to downward statistical fluctuations [195]. We accomplish this by recording the actual limit as the maximum of the $16^{\text{th}}$ percentile of the distribution of expected limits under the null hypothesis, as computed using the Asimov procedure [172], and the limit observed on the actual data. Our test-statistic (TS) for comparing signal and null hypotheses for evidence of an axion is the log-likelihood ratio $\text{TS}_i \equiv 2 \times [\ln \mathcal{L}_i(\vec{d}|\hat{A}) - \ln \mathcal{L}_i(\vec{d}|0)]$, for $\hat{A} > 0$, and $\text{TS}_i = 0$ if $\hat{A} < 0$.

We additionally analyze the stacked but uncalibrated OFF spectra. This is valuable because the OFF data are subtracted and divided from the ON data to remove the instrumental baselines, but this may cause features in the OFF spectra to be imprinted on the calibrated flux densities. Therefore, statistically significant excesses that appear in both the calibrated source flux density spectra and the OFF system temperatures can be vetoed as they are inconsistent with, or at least do not require, an axion interpretation. In our analysis, we veto any excess in the calibrated ON data which appears with a $97.5^{\text{th}}$ percentile discovery TS in the OFF data. Note that we determine the TS percentiles by using the full distribution of observed TSs.

The 95% upper limits on the flux densities, defined relative to the single-channel frequency bin widths $\delta f_{\text{fid}}$ given in Tab. 7.1, are shown in Fig. 7.1. We compare the upper limits to the expected limits from the ideal radiometer equation, which assumes that all of the noise is thermal at the system temperature. The true limits are slightly weaker likely because of sources of systematic uncertainty, such as uncertainties in the background model and instrumental uncertainties not fully

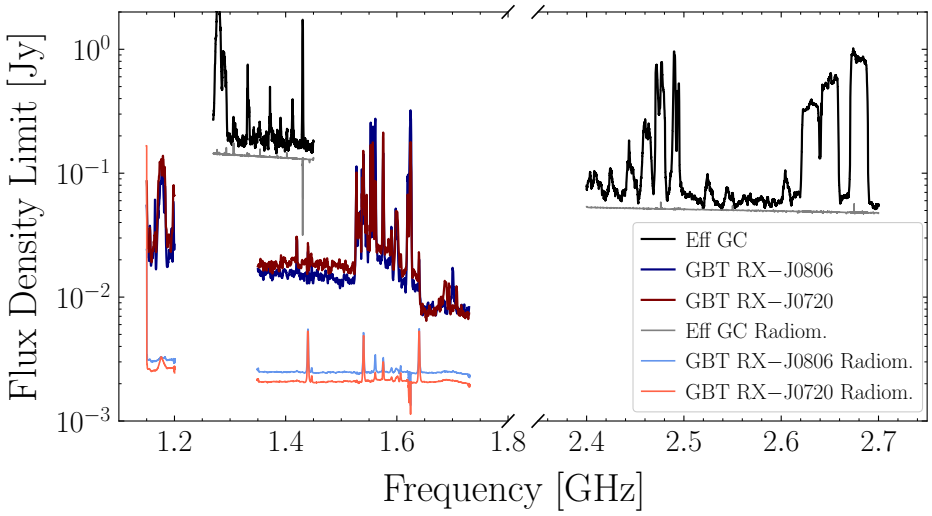



Figure 7.1: The 95% upper limits on the signal flux for the indicated sources from the GBT and Effelsberg observations. These upper limits apply to monochromatic signals at the widths $\delta f_{\rm fid}$ given in Tab. 7.1. These curves have been down-sampled for visualization purposes. We compare these limits with the 95% upper limits expected from the ideal radiometer equation under the assumption that the only source of statistical uncertainty is thermal noise at the total system temperature.

captured by the calibration procedure.

We search for evidence of an axion signal by using the discovery TSs. We apply a discovery threshold of $\mathrm{TS} > 100$, which was defined before performing the analysis and not modified afterwards. From Monte Carlo (MC) simulations of the null hypothesis we find that this TS threshold corresponds to approximately $5\sigma$ local significance (see the appendices for details). After applying the analysis procedure described above we find no axion signal candidates at or beyond the detection significance in any of the observations, and the distributions of observed TSs are consistent with the null hypothesis. Note that HI emission frequencies are excluded automatically in our analysis by the OFF veto criterion.

## 7.4 Results

To translate the flux-density limits from Fig. 7.1 into limits on the axion-photon coupling, we closely follow the theoretical modeling presented in [103, 104] for computing the axion-induced radio fluxes from these specific sources.

The radiated power for a single INS depends on $g_{a\gamma\gamma}$, the polar magnetic field strength $B_0$ (assuming a dipole field configuration), the NS mass (which we fix at $1\ M_\odot$, since this value does

not significantly affect the flux), the NS spin period $P$, the axion mass $m_a$, the DM density $\rho_\infty$ in the neighborhood of the NS but asymptotically far away from its gravitational potential, and the velocity dispersion $v_0$ of the ambient DM. For the local INSs we take $v_0 = 200\,\mathrm{km/s}$ and $\rho_\infty = 0.4$ GeV$/$cm$^3$ [28, 65, 296]. For the GC analysis we assume the DM follows an Navarro-Frenk-White (NFW) [297, 298] density profile near the GC, normalized to give the local DM density above and with a scale radius of 20 kpc (see, *e.g.*, [104]). For RX J0806.4−4123 we take $\log_{10}(B_0/G) = 13.40$ and $P = 11.4\,\mathrm{s}$, while for RX J0720.4−3125 we use $\log_{10}(B_0/G) = 13.53$ and $P = 8.4\,\mathrm{s}$. We note that these parameters were inferred from spin-down measurements performed in the $X$-ray band [299–301]. We take RX J0806.4−4123 and RX J0720.4−3125 to be at distances of 250 pc and 360 pc from Earth, respectively [300].

Given these parameters, we estimate the radiated power following [103, 288]. However, we note that a fully self consistent calculation of the axion-induced radiation has yet to be performed. Ref. [288] corrected the assumption in [103] that the axions travel along radial trajectories, but [288] did not account for the fact that the outgoing radiation is strongly refracted in the inhomogeneous magnetosphere, as we point out in the appendices. As a dedicated simulation of the axion-induced radiation is beyond the scope of this work, we estimate the power with the following approximation. We assume that (i) all axions travel along radial trajectories, as in [103], (ii) that all NSs are aligned rotators (magnetically-misaligned rotators give nearly identical results [103]), and (iii) that the magnetosphere is well-described by the Goldreich-Julian (GJ) model [302]. Then, following [103] we compute the angular power distribution $dP/d\theta$ of radio emission as a function of the angle from the polar axis $\theta$. However, we assign to each NS a single power value equal to $\int \frac{dP}{d\theta}\, d\theta$, and we assume that the flux is radiated from each NS isotropically. With the latter assumption we find results are consistent with those in [288], which correctly accounted for the isotropic axion phase space. For example, taking NS parameters describing the nearby isolated NSs studied in this work (and assuming aligned rotators) the formalism in [288] predicts fluxes $\sim$50% larger than inferred by our simpler calculation. We chose this simpler formalism, however, because it is likely that the more complicated computation in [288] must be modified due to the refraction of outgoing radio photons, which could result in an anisotropic signal(though from the studies in [288] we do not expect such calculations to change the flux predictions by more than an $\mathcal{O}(1)$ amount). Given an improved theoretical predictions in the future, our results may be reinterpreted using the Supplementary Data [303].

In [104] it is shown that more complicated magnetosphere models, such as the electrosphere model, give similar results. In particular, the total radiated power averaged over NS populations differed by $\sim$20% between the electrosphere and GJ models in [104]. Active pulsars and magnetars could have magnetospheres which deviate more substantially from the GJ model by having large charge-pair multiplicities, though this is expected to only affect a small fraction of the NSs

within the populations and to not affect the nearby isolated NSs studied in this work (see [104] and references therein).

The width of the signal in frequency space is determined in part by the asymptotic energy dispersion of the DM, which is set by $v_0$. This induces a $\delta f/f \lesssim 10^{-6}$ contribution to the width from the INSs. However, as discussed more in the appendices and in [287], the signals are Doppler-broadened when refracting or reflecting from the rotating plasma, inducing a frequency broadening closer to $\delta f/f \sim 5 \times 10^{-6}$ and justifying the bin widths taken in Tab. 7.1.

Since we do not actually know which specific NSs are being targeted in the Effeslberg GC analysis (and similarly in the GBT population analyses discussed in the appendices), we model the population of NSs (number density, spatial distribution, magnetic field, and spin period) within the GC region as a whole, closely following [104]. In particular, two models for the NS magnetic field and period distributions were developed in that work, based on fits to existing pulsar data. We conservatively choose the model which yields weaker constraints as our fiducial model. In practice, our fiducial NS population model (Model II in [104]) assumes that magnetic fields quickly decay after the NSs cross the pulsar death-line, while the optimistic model (Model I in [104]) assumes that the magnetic fields decay more slowly. We also follow [104] when modeling the spatial distribution of NSs within the Galactic bulge and disk. For the Effelsberg analysis, we perform $\mathcal{O}(10^3)$ MC simulations of the NS population model and profile over the simulation results when calculating the expected flux and associated 95% limit.

Given the fiducial models we have described, we obtain the limits on $g_{a\gamma\gamma}$ shown in Fig. 7.2. The orange band represents the predicted $g_{a\gamma\gamma}$ for the QCD axion, and the grey shaded regions represent existing constraints from other experiments. We obtain limits that are stronger than those from CAST [263] and comparable to constraints from the UF [260] and RBF [261,290] haloscopes, while the S-band Effelsberg constraints exclude previously unexplored parameter space. The green shaded band in Fig. 7.2 represents two dominant sources of uncertainty for the GC analysis. The top of the band is derived by assuming that the DM profile follows a cored density profile with a core radius of 0.6 kpc; this radius is chosen based on recent hydrodynamic simulations which suggest that the DM density may be modified in the inner $\sim$0.6 kpc where the baryons dominate the gravitational potential, though these same simulations suggest an enhancement of the central DM density may also be possible [304]. The lower boundary of the band assumes the fiducial NFW DM profile but takes the alternate NS population model (Model I) from [104].

## 7.5 Conclusion

In this work we performed the first dedicated radio telescope search for signatures of axion DM from axion-photon conversion in NS magnetospheres. We found no evidence for axion DM and set

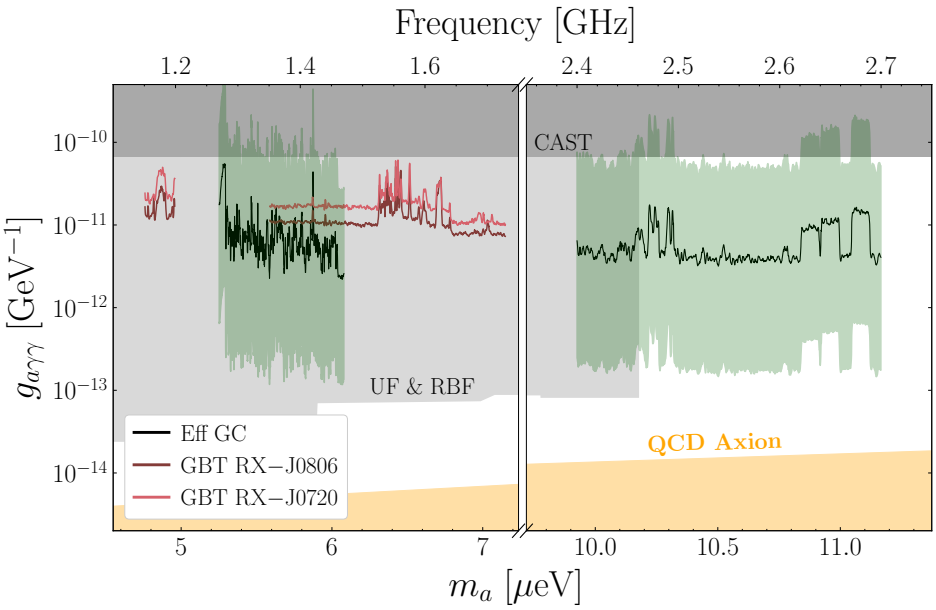



Figure 7.2: The one-sided 95% upper limits on $g_{a\gamma\gamma}$ as a function of the axion mass $m_a$ from this work are shown as colored lines (GBT INS observations) and black lines (Effelsberg GC observations). Previous limits from the CAST helioscope and the UF and RBF haloscopes are shown in shaded grey. The range of couplings expected for the QCD axion is shaded in orange. Note that the fiducial GC limits assume an NFW DM profile and the conservative NS population model (Model II) from [104]. The green band depicts theoretical uncertainties on the $g_{a\gamma\gamma}$ limit associated with the GC analysis for the Effelsberg data. The top of the band assumes an NFW DM density profile with a 0.6 kpc core, while the bottom of the band uses the alternate NS population model in [104] (Model I).

some of the strongest constraints to date on the axion DM scenario. These results show that radio searches for axion DM are a promising path forward, analogous to indirect detection for WIMP DM searches, which should proceed in parallel with laboratory experiments for discovering or excluding axion DM. Additional flux sensitivity is needed in order to reach the QCD axion band at the frequencies targeted in this work. This sensitivity may be available with the upcoming Square Kilometer Array-mid [305] or may already be achievable with the FAST radio telescope [306], since at constant system temperature the sensitivity to $g_{a\gamma\gamma}$ scales inversely with the square root of the effective area [104].

Our work strongly motivates searching with the GBT or Effelsberg radio telescope for evidence of axion DM at higher frequencies, closer to 6 GHz, to probe the axion mass window around $m_a \approx 25\ \mu$eV. There is mounting evidence that points towards 25 $\mu$eV as a likely mass for the axion [79, 112], and the axion-photon coupling may also be enhanced [169] and thus within reach of GBT and Effelsberg searches. This work also motivates additional effort in modeling the population

evolution of NS magnetic fields and spin periods, as these are the largest sources of uncertainty in our population analyses, as well as further efforts to understand the distribution of DM in the inner Galaxy. More work on the axion-induced signal itself from individual INSs would be also useful, as a full calculation of the axion-induced radio signal does not yet exist; such results could lead to reinterpretations of the limits presented in this Letter using the Supplementary Data [303].

# CHAPTER 8

# X-Ray Searches for Axions from Super Star Clusters

Ultralight axion-like particles that couple weakly to ordinary matter are natural extensions to the Standard Model. For example, string compactifications often predict large numbers of such pseudo-scalar particles that interact with the Standard Model predominantly through dimension-five operators [51, 52]. If an axion couples to quantum chromodynamics (QCD) then it may also solve the strong *CP* problem [38, 39, 43, 44]; in this work we refer to both the QCD axion and axion-like particles as axions.

Axions may interact electromagnetically through the operator $\mathcal{L} = -g_{a\gamma\gamma} a F_{\mu\nu} \tilde{F}^{\mu\nu}/4$, where $a$ is the axion field, $F$ is the electromagnetic field-strength tensor, with $\tilde{F}$ its Hodge dual, and $g_{a\gamma\gamma}$ is the dimensionful coupling constant of axions to photons. This operator allows both the production of axions in stellar plasmas through the Primakoff Process [307,308] and the conversion of axions to photons in the presence of static external magnetic fields. Strong constraints on $g_{a\gamma\gamma}$ for low-mass axions come from the CERN Axion Solar Telescope (CAST) experiment [263], which searches for axions produced in the Solar plasma that free stream to Earth and then convert to $X$-rays in the magnetic field of the CAST detector. CAST has excluded axion couplings $g_{a\gamma\gamma} \gtrsim 6.6 \times 10^{-11}$ GeV$^{-1}$ for axion masses $m_a \lesssim 0.02$ eV at 95% confidence [263]. Primakoff axion production also opens a new pathway by which stars may cool, and strong limits ($g_{a\gamma\gamma} \lesssim 6.6 \times 10^{-11}$ GeV$^{-1}$ at 95% confidence for $m_a \lesssim$ keV) are derived from observations of the horizontal branch (HB) star lifetime, which would be modified in the presence of axion cooling [309].

In this work, we produce some of the strongest constraints to-date on $g_{a\gamma\gamma}$ for $m_a \lesssim 10^{-9}$ eV through $X$-ray observations with the Nuclear Spectroscopic Telescope Array (NuSTAR) telescope [310] of super star clusters (SSCs), which are relatively young, luminous, and compact clusters of stars produced in regions of particularly high stellar formation rates. Thought to be the progenitors of globular clusters, SSCs contain a large number of hot, young, and massive stars, such as Wolf-Rayet (WR) stars. Many of the known SSCs have been detected in extragalactic targets by the Hubble Space Telescope, but several within the Milky Way have been detected. As we will show, the stars within SSCs are highly efficient at producing axions with energies ∼10–100 keV through the Primakoff process. Axions produced in Milky Way SSCs may then convert

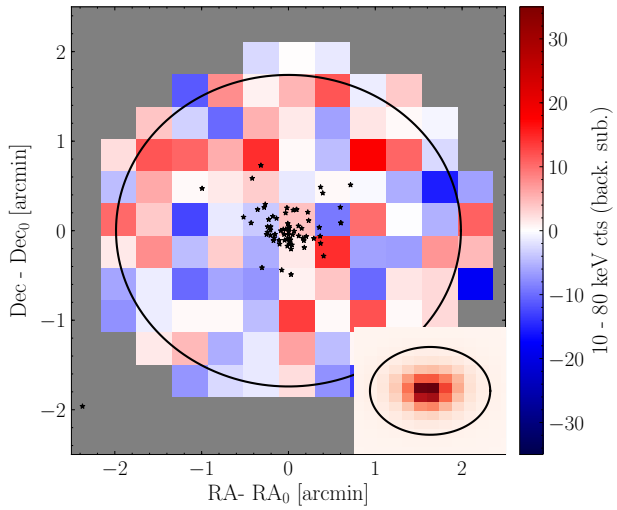



Figure 8.1: The stacked and pixelated background-subtracted count data (10 - 80 keV) from the NuSTAR observations of the Quintuplet SSC. The locations of the stars are indicated in black, while the 90% energy containment region for emission associated with the SSC is indicated by the black circle, accounting for the NuSTAR point spread function (PSF). $\mathrm{RA}_0$ and $\mathrm{DEC}_0$ denote the locations of the cluster center. We find no evidence for axion-induced emission from this SSC, which would follow the spatial counts template illustrated in the inset panel.

into photons in the Galactic magnetic fields, leading to signatures observable with space-based $X$-ray telescopes such as NuSTAR. We analyze archival NuSTAR data from the Quintuplet SSC near the Galactic Center (GC) along with the nearby Westerlund 1 (Wd1) cluster and constrain $g_{a\gamma\gamma} \lesssim 3.6 \times 10^{-12}$ GeV$^{-1}$ at 95% confidence for $m_a \lesssim 5 \times 10^{-11}$ eV. In Fig. 8.1 we show the locations of the stars within the Quintuplet cluster that are considered in this work on top of the background-subtracted NuSTAR counts, from 10 - 80 keV, with the point-spread function (PSF) of NuSTAR also indicated. In the appendices we show that observations of the Arches SSC yield similar but slightly weaker limits.

Our work builds upon significant previous efforts to use stars as laboratories to search for axions. Some of the strongest constraints on the axion-matter couplings, for example, come from examining how HB, white dwarf (WD), red giant, and neutron star (NS) cooling would be affected by an axion [309, 311–319]. When the stars have large magnetic fields, as is the case for WDs and NSs, the axions can be converted to $X$-rays in the stellar magnetospheres [301, 320–322]. Intriguingly, in [301, 322] observations of the Magnificent Seven nearby isolated NSs found evidence for a hard $X$-ray excess consistent with the expected axion spectrum from nucleon bremsstrahlung. This work extends these efforts by allowing the axions to convert to $X$-rays not just in the stellar magnetic fields but also in the Galactic magnetic fields [323–325].

## 8.1 Organization of this Chapter

A brief description of the set of procedures implemented in this Chapter to constrain axions is as follows.

- Use the Modules for Experiments in Stellar Astrophysics (MESA) stellar evolution code to model the interiors of the hot, young stars within Milky Way SSCs.
- Predict the axion production from SSC stars using the MESA-generated stellar interior models combined with the axion emissivities through the Primakoff process.
- Convolve the radiated spectrum axions with the energy-dependent conversion probability of axions propagating through the Galactic Magnetic Field from their stellar origin to Earth.
- Generate an energy-dependent template of axion-induced X-ray flux from stars in SSCs that is then convolved with the NuSTAR instrumental response.
- Reduce archival data collected with the NuSTAR telescope at three known Milky Way SSCs: Arches, Quintuplet, and Westerlund 1 to produce X-ray flux maps.
- Analyze the X-ray images in search of emission consistent with the previously generated axion-induced X-ray flux templates, setting limits in the absence of a detection.

These procedures are explained in greater detail in the subsequent sections and the corresponding Appendices.

## 8.2 Axion production in SSCs

During helium burning, particularly massive stars may undergo considerable mass loss, especially through either rotation or binary interaction, which can begin to peel away the hydrogen envelope, revealing the hot layers underneath and reversing the cooling trend. Stars undergoing this process are known as WR stars, and these stars are the most important in our analyses. If the star has a small ($<$40% abundance) remaining hydrogen envelope, it is classified as a WNh star; at $<$5% hydrogen abundance it is classified as a WN star; otherwise, it is classified as WC or WO, which indicates the presence of $>$2% carbon, and oxygen, respectively, in the atmosphere.

Axions are produced through the photon coupling $g_{a\gamma\gamma}$ in the high-mass stars in SSCs through the Primakoff process $\gamma + (e^-, Z) \rightarrow a + (e^-, Z)$. This process converts a stellar photon to an axion in the screened electromagnetic field of the nucleons and electrons. The massive stars are high-temperature and low-density and therefore form nonrelativistic nondegenerate plasmas. The

Primakoff emission rate was calculated in [308, 326] as a function of temperature, density, and composition, and is described in detail in the appendices.

To compute the axion luminosity in a given star, we use the stellar evolution code Modules for Experiments in Stellar Astrophysics (MESA) [327, 328] to find, at any particular time in the stellar evolution, radial profiles of temperature, density, and composition. The simulation states are specified by an initial metallicity $Z$, an initial stellar mass, an initial rotation velocity, and an age. The initial metallicity is taken to be constant for all stars. In the appendices we show that the Quintuplet and Arches clusters, which are both near the GC, are likely to have initial metallicities in the range $Z \in (0.018, 0.035)$, consistent with the conclusions of previous works which place the initial metallicities of these clusters near solar (solar metallicity is $Z \approx 0.02$) [329, 330]. Note that higher metallicities generally lead to the stars entering the WR classifications sooner, when their cores are cooler. Rotation may also cause certain massive stars to be classified as WR stars at younger ages. We model the initial rotation distribution as a Gaussian distribution with mean $\mu_{\rm rot}$ and standard deviation $\sigma_{\rm rot}$ for non-negative rotation speeds [331,332]. Refs. [331,332] found $\mu_{\rm rot} \approx 100$ km/s and $\sigma_{\rm rot} \approx 140$ km/s, but to assess systematic uncertainties we vary $\mu_{\rm rot}$ between $50$ and $150$ km/s [331].

We draw initial stellar velocities from the velocity distribution described above (from 0 to 500 km/s) and initial stellar masses from the Kroupa initial mass function [333] (from 15 to 200 $M_\odot$). We use MESA to evolve the stars from pre-main-sequence (pre-MS)–before core hydrogen ignition–to near-supernova. At each time step we assign each stellar model a spectroscopic classification using the definitions in [334,335]. We then construct an ensemble of models for each spectroscopic classification by joining together the results of the different simulations that result in the same classification for stellar ages within the age range for star formation in the cluster; for Quintuplet, this age range is between 3.0 and 3.6 Myr [336]. Note that each simulation generally provides multiple representative models, taken at different time steps. In total we compute $10^5$ models per stellar classification.

Quintuplet hosts 71 stars of masses $\gtrsim 50 M_\odot$, with a substantial WR cohort [336]. In particular it has 14 WC + WN stars, and we find that these stars dominate the predicted axion flux. For example, at $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ we compute that the total axion luminosity from the SSC (with $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s) is $2.1^{+0.7}_{-0.4} \times 10^{35}$ erg/s, with WC + WN stars contributing $\sim$70% of that flux. Note that the uncertainties arise from performing multiple (500) draws of the stars from our ensembles of representative models. In the 10 - 80 keV energy range relevant for NuSTAR the total luminosity is $1.7^{+0.4}_{-0.3} \times 10^{35}$ erg/s. We take $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s because these choices lead to the most conservative limits. For example, taking the metallicity at the lower-end of our range ($Z = 0.018$) along with $\mu_{\rm rot} = 100$ km/s the predicted 10 - 80 keV flux increases by $\sim$60%. At fixed $Z = 0.035$ changing $\mu_{\rm rot}$ from 150 km/s to 100 km/s increases the total luminosity

(over all energies) by $\sim$10%, though the luminosity in the 10 - 80 keV range is virtually unaffected.

The Wd1 computations proceed similarly. Wd1 is measured from parallax to be a distance $d \in (2.2, 4.8)$ kpc from the Sun [337], accounting for both statistical and systematic uncertainties [338]. Wd1 is estimated to have an age between 4.5 and 7.1 Myr from isochrone fitting, which we have broadened appropriately from [339] accounting for expanded distance uncertainties. In our fiducial analysis we simulate the stars in Wd1 for initial metallicity $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s as this leads to the most conservative flux predictions, even though it is likely that the metallicity is closer to solar for Wd1 [340], in which cases the fluxes are larger by almost a factor of two (see the appendices). We model 153 stars in Wd1 [339], but the axion flux is predominantly produced by the 8 WC and 14 WN stars. In total we find that the 10 - 80 keV luminosity, for $g_{a\gamma\gamma} = 10^{-12}$ GeV, is $9.02^{+1.2}_{-1.1} \times 10^{35}$ erg/s, which is $\sim$5 times larger than that from Quintuplet.

## 8.3 Axion conversion in Galactic fields

The axions produced within the SSCs may convert to $X$-rays in the Galactic magnetic fields. The axion Lagrangian term $\mathcal{L} = g_{a\gamma\gamma} a \mathbf{E} \cdot \mathbf{B}$, written in terms of electric and magnetic fields $\mathbf{E}$ and $\mathbf{B}$, causes an incoming axion state to rotate into a polarized electromagnetic wave in the presence of an external magnetic field (see, *e.g.*, [341]). The conversion probability $p_{a\to\gamma}$ depends on the transverse magnetic field, the axion mass $m_a$, and the plasma frequency $\omega_{\rm pl} \approx 3.7 \times 10^{-12} (n_e/10^{-2}\ {\rm cm}^{-3})^{-1/2}$ eV, with $n_e$ the free-electron density (see the appendices for an explicit formula). Note that hydrogen absorption towards all of our targets is negligible, being at most $\sim$5% in the 15-20 keV bin of the Quintuplet analysis [342].

To compute the energy-dependent conversion probabilities $p_{a\to\gamma}$ for our targets we need to know the magnetic field profiles and electron density distributions along the lines of sight. For our fiducial analysis we use the regular components of the JF12 Galactic magnetic field model [343, 344] and the YMW16 electron density model [345] (though in the appendices we show that the `ne2001` [346] model gives similar results), though the JF12 model does not cover the inner kpc of the Galaxy. Outside of the inner kpc the conversion probability for Quintuplet is dominated by the out-of-plane (X-field) component in the JF12 model. We conservatively assume that the magnitude of the vertical magnetic field within the inner kpc is the same as the value at 1 kpc ($|B_z| \approx 3$ $\mu$G), as illustrated in Fig. G.6. In our fiducial magnetic field model the conversion probability is $p_{a\to\gamma} \approx 2.4 \times 10^{-4}$ ($7 \times 10^{-5}$) for $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ for axions produced in the Quintuplet SSC with $m_a \ll 10^{-11}$ eV and $E = 80$ keV ($E = 10$ keV). Completely masking the inner kpc reduces these conversion probabilities to $p_{a\to\gamma} \approx 1.0 \times 10^{-4}$ ($p_{a\to\gamma} \approx 3.2 \times 10^{-5}$), for $E = 80$ keV ($E = 10$ keV). On the other hand, changing global magnetic field model to that presented in [347] (PTKN11), which has a larger in-plane component than the JF12 model but no out-of-

plane component, leads to conversion probabilities at $E = 80$ and 10 keV of $p_{a\to\gamma} \approx 4.9 \times 10^{-4}$ and $p_{a\to\gamma} \approx 4.2 \times 10^{-5}$, respectively, with the inner kpc masked.

The magnetic field is likely larger than the assumed 3 $\mu$G within the inner kpc. Note that the local interstellar magnetic field, as measured directly by the *Voyager* missions [348], indirectly by the Interstellar Boundary Explorer [349], inferred from polarization measurements of nearby stars [350], and inferred from pulsar dispersion measure and the rotation measure data [351], has magnitude $B \sim 3$ $\mu$G, and all evidence points to the field rising significantly in the inner kpc [352]. For example, Ref. [353] bounded the magnetic field within the inner 400 pc to be at least 50 $\mu$G, and more likely 100 $\mu$G (but less than $\sim$400 $\mu$G [354]), by studying non-thermal radio emission in the inner Galaxy. Localized features in the magnetic field in the inner kpc may also further enhance the conversion probability beyond what is accounted for here. For example, the line-of-sight to the Quintuplet cluster overlap with the GC radio arc non-thermal filament, which has a $\sim$3 mG vertical field over a narrow filament of cross-section $\sim(10\ \mathrm{pc})^2$ (see, *e.g.*, [355]). Accounting for the magnetic fields structures described above in the inner few hundred pc may enhance the conversion probabilities by over an order of magnitude relative to our fiducial scenario (see the appendices).

When computing the conversion probabilities for Wd1 we need to account for the uncertain distance $d$ to the SSC (with currently-allowable range given above). In the JF12 model we find the minimum $p_{a\to\gamma}/d^2$ (for $m_a \ll 10^{-11}$ eV) is obtained for $d \approx 2.6$ kpc, which is thus the value we take for our fiducial distance in order to be conservative. At this distance the conversion probability is $p_{a\to\gamma} \approx 2.4 \times 10^{-6}$ ($p_{a\to\gamma} \approx 1.5 \times 10^{-6}$) for $E = 10$ keV ($E = 80$ keV), assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ and $m_a \ll 10^{-11}$ eV. We note that the conversion probabilities are over 10 times larger in the PTKN11 model (see the appendices), since there is destructive interference (for $d \approx 2.6$ kpc) in the JF12 model towards Wd1. We do not account for turbulent fields in this analysis; inclusion of these fields may further increase the conversion probabilities for Wd1, although we leave this modeling for future work.

## 8.4 Data analysis

We reduce and analyze 39 ks of archival NuSTAR data from Quintuplet with observation ID `40010005001`. This observation was performed as part of the NuSTAR Hard X-ray Survey of the GC Region [356, 357]. The NuSTAR data reduction was performed with the HEASoft software version 6.24 [358]. This process leads to a set of counts, exposure, and background maps for every energy bin and for each exposure (we use data from both Focal Plane Modules A and B). The astrometry of each exposure is calibrated independently using the precise location of the source 1E 1743.1-2843 [359], which is within the field of view. The background maps account

for the cosmic $X$-ray background, reflected solar X-rays, and instrumental backgrounds such as Compton-scattered gamma rays and detector and fluorescence emission lines [360]. We then stack and re-bin the data sets to construct pixelated images in each of the energy bins. We use 14 5-keV-wide energy bins between 10 and 80 keV. We label those images $d_i = \{c_i^p\}$, where $c_i^p$ stands for the observed counts in energy bin $i$ and pixel $p$. The pixelation used in our analysis is illustrated in Fig. 8.1.

For the Wd1 analysis we reduced Focal Plane Module A and B data totaling 138 ks from observation IDs `80201050008`, `80201050006`, and `80201050002`. This set of observations was performed to observe outburst activity of the Wd1 magnetar CXOU J164710.2–45521 [361], which we mask at $0.5'$ in our analysis. (The magnetar is around 1.5' away from the cluster center.) Note that in [361] hard $X$-ray emission was only detected with the NuSTAR data from 3 - 8 keV from CXOU J164710.2–45521 – consistent with this, removing the magnetar mask does not affect our extracted spectrum for the SSC above 10 keV. We use the magnetar in order to perform astrometric calibration of each exposure independently. The Wd1 exposures suffer from ghost-ray contamination [362] from a nearby point source that is outside of the NuSTAR field of view at low energies (below $\sim$15 keV) [361]. (Ghost-ray contamination refer to those photons that reflect only a single time in the mirrors.) The ghost-ray contamination affects our ability to model the background below 15 keV and so we remove the 10 - 15 keV energy bin from our analysis.

In each energy bin we perform a Poissonian template fit over the pixelated data to constrain the number of counts that may arise from the template associated with axion emission from the SSC. To construct the signal template we use a spherically-symmetric approximation to the NuSTAR PSF [363] and we account for each of the stars in the SSC individually in terms of spatial location and expected flux, which generates a non-spherical and extended template. We label the set of signal templates by $S_i^p$. We search for emission associated with the signal templates by profiling over background emission. We use the set of background templates described above and constructed when reducing the data, which we label $B_i^p$.

Given the set of signal and background templates we construct a Poissonian likelihood in each energy bin:

$$p_i(d_i|\{S_i, A_B\}) = \sum_p \frac{(\mu_i^p)^{c_i^p} e^{-\mu_i^p}}{c_i^p!}\,, \tag{8.1}$$

with $\mu_i^p = S_i S_i^p + A_B B_i^p$. We then construct the profile likelihood $p_i(d_i|\{S_i\})$ by maximizing the log likelihood at each fixed $S_i$ over the nuisance parameter $A_B$. Note that when constructing the profile likelihood we use the region of interest (ROI) where we mask pixels further than 2.0' from the SSC center. The 90% containment radius of NuSTAR is $\sim$1.74', independent of energy, as

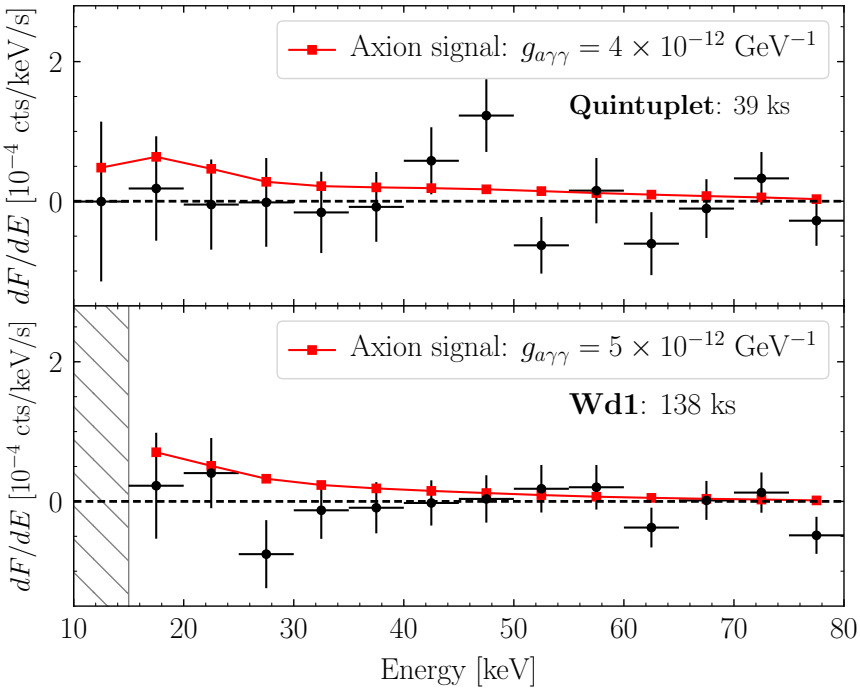



Figure 8.2: The spectra associated with the axion-induced templates from the Quintuplet and Wd1 SSCs constructed from the NuSTAR data analyses, with best-fit points and $1\sigma$ uncertainties indicated. In red we show the predicted spectra from an axion with $m_a \ll 10^{-11}$ eV and indicated $g_{a\gamma\gamma}$. Note that for Wd1 we do not analyze the 10 - 15 keV energy bin because of ghost-ray contamination.

indicated in Fig. 8.1. We use a localized region around our source to minimize possible systematic biases from background mismodeling. However, as we show in the appendices our final results are not strongly dependent on the choice of ROI. We also show in the appendices that if we inject a synthetic axion signal into the real data and analyze the hybrid data, we correctly recover the simulated axion parameters.

The best-fit flux values and $1\sigma$ uncertainties extracted from the profile likelihood procedure are illustrated in Fig. 8.2 for the Quintuplet and Wd1 data sets. We compare the spectral points to the axion model prediction to constrain the axion model. More precisely, we combine the profile likelihoods together from the individual energy bins to construct a joint likelihood that may be used to search for the presence of an axion signal: $p(d|\{m_a, g_{a\gamma\gamma}\}) = \prod_i p_i\big[d_i|R_i(m_a, g_{a\gamma\gamma})\big]$, where $R_i(m_a, g_{a\gamma\gamma})$ denotes the predicted number of counts in the $i^{\rm th}$ energy bin given an axion-induced $X$-ray spectrum with axion model parameters $\{m_a, g_{a\gamma\gamma}\}$. The values $R_i(m_a, g_{a\gamma\gamma})$ are computed using the forward-modeling matrices constructed during the data reduction process.

In Fig. 8.3 we illustrate the 95% power-constrained [195] upper limits on $g_{a\gamma\gamma}$ as a function of the axion mass $m_a$ found from our analyses. The joint limit (red in Fig. 8.3), com-

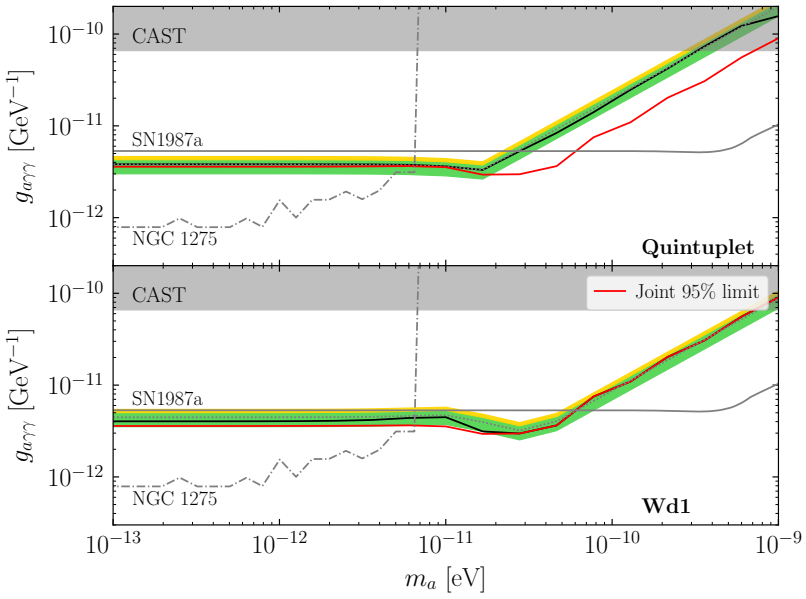



Figure 8.3: The 95% upper limits (black) on $g_{a\gamma\gamma}$ as a function of the axion mass from the Quintuplet and Wd1 data analyses. We compare the limits to the $1\sigma$ (green band) and $2\sigma$ (yellow band) expectations under the null hypothesis, along with the median expectations (dotted). The joint 95% upper limit, combining Quintuplet and Wd1, is also indicated (expected joint limit not shown). At low masses our limits may be surpassed by those from searches for $X$-ray spectral modulations from NGC 1275 [364], though we caution that those limits have been called into question recently, as discussed further in the text [365].

bining the Quintuplet and Wd1 profile likelihoods, becomes $g_{a\gamma\gamma} \lesssim 3.6 \times 10^{-12}$ GeV$^{-1}$ at low axion masses. At fixed $m_a$ the upper limits are constructed by analyzing the test statistic $q(g_{a\gamma\gamma}|m_a) \equiv 2\ln p(d|\{m_a, g_{a\gamma\gamma}\}) - 2\ln p(d|\{m_a, \bar{g}_{a\gamma\gamma}\})$, where $\bar{g}_{a\gamma\gamma}$ is the signal strength that maximizes the likelihood, allowing for the possibility of negative signal strengths as well. The 95% upper limit is given by the value $g_{a\gamma\gamma} > \bar{g}_{a\gamma\gamma}$ such that $q(g_{a\gamma\gamma}|m_a) \approx 2.71$ (see, *e.g.*, [172]). The $1\sigma$ and $2\sigma$ expectations for the 95% upper limits under the null hypothesis, constructed from the Asimov procedure [172], are also shown in Fig. 8.3. The evidence in favor of the axion model is $\sim 0.3\sigma$ ($0\sigma$) local significance at low masses for Quintuplet (Wd1).

We compare our upper limits with those found from the CAST experiment [263], the non-observation of gamma-rays from SN1987a [366] (see also [367–369] along with [370], who recently questioned the validity of these limits), and the NGC 1275 $X$-ray spectral modulation search [364]. It was recently pointed out, however, that the limits in [364] are highly dependent on the intracluster magnetic field models and could be orders of magnitude weaker, when accounting for both regular and turbulent fields [365]. The CAST limits are stronger than ours for $m_a \gtrsim 10^{-9}$ eV and rely on less modeling assumptions, since CAST searches for axions produced in the Sun,

though we have made conservative choices in our stellar modeling.

## 8.5 Conclusion

We present limits on the axion-photon coupling $g_{a\gamma\gamma}$ from a search with NuSTAR hard $X$-ray data for axions emitted from the hot, young stars within SSCs and converting to $X$-rays in the Galactic magnetic fields. We find the strongest limits from analyses of data towards the Quintuplet and Wd1 clusters. Our limits represent some of the strongest and most robust limits to-date on $g_{a\gamma\gamma}$ for low-mass axions. We find no evidence for axions. Promising targets for future analyses could be nearby supergiant stars, such as Betelgeuse [323, 371], or young NSs such as Cas A.

# CHAPTER 9

# X-Ray Searches for Decaying Dark Matter in the Milky Way

Sterile neutrino dark matter (DM) is a well-motivated DM candidate that may give rise to observable nearly monochromatic X-ray signatures [372–374]. In this scenario the DM has a mass in the keV range and may decay into an active neutrino and an X-ray, with energy set by half the rest mass of the sterile neutrino [375]. Sterile neutrino DM is motivated in part by the seesaw mechanism for explaining the active neutrino masses [376, 377]. In this work we present one of the most sensitive searches for sterile neutrino DM, along with other DM candidates that may decay to monochromatic X-rays, over the mass range $m_\chi \in [5, 16]$ keV. We do so by searching for DM decay from the ambient halo of the Milky Way using all archival data from the *XMM-Newton* telescope collected from its launch until September 5, 2018.

This work builds heavily off the method developed in [378], which used *XMM-Newton* blank-sky observations (BSOs) to strongly disfavor the decaying DM explanation of the previously-observed 3.5 keV unidentified X-ray line (UXL). This UXL was found in nearby galaxies and clusters [379–383]. However the analysis performed in [378] was able to robustly rule out the DM decay rate required to explain the previous 3.5 keV UXL signals [384]. (For additional non-observations, see Refs. [385–391].) We extend the search in [378] to the broader mass range $m_\chi \in [5, 16]$, and in doing so implement the following notable differences: (i) we use a data-driven approach to construct stacked, background-subtracted data sets in rings around the Galactic Center (GC), while Ref. [378] performed a joint-likelihood analysis at the level of individual exposures, and (ii) we use Gaussian Process (GP) modeling to describe continuum residuals, instead of parametric modeling as used in [378].

As demonstrated in [378], BSO searches for DM decaying in the Milky Way halo can be both more sensitive and more robust than extra-galactic searches, because (i) the expected DM flux, even at angles $\sim$45° away from the GC, rivals the expected flux from the most promising extra-galactic objects, such as M31 and the Perseus cluster; (ii) promising extra-galactic targets have continuum and line-like X-ray features that are confounding backgrounds for DM searches (dwarf galaxies

being an exception [389, 392]), while BSOs instead focus on the lowest-background regions of the sky; (iii) extra-galactic targets require pointed observations, while in principle any observation collected by *XMM-Newton* is sensitive to DM decay in the Milky Way, opening up considerably more exposure time.

The limits presented in this work represent the strongest found using the *XMM-Newton* instrument over the energy range ∼2.5–8 keV. At higher energies our limits are superseded with those found using the NuSTAR satellite [393–397]. Ref. [395] performed a search similar in spirit to that in this work (though with NuSTAR) in that they looked for DM decay from the Milky Way halo near the GC ($\sim 10^\circ$ away in their case), while Ref. [397] searched for DM decay from M31 with NuSTAR. Our results put in tension efforts to explain the abundance of DM with sterile neutrinos. DM models such as axion-like-particle DM [398] and moduli DM [399] predict similar UXL signatures from DM decay and are also constrained by this work.

## 9.1 Data Reduction and Processing

We process and analyze all publicly-available data collected before 5 September 2018 by the metal oxide semiconductor (MOS) and positive-negative (PN) cameras on board *XMM-Newton*. We subject each exposure to a set of quality cuts, which are described shortly. Those exposures satisfying the quality cuts are included in our angularly-binned data products. In particular, we divide the sky into 30 concentric annuli centered around the GC, each with a width of $6^\circ$ in angular radius from the GC, $r_{\rm GC}$, where $\cos(r_{\rm GC}) = \cos(l)\cos(b)$ in terms of the Galactic longitude, $l$, and latitude, $b$. We label these from 1 to 30, starting from the innermost ring. We further mask the Galactic Plane such that we only include the region $|b| \geq 2^\circ$. In each ring we then produce stacked spectra where, in each energy bin, we sum over the counts from each exposure whose central position lies within that annulus. We produce separate data sets for the MOS and PN cameras, which have 2400 and 4096 energy channels, respectively. In addition to stacking the counts in each ring and energy channel, we also construct the appropriately weighted detector response matrices in every ring for forward modeling an incident astrophysical flux. The full-sky maps and associated modeling data are provided as Supplementary Data [400] in both the annuli and in finer-resolution `HEALPix` binning [401]. We analyze the MOS data from 2.5 to 8 keV and the PN data from 2.5 to 7 keV, in order to exclude intervals containing large instrumental features.

## 9.2 Data Analysis

Having constructed our data in all 30 rings, we divide the full sky into two regions of interest (ROI): a signal ROI, consisting of annuli 1 through 8 ($0^\circ \leq r_{\rm GC} \leq 48^\circ$ with $|b| \geq 2^\circ$), inclusive,

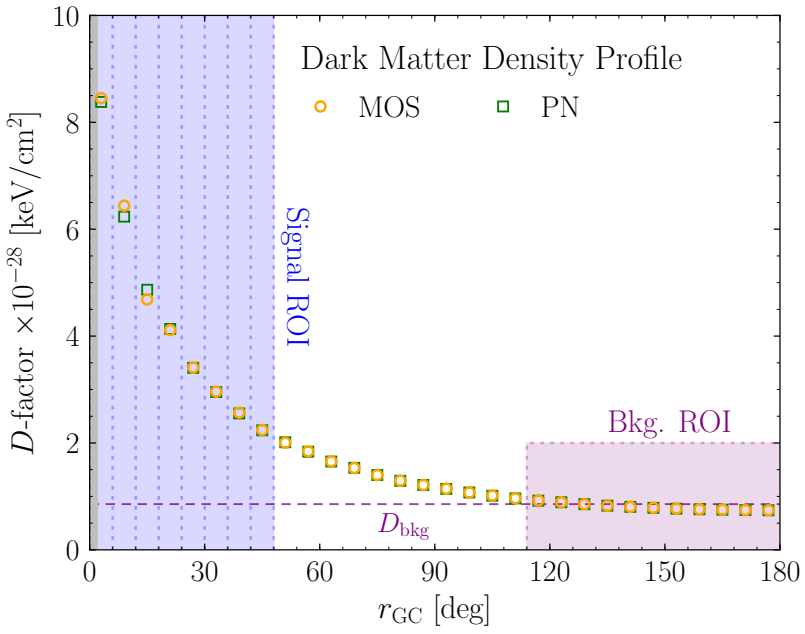



Figure 9.1: Our fiducial $D$-factor, which is proportional to the expected DM signal flux. Values are given in all 30 annuli, which are $6^\circ$ wide in angular distance from the GC (with $|b| > 2^\circ$), and we define a signal and background ROI as shown. In each ring, we compute the $D$-factor of all MOS or PN exposures, weighted according to the observation time and field of view. The horizontal line indicates $D_{\rm bkg}$, the mean $D$-factor in the background ROI.

and the background ROI, consisting of annuli 20 through 30 ($114^\circ \leq r_{\rm GC} \leq 180^\circ$ with $|b| \geq 2^\circ$). The regions are illustrated in Fig. 9.1. The MOS (PN) exposure time in the signal ROI is 25.27 Ms (5.56 Ms), whereas in the background ROI it is 62.51 Ms (17.54 Ms). The signal flux of decay producing a single photon, given by

$$F = \frac{D}{4\pi m \tau} \tag{9.1}$$

is proportional to the $D$-factor which is defined by the line-of-sight integral of the Galactic DM density $\rho_{\rm DM}$ by

$$D(\Omega) \equiv \int ds\, \rho_{\rm DM}(s, \Omega). \tag{9.2}$$

Here we have defined $m$ the mass of the decaying particle and $\tau$ as its lifetime. Hence we see that a model-independent constraint on the DM lifetime can immediately be inferred given an known $D$-factor while translating these to bounds on a given model merely requires translation of $\tau$ to the relevant particle physics parameters. Decay to two photons can be trivially accommodated by multiply the flux in Eq. 9.1 by a factor of two.

In Fig. 9.1 we show the appropriately weighted $D$-factor in each annuli. The motivation for the two ROIs is that the signal should dominate in the inner regions of the Galaxy and become pro-

gressively weaker further away from the GC. The background ROI is chosen to be large enough to have significantly more exposure time than the signal ROI, so that using the background-subtracted data does not significantly broaden the statistical uncertainties. We stack the data over the full background ROI, which has $D$-factor $D_{\rm bkg}$, and use this as an estimate of the instrumental and astrophysical baseline fluxes by subtracting this data from the data in each ring of the signal ROI. This subtraction mostly removes large instrumental lines, as illustrated in Fig. H.1.

We analyze the background-subtracted data in each annulus for evidence of a UXL. The data is modeled as a combination of narrow spectral features at the locations of known astrophysical and instrumental lines, and a continuum flux which we account for using GP modeling. Note that the instrumental lines need not be completely removed by the data-subtraction procedure, leaving a residual flux or flux deficit that must be modeled. Astrophysical emission lines from the Milky Way plasma should be brighter in the signal ROI, and so are also expected to appear in the background-subtracted data. For both astrophysical and instrumental lines, the lines are modeled using the forward modeling matrices for MOS and PN. We allow the instrumental lines to have either positive or negative normalizations, while the astrophysical lines are restricted to have positive normalizations. To decide which lines to include in our residual background model we start with an initial list of known instrumental and astrophysical lines. The instrumental lines are determined from an analysis of the background ROI data, while the astrophysical lines are those expected to be produced by the Milky Way. In each ring, and for MOS and PN independently, we then determine the significance of each emission line, keeping those above $\sim 2\sigma$. As a result, every ring has a different set of lines included in the analysis. We note that it is conceivable that a UXL might be inadvertently removed by an overly-subtracted instrumental line at the same energy; however, it would be highly unlikely for such a conspiracy to occur in every ring, given the varying $D$-factor. The effects of sub-threshold instrumental lines are mitigated through a *spurious-signal* nuisance parameter [402], as discussed in the SM.

The unprecedented data volume incorporated into this analysis necessitates a flexible approach to modeling the residual continuum emission, which is accomplished with GP modeling, in order to minimize background mismodeling. As opposed to parametric modeling, where the model is specified by a specific functional form and associated list of model parameters, GP modeling is non-parametric: the model expectations for the data at two different energies, $E$ and $E'$, are assumed to be normally distributed with non-trivial covariance. Taking the model expectation to have zero mean, the GP model is then fully specified by the covariance kernel, $K(E, E')$. We model the mean-subtracted data using the non-stationary kernel $K(E, E') = A_{\rm GP} \exp\left[-(E - E')^2/(2EE'\sigma_E^2)\right]$, implemented in `george` [403], where $\sigma_E$ is the correlation-length hyperparameter and $A_{\rm GP}$ is the amplitude hyperparameter. We fix $\sigma_E$ such that it is larger than the energy resolution of the detector, which is $\delta E/E \sim 0.03$ across most energies for MOS

and PN, while ensuring $\sigma_E$ is kept small enough to have the flexibility to model real variations in the data. The goal is to balance two competing effects. If $\sigma_E$ approaches the lower limit imposed by the energy resolution of the detector, then the GP model would have the flexibility to account for line-like features, which would reduce our sensitivity when searching for such features over the continuum background. On the other hand, if $\sigma_E$ is too large then the GP continuum model may not accurately model real small-scale variations in the data. In our fiducial analysis we fix $\sigma_E = 0.3$, though in the appendices we show that our results are robust to variations not only in this choice, but also to modifications to the form of the kernel itself. In contrast, the hyperparameter $A_{\rm GP}$ is treated as a nuisance parameter that is profiled over when searching for UXLs.

We then follow the statistical approach developed in [404], which used GP modeling to perform an improved search for narrow resonances over a continuum background in the context of the Large Hadron Collider. In particular, we construct a likelihood ratio $\Lambda$ between the model with and without the signal component, where the signal is the UXL line at fixed energy $E_{\rm sig}$. The null model is as above, the combination of a GP model with a single nuisance parameter $A_{\rm GP}$, and a set of background lines, whose amplitudes are treated as nuisance parameters. We use the marginal likelihood from the GP fit in the construction of the likelihood ratio [404]. Note that as the number of counts in all energy bins is large ($\gg 100$), we are justified in assuming normally-distributed errors in the context of the GP modeling. We then profile over all nuisance parameters. Finally, the discovery significance is quantified by the test statistic (TS) $t = -2\ln\Lambda$. We verify explicitly in the appendices that under the null hypothesis $t$ follows a $\chi^2$-distribution. The 95% one-sided upper limits are constructed from the profile likelihood, as a function of the signal amplitude.

We implement this procedure and scan for a UXL from 2.5 to 8 keV in 5 eV intervals. At each test point we construct profile likelihoods for signal flux independently for each ring using the background-subtracted MOS and PN data. We then combine the likelihoods between rings – and eventually cameras – in a joint likelihood in the context of the DM model, as discussed shortly. As an example, Fig. 9.2 illustrates the signal and null model fits to the innermost MOS background-subtracted signal-annulus data for a putative UXL at 3.5 keV (indicated by the vertical dashed line). Note that while the fit is performed over the full energy range (2.5−8 keV) for clarity we show the data zoomed in to the range 3 to 4 keV. In this case the data have a deficit, which manifests itself as a signal with a negative amplitude.

## 9.3 Interpretation

We combine together the profile likelihoods from the individual annuli to test the decaying DM model. In the context of sterile neutrino DM with mass $m_\chi$ and mixing angle $\theta$, the DM decay in the Galactic halo produces an X-ray flux at energy $m_\chi/2$ that scales as $\Phi \propto m_\chi^4 D \sin^2(2\theta)$ [405].

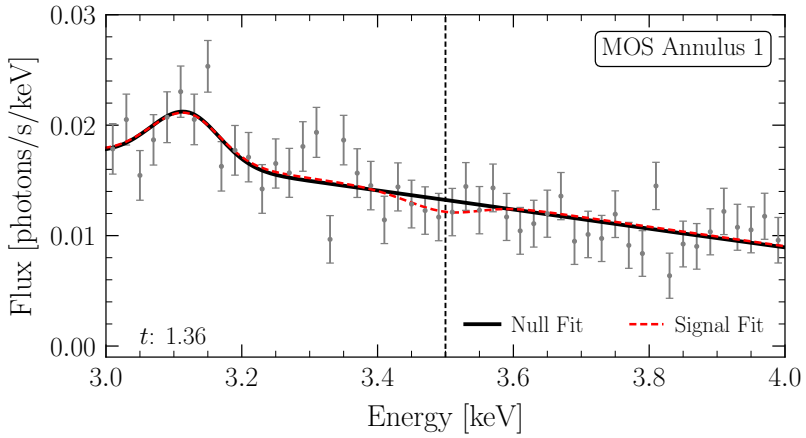



Figure 9.2: The background-subtracted MOS data for the innermost annulus, downbinned by a factor of four for presentation purposes. The indiciated best fit null and signal models, for a 3.5 keV UXL, are constructed using the GP modeling described in the text.

Specifically, the translation from the model-independent $\tau$ to $\sin^2(2\theta)$ is given by

$$\frac{1}{\tau} = 1.361 \times 10^{-29}\,\mathrm{s}^{-1} \left(\frac{\sin^2 2\theta}{10^{-7}}\right) \left(\frac{m_s}{1\,\mathrm{keV}}\right)^5 . \tag{9.3}$$

It is in terms of this $\sin^2(2\theta)$ that we present the majority of our results, though we do provide general limits on $\tau$ and on $g_{a\gamma\gamma}$ for the axion scenario. Note that the $D$-factors, appropriately averaged over observations in the individual annuli, are illustrated in Fig. 9.1. Thus, at fixed DM mass $m_\chi$ we may construct profile likelihoods as functions of $\sin^2(2\theta)$ to appropriately combine the profile likelihoods as functions of flux in the individual annuli. We subtract off $D_{\mathrm{bkg}}$ from the $D$-factors in each signal ring since any UXL would also appear in the background ROI and thus be included in the background subtraction.

The $D$-factors may be computed from the DM density profile of the Milky Way. Modern hydrodynamic cosmological simulations indicate that the DM density profile in Milky Way mass halos generally have a high degree of spherical symmetry (for a review, see Ref. [406]). Further, the presence of baryons contracts the inner $\sim$10 kpc of the profile away from the canonical Navarro, Frenk, and White (NFW) DM distribution [297, 298], so that there is an enhancement of the DM density at smaller radii versus the NFW expectation [304, 407–411], though cores could develop on top of this contraction at radii $\lesssim 2$ kpc [412–415]. For example, in Milky Way analogue halos within the `Fire-2` simulations the DM-only and hydrodynamic simulations produce DM density profiles that agree within $\sim$25% at 10 kpc, but with baryons the density profiles are typically around twice as large as the NFW DM-only expectation at distances $\sim$1 kpc away from the GC [304]. To be

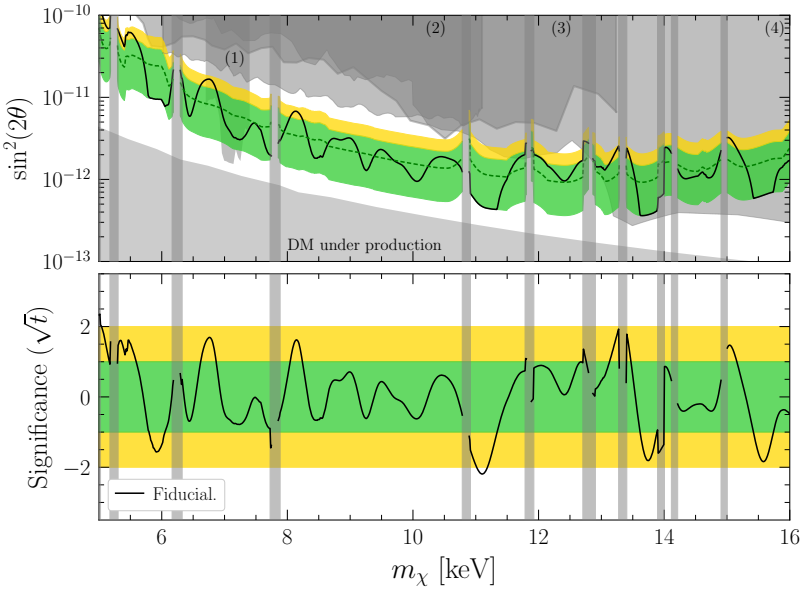



Figure 9.3: (Upper) The power-constrained 95% upper limit on the DM lifetime from this work, presented in the context of the sterile-neutrino mixing angle $\sin^2(2\theta)$, as a function of the DM mass $m_\chi$. The dark grey regions correspond to theoretical bounds for DM underproduction in the $\nu$MSM or bounds from previous X-ray searches (1)–(5); see text for details. (Lower) The associated sign-weighted significance for the UXL. Vertical grey regions denote background lines and are at least partially masked. Green and gold regions indicate 1/2$\sigma$ expectations under the null hypothesis. These results are shown in the context of more general DM models as constraints on the DM lifetime in Fig. H.7.

conservative we assume the canonical NFW density profile for all radii, though in the appendices we discuss how our results change for alternate density profiles.

The NFW profile is specified by a characteristic density $\rho_0$ and a scale radius $r_s$: $\rho_{\rm DM}(r) = \rho_0/(r/r_s)/(1+r/r_s)^2$. We use the recent results from [416], who combined *Gaia* DR2 Galactic rotation curve data [417] with total mass estimates for the Galaxy from satellite observations [418, 419]. These data imply, in the context of the NFW model, a virial halo mass $M_{200}^{\rm DM} = 0.82^{+0.09}_{-0.18} \times 10^{12}\ M_\odot$ and a concentration $c = r_{200}/r_s = 13.31^{+3.60}_{-2.68}$, with a non-trivial covariance between $M_{200}^{\rm DM}$ and $c$ [416] such that lower concentrations prefer higher halo masses. Within the 2D 68% containment region for $M_{200}^{\rm DM}$ and $c$ quoted in Ref. [416], the lowest DM density at $r \approx 0.5$ kpc, and thus the most conservative profile for the present analysis, is obtained for $\rho_0 = 6.6 \times 10^6\ M_\odot/{\rm kpc}^3$ and $r_s = 19.1$ kpc. We adopt these values for our fiducial analysis. With our choice of NFW DM parameters the local DM density, at the solar radius, is $\sim$0.29 GeV/cm$^3$ (*cf.* 0.4 GeV/cm$^3$ used in [378]), which is consistent with local measurements of the DM density using the vertical motion of tracer stars perpendicular to the Galactic plane, see, *e.g.*, Refs. [28, 420].

We search for evidence of decaying DM in 10 eV intervals in mass between $5-16$ keV, masking 0.1 keV windows around the locations of known lines, as indicated in Fig. 9.3. We construct the joint likelihoods for the MOS and PN data sets. We test and account for additional background mismodeling in the MOS and PN analyses by looking at the distribution of best-fit mixing angles in the energy side-bands, using a technique similar to the "spurious signal" used by ATLAS in the search for the Higgs boson [402]. This procedure is described in the appendices and only has a small effect at low masses. We then combine, at a given mass, the resulting MOS and PN profile likelihoods to obtain the final profile likelihood used to construct the limit and discovery significance shown in Fig. 9.3. In that figure we show the one-sided 95% upper limit on $\sin^2(2\theta)$ in the upper panel, along with the 1 and $2\sigma$ expectations for the power-constrained upper limit [195] under the null hypothesis (shaded green and gold, respectively).

We find no evidence for decaying DM signals above our pre-determined significance threshold of $5\sigma$ global significance (corresponding to $\sim 6\sigma$ local significance), as shown in the bottom panel. In that figure we compare our upper limit to previous limits in the literature, adjusted to our fiducial DM model for the Milky Way where appropriate. In the context of the $\nu$MSM it is impossible to explain all of the observed DM in the region marked "DM under production" because of the big bang nucleosynthesis bound on the lepton chemical potential [421–423]. Note that the $\nu$MSM also predicts that the DM becomes increasingly warm for decreasing $m_\chi$, which leads to tension with Milky Way satellite galaxy counts for low $m_\chi$: data from the Dark Energy Survey and other Galactic satellite surveys [424] constrains $m_\chi$ greater than $\sim$15–20 keV in the $\nu$MSM [425] (which can be strengthened further when combined with strong lensing measurements [22]), though we note that our results apply to more general DM production mechanisms that do not predict modifications to small-scale structure. In Fig. 9.3 we also show previous X-ray limits from (1) [378], (2) a *Chandra* search for DM decay in the Milky Way [426], (3) a *Chandra* search for DM decay in M31 [385], and (4) combined *NuSTAR* searches for DM decay: in the Milky Way [393–395], the Bullet Cluster [396], and M31 [397]. Note that the results from Milky Way searches have been adjusted to use the same DM density profile as in our fiducial analysis.

### 9.3.1 Axion Interpretation

Although this chapter is presented within the context of the sterile neutrino, these X-ray searches can also be interpreted in the context of a search for the decay of keV-scale axions to two photons. The axion lifetime is given in terms of its coupling as

$$\frac{1}{\tau} = \frac{g_{a\gamma\gamma}^2 m_a^3}{64\pi} \tag{9.4}$$

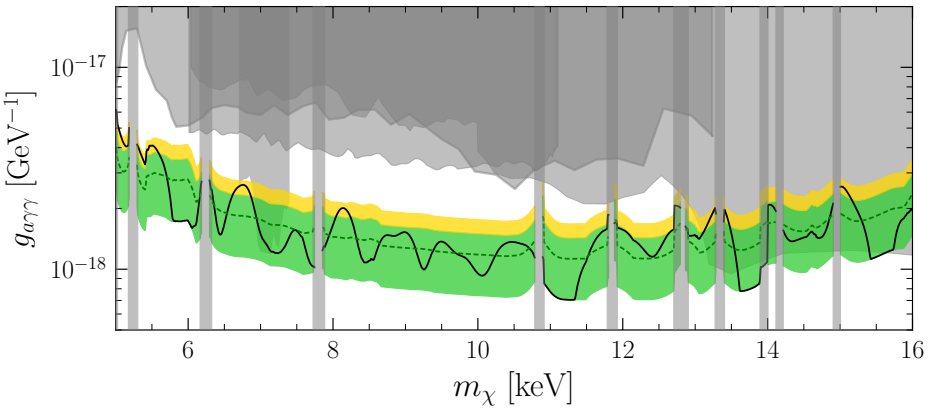



Figure 9.4: Our decay search interpreted in the context of limits on the axion decay to two photons. Our limits, along with those from additional blank sky searches are indicated as in Fig. 9.3

allowing for simple translation of our results to constraints on $g_{a\gamma\gamma}$. These constraints are presented in Fig. 9.4.

## 9.4 Conclusion

We find no significant evidence for decaying DM, which leads us to set some of the strongest constraints to-date on the DM lifetime. We confirm the results of Dessert *et al.* [378] for the non-observation of a DM decay line near 3.5 keV using a more robust and flexible analysis strategy, leaving little room for a decaying DM explanation of the previously-observed 3.5 keV anomalies [379–383]. (See the appendices for further discussion.)

Given the data volume incorporated into this analysis it is unlikely that further analyses of *XMM-Newton* data, or *Chandra* data, could produce qualitatively stronger results on the DM lifetime in the mass range considered here. However, the approach taken in this work may lead to a powerful advancement in discovery power with future data sets from surveys such as those by the upcoming *Athena* [427] and *XRISM* [428] telescopes. A combination of the data collected by those missions and the analysis framework introduced in this work may lead to the discovery of decaying DM in the few-keV mass range at lifetimes beyond those probed in this work.

# CHAPTER 10

# Conclusion

In this thesis, I have shown that promisingly open parameter space for axions can be effectively probed with new phenomenological insights combined with novel detection methods possible through astrophysical observation and precision laboratory measurement. As these searches are currently probing only the most accessible possibilities for axion DM, there remains considerable work to be done. However, there are many reasons to be optimistic. New ideas and steadily improving computational power will enable us to refine our understanding of axion phenomenology. Simultaneously, upcoming telescopes and experiments with improved sensitivities will lead to new and improved searches that will push the frontiers of axion detection further than previously thought possible. Moreover, there exists a wealth of unexamined archival data from astrophysics, some of which might already contain hints of new physics. There is every reason to believe that axions exist in nature, and with now unprecedented effort devoted to their discovery, prospects for shedding light on the nature of dark matter may be better than ever.

# APPENDIX A

# Simulations of Axion Production in the Post-Inflationary Misalignment Scenario

## A.1 Simulation Equations of Motion

Our phenomenological Lagrangian describing the PQ field is adopted from the construction of [121] and is of the form

$$\mathcal{L}_{PQ} = \frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}\left(|\Phi|^2 - f_a^2\right)^2 - \frac{\lambda T^2}{6}|\Phi|^2 - m_a(T)^2 f_a^2[1 - \cos \mathrm{Arg}(\Phi)], \tag{A.1}$$

where $\Phi$ is the complex PQ scalar, $T$ is the temperature, $\lambda$ is the PQ quartic coupling strength, $f_a$ is the PQ-scale identified as the axion decay constant, and $m_a(T)$ is the temperature-dependent axion mass [121]. The parametrization of the temperature-dependent mass is adopted from the leading order term in the fit in [116]. Explicitly, the axion mass is parametrized by

$$m_a(T)^2 = \min\left[\frac{\alpha_a \Lambda^4}{f_a^2 (T/\Lambda)^n},\, m_a\right], \tag{A.2}$$

for $\alpha = 1.68 \times 10^{-7}$, $\Lambda = 400\,\mathrm{MeV}$ and $n = 6.68$. The growth of the mass is truncated at $T \approx 100\,\mathrm{MeV}$. The zero-temperature mass is given by

$$m_a^2 = \frac{m_\pi^2 f_\pi^2}{f_a^2} \frac{m_u m_d}{(m_u + m_d)^2}, \tag{A.3}$$

where $m_\pi$ is the pion mass, $f_\pi$ is the pion decay constant, $m_{u/d}$ is the up/down quark mass. Details of the temperature-dependent axion mass, or equivalently, the topological susceptibility, remain uncertain, especially at low temperatures. Note that we do not explore here how our results are affected by uncertainties in the temperature-dependent axion mass, though doing so is a worthwhile direction for future work.

Decomposing the complex scalar as $\Phi = \phi_1 + i\phi_2$, and assuming a radiation-dominated cosmological background, leads to equations of motion in metric coordinates of the form

$$\begin{aligned}
\ddot{\phi}_1 + 3H\dot{\phi}_1 - \frac{1}{R^2}\nabla^2\phi_1 + \frac{1}{3}\lambda\phi_1\left[3\left(\phi_1^2+\phi_2^2-f_a^2\right)+T^2\right] - \frac{m_a(T)^2\phi_2^2}{\left(\phi_1^2+\phi_2^2\right)^{3/2}} = 0 \\
\ddot{\phi}_2 + 3H\dot{\phi}_2 - \frac{1}{R^2}\nabla^2\phi_2 + \frac{1}{3}\lambda\phi_2\left[3\left(\phi_1^2+\phi_2^2-f_a^2\right)+T^2\right] + \frac{m_a(T)^2\phi_1\phi_2}{\left(\phi_1^2+\phi_2^2\right)^{3/2}} = 0\,.
\end{aligned} \tag{A.4}$$

Over temperatures $T \gtrsim 100\,\mathrm{MeV}$, the number of relativistic degrees of freedom $g_*$ in the Standard Model is expected to vary only mildly. For simplicity, we therefore assume $g_* = 81$, which is a typical value adopted at high temperatures (though later in the Appendices we explore the systematic uncertainty introduced by this assumption). It is useful to define a dimensionless conformal time $\hat{\eta}$ such that

$$\hat{\eta} = \frac{R}{R(T=T_1)} = \frac{R}{R_1} = \left(\frac{t}{t_1}\right)^{1/2}\,, \tag{A.5}$$

where $R$ is the scale factor and the time $t_1$ (with $T(t_1) \equiv T_1$) is a reference time that will be defined differently in the PQ and QCD epoch simulations.

The axion-mass term is not included in our PQ-epoch simulations. In our QCD-epoch simulations, on the other hand, the mass term is included and drives the dynamics. In this case, the mass grows until the cutoff temperature $T_c$ at which point the axion mass reaches its zero-temperature value; the corresponding conformal time is given by $\hat{\eta}_c = R(T=T_c)/R_1$. Rewriting (A.4) with the dimensionless coordinates, we then find

$$\begin{aligned}
\psi_1'' + \frac{2}{\hat{\eta}}\psi_1' - \bar{\nabla}^2\psi_1 + \frac{1}{H_1^2}\Bigg[\lambda\psi_1\left(\hat{\eta}^2 f_a^2(\psi_1^2+\psi_2^2-1)+\frac{1}{3}T_1^2\right) \\
- m_a^2(T_1)\hat{\eta}^2\min(\hat{\eta},\hat{\eta}_c)^n\left(\frac{\psi_2^2}{(\psi_1^2+\psi_2^2)^{3/2}}\right)\Bigg] = 0
\end{aligned} \tag{A.6}$$

$$\begin{aligned}
\psi_2'' + \frac{2}{\hat{\eta}}\psi_2' - \bar{\nabla}^2\psi_2 + \frac{1}{H_1^2}\Bigg[\lambda\psi_2\left(\hat{\eta}^2 f_a^2(\psi_1^2+\psi_2^2-1)+\frac{1}{3}T_1^2\right) \\
+ m_a^2(T_1)\hat{\eta}^2\min(\hat{\eta},\hat{\eta}_c)^n\left(\frac{\psi_1\psi_2}{(\psi_1^2+\psi_2^2)^{3/2}}\right)\Bigg] = 0\,,
\end{aligned} \tag{A.7}$$

where $\phi = f_a\psi$, primes denote derivatives with respect to $\hat{\eta}$, and the spatial gradient is taken with respect to $\bar{x} = a_1 H_1 x$.

### A.1.1 The PQ Epoch

Simulations in the PQ epoch occur at $T \gg \Lambda_{\rm QCD}$ and so the temperature-dependent axion mass may be neglected. We therefore take our equations of motion to be

$$\begin{aligned}\psi_1'' + \frac{2}{\tilde{\eta}}\psi_1' - \bar{\nabla}^2\psi_1 + \lambda\psi_1\left[\tilde{\eta}^2\left(\psi_1^2+\psi_2^2-1\right)+\frac{T_1^2}{3f_a^2}\right] &= 0 \\ \psi_2'' + \frac{2}{\tilde{\eta}}\psi_2' - \bar{\nabla}^2\psi_2 + \lambda\psi_2\left[\tilde{\eta}^2\left(\psi_1^2+\psi_2^2-1\right)+\frac{T_1^2}{3f_a^2}\right] &= 0\,,\end{aligned} \tag{A.8}$$

and we fix $\tilde{\eta} = 1$ to be the time at which $H_1 = f_a$. Note that for our PQ-epoch simulations we refer to $\hat{\eta}$, defined in (A.5), as $\tilde{\eta}$ in order to avoid confusion with the dimensionless conformal time $\hat{\eta}$ used in the QCD-epoch simulations. The ratio $(T_1/f_a)^2$ is determined by

$$\left(\frac{T_1}{f_a}\right)^2 \approx 8.4\times10^5\left(\frac{10^{12}\,\mathrm{GeV}}{f_a}\right)\,. \tag{A.9}$$

In principle, it would seem that axions of different decay constants would require different simulations in the PQ epoch. However, this ratio is degenerate with our choice of physical box size and dynamical range in $\tilde{\eta}$ in a particular simulation, allowing us to perform only one PQ simulation and interpret its output as the initial state of the axion field for several different values of $f_a$. The key assumption behind this, however, is that at late times after the PQ phase transition the field enters the scaling regime so that we may reinterpret the output of the PQ simulation in the appropriately rescaled box as the initial state of the QCD simulation at much lower temperatures. Note that the value of $\lambda$ is a free parameter, which we naturally choose to be $\lambda = 1$ though it has little effect.

#### A.1.1.1 Initial Conditions for a PQ Scalar

We generate initial conditions for our PQ scalar by taking it to be described by a thermal distribution characterized by the temperature $T$ at the initial early time. As can be read off from the Lagrangian, each of the two fields has an effective mass of the form

$$m_{\rm eff}^2 = \lambda\left(\frac{T^2}{3} - f_a^2\right)\,. \tag{A.10}$$

Correlation functions of the initially-free massive scalar fields are given by

$$\langle \phi_i(x)\phi_j(y)\rangle = \delta_{ij}\int \frac{dk}{2\pi}\frac{n_k}{\omega_k}e^{ik\cdot(x-y)} \tag{A.11}$$

$$\langle \dot{\phi}_i(x)\dot{\phi}_j(y)\rangle = \delta_{ij}\int \frac{dk}{2\pi}n_k\omega_k e^{ik\cdot(x-y)} \tag{A.12}$$

$$\langle \dot{\phi}_i(x)\phi_j(y)\rangle = 0\,, \tag{A.13}$$

where overdots denote differentiation with respect to time, and we have defined

$$n_k = \frac{1}{e^{\omega_k/T}-1}, \qquad \omega_k = \sqrt{k^2+m_{\rm eff}^2}. \tag{A.14}$$

In momentum space, these correlation functions take the form

$$\langle \phi_i(k)\phi_j(k')\rangle = \frac{2\pi n_k}{\omega_k}\delta(k+k')\delta_{ij} \tag{A.15}$$

$$\langle \dot{\phi}_i(k)\dot{\phi}_j(k')\rangle = 2\pi n_k\omega_k\delta(k+k')\delta_{ij} \tag{A.16}$$

$$\langle \dot{\phi}_i(k)\phi_j(k')\rangle = 0\,. \tag{A.17}$$

Our simulations occur on a discrete lattice of finite size, so the correlation functions above lead to initial conditions set by a realization of a Gaussian random field specified in Fourier space by

$$\langle \phi_i(k)\rangle = 0, \qquad \langle |\phi_i(k)|^2\rangle = \frac{n_k}{\omega_k}L, \tag{A.18}$$

$$\langle \dot{\phi}_i(k)\rangle = 0, \qquad \langle |\dot{\phi}_i(k)|^2\rangle = n_k\omega_k L\,. \tag{A.19}$$

Note that we include the 50 lowest $k$-modes in each of the three directions when constructing the initial conditions, and we have verified that including more modes does not affect our results.

### A.1.2 Early Times in the QCD Epoch

During the QCD epoch, $T \sim \Lambda_{\rm QCD}$, and so the axion mass is non-negligible. Here, we define $\hat{\eta} = 1$ to be the time at which $H_1 = m_a(T_1)$, with the axion field beginning to oscillate shortly thereafter when $m_a = 3H$. The equations of motion are then given by

$$\begin{aligned}
&\psi_1'' + \frac{2}{\hat{\eta}}\psi_1' - \bar{\nabla}^2\psi_1 + \tilde{\lambda}\hat{\eta}^2\psi_1(\psi_1^2+\psi_2^2-1) - \min(\hat{\eta},\hat{\eta}_c)^n\hat{\eta}^2\left(\frac{\psi_2^2}{(\psi_1^2+\psi_2^2)^{3/2}}\right) = 0\\
&\psi_2'' + \frac{2}{\hat{\eta}}\psi_2' - \bar{\nabla}^2\psi_2 + \tilde{\lambda}\hat{\eta}^2\psi_2(\psi_1^2+\psi_2^2-1) + \min(\hat{\eta},\hat{\eta}_c)^n\hat{\eta}^2\left(\frac{\psi_1\psi_2}{(\psi_1^2+\psi_2^2)^{3/2}}\right) = 0\,,
\end{aligned} \tag{A.20}$$

where we have neglected the $T_1$ contribution to the PQ scalar mass as it is small compared to $f_a$. The parameter $\tilde{\lambda}$ is defined by

$$\tilde{\lambda} = \lambda \left( \frac{f_a}{m_a(T_1)} \right)^2 \tag{A.21}$$

and can be interpreted as the squared mass of the radial mode $|\Phi/f_a|$. For physical parameters we expect $\tilde{\lambda} \gg 1$, though in practice we find that the final results are relatively independent of $\tilde{\lambda}$ for moderately sized values of the parameter, as described in the main text and later in the Appendices. Indeed, our choices for $\tilde{\lambda}$ allow us to resolve the radial mode mass by more than a few grid-spacings, satisfying the requirement of [124] to accurately study the axion spectrum from string radiation. There exist additional criteria on the largeness of $\tilde{\lambda}$ such that the metastability of topological defects is preserved despite the unphysical smallness of simulated $\tilde{\lambda}$ in comparison with the rapidly increasing axion mass. At all times prior to expected defect collapse, our choices of $\tilde{\lambda}$ satisfy the simplest construction of these conditions [96], with our choice of $\tilde{\lambda} = 5504$ satisfying the most stringent criteria established in [125]. We note that we are largely unable to differentiate between simulations at any two particular values of $\tilde{\lambda}$, and that our choice of values appear to have minimal impact, as illustrated further below.

### A.1.3 Late Times in the QCD Epoch

The presence of topological defects in the axion field at early times during the QCD epoch requires that we fully simulate both degrees of freedom of the PQ field. Once the topological defects have collapsed, however, we are free to use the axion-only equations of motion. Our axion is defined by $a = f_a \mathrm{arctan2}(\phi_1, \phi_2)$ and has the Lagrangian

$$\mathcal{L} = \frac{1}{2}(\partial a)^2 - m_a^2(T) f_a^2 \left[ 1 - \cos\left( \frac{a}{f_a} \right) \right] , \tag{A.22}$$

along with corresponding equations of motion

$$\theta'' + \frac{2}{\hat{\eta}}\theta' - \bar{\nabla}^2\theta + \min(\hat{\eta}, \hat{\eta}_c)^n \hat{\eta}^2 \sin\theta = 0 \, . \tag{A.23}$$

Above, we define $\theta = a/f_a$. Evolving these equations of motion is formally equivalent to freezing out excitations of the radial mode by taking $\tilde{\lambda} \to \infty$, which more accurately recovers the true physics of the evolution of the axion field for realistic values of $f_a$. Note that the coordinate $\bar{x}$ and $\hat{\eta}$ here are identical to those used in evolving the two degrees of freedom of the complex scalar performed prior to defect collapse.

### A.1.4 Analytically Evolving in the Fixed-Mass Small-Field Limit

At late times when the axion mass has reached its zero-temperature value and the axion field has redshifted considerably so that $|\theta| \ll 1$, the equations of motion are linear and well-approximated by

$$\theta'' + \frac{2}{\hat{\eta}}\theta' - \nabla^2\theta + \hat{\eta}_c^n \hat{\eta}^2 \theta = 0. \tag{A.24}$$

We may solve this equation analytically by going to Fourier space and adopting an ansatz for the solution as

$$\theta(\hat{\eta}) = f(\hat{\eta}) \exp(i\mathbf{k}\cdot\mathbf{x})\,. \tag{A.25}$$

This ansatz leads to the equation

$$f''(\hat{\eta}) + \frac{2f'(\hat{\eta})}{\hat{\eta}} + f(\hat{\eta})\left(\hat{\eta}^2\hat{\eta}_c^n + \mathbf{k}^2\right) = 0\,, \tag{A.26}$$

which has the general solution

$$\begin{aligned} f(\hat{\eta}) = \frac{\exp(-\frac{i}{2}\hat{\eta}^2\hat{\eta}_c^{n/2})}{\hat{\eta}} \Bigg[ & C_1 H_{-\frac{1}{2}\hat{\eta}_c^{-n/2}\left(\hat{\eta}_c^{n/2}+i\mathbf{k}^2\right)}\left(\sqrt[4]{-1}\hat{\eta}\hat{\eta}_c^{n/4}\right) + \\ & C_2\, {}_1F_1\left(\frac{1}{4}\hat{\eta}_c^{-n/2}\left(\hat{\eta}_c^{n/2} + i\mathbf{k}^2\right); \frac{1}{2}; i\hat{\eta}^2\hat{\eta}_c^{n/2}\right)\Bigg]\,, \end{aligned} \tag{A.27}$$

for coefficients $C_1$ and $C_2$ determined by boundary conditions, and where $H_n$ and ${}_1F_1$ are the analytic continuations of the Hermite polynomials and the confluent hypergeometric function of the first kind, respectively. From this analytic solution, we can transfer late-time field configurations from our simulation to arbitrary large $\hat{\eta}$. Differentiation with respect to $\hat{\eta}$ may be straightforwardly performed to find $f'(\hat{\eta})$ at large $\hat{\eta}$ as well. The computation of the analytically continued Hermite polynomials and hypergeometric functions was performed with the python package `mpmath`.

We directly compare the differential mass spectrum at $\hat{\eta} = 7$ with the same field analytically evolved to $\hat{\eta} = \eta_{\text{MR}}$ in Fig. A.1. While the basic differential shape is the same, the $\hat{\eta} = 7$ results have a much wider distribution in $\delta$. In particular, all overdensities above $\delta > 10$ have vanished by the time matter-radiation equality is reached. However, the peak of the distribution is still around $\delta = 1$. Evolving the fields down to matter-radiation equality is important because many of the modes are generated with high momentum at the QCD epoch, causing the large overdensities to disperse by the time of matter-radiation equality.

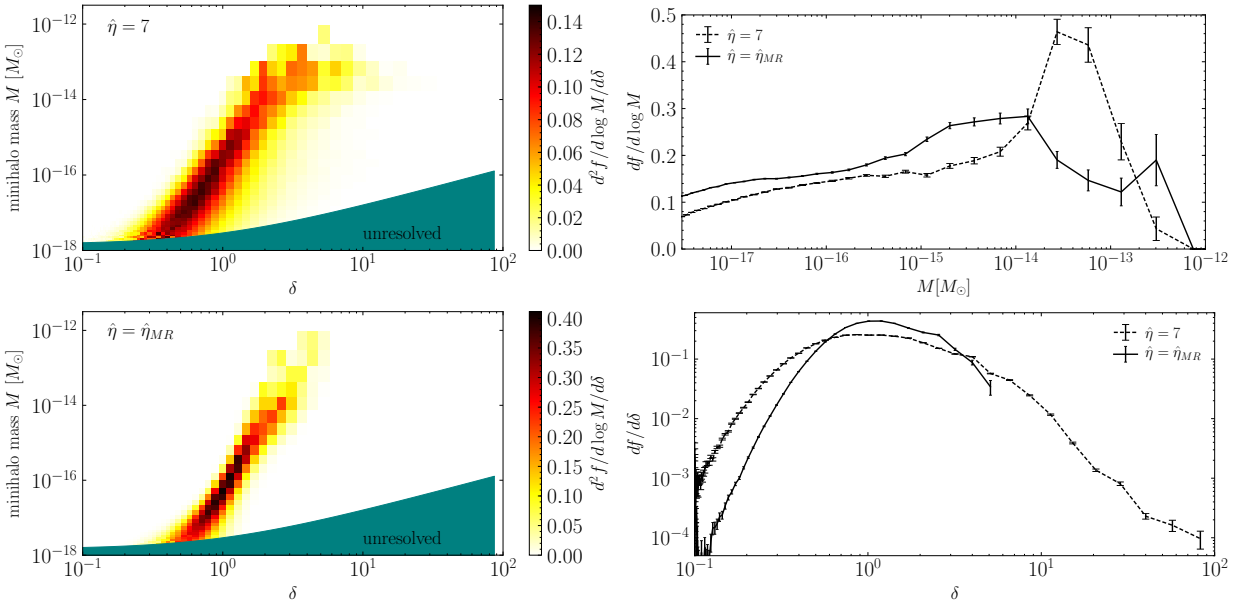



Figure A.1: Double differential mass fractions for axion minihalos as a function of the concentration parameter $\delta$ and mass $M$. In the top left we compute that mass function using the field immediately after the QCD phase transition, at $\hat{\eta} = 7$, while in the bottom left we use the more correct procedure of first evolving to $\hat{\eta} = \hat{\eta}_{\rm MR}$ before performing the clustering procedure. Evolving to matter-radiation equality gives the most over-dense regions time to expand and results in less dense overdensities, as compared to the incorrect procedure shown in the top left. This is perhaps even more apparent in the single differential mass fractions as a function of the mass $M$ (top right) and concentration parameter $\delta$ (bottom right). These results are based on our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$. Error bars are statistical, and we do not extend the $df/d\log M$ curves to lower masses as we are unable to resolve those properly.

## A.2 Studying the (Over)Density Field

Our interest in this work is studying the energy density field $\rho$ and the overdensity field $\delta = (\rho - \bar{\rho})/\rho$ realized in the axion field from our simulations. The axion energy density for the axion field $a = f_a\theta$ is computed by the Hamiltonian density

$$\mathcal{H} = f_a^2 \left[ \frac{1}{2}\dot{\theta}^2 + \frac{1}{2R^2}(\nabla\theta)^2 + m_a^2(1-\cos\theta) \right], \tag{A.28}$$

which can be rewritten in simulation units as

$$\mathcal{H} = m_a^2 f_a^2 \left[ \frac{\theta'^2 + (\bar{\nabla}\theta)^2}{2\hat{\eta}_c^{6.68}\hat{\eta}^2} + (1-\cos\theta) \right], \tag{A.29}$$

assuming $\hat{\eta} > \hat{\eta}_c$. At late times, the Hamiltonian is approximately

$$\mathcal{H} \approx \frac{m_a^2 f_a^2}{2} \left( \frac{\theta'^2}{\hat{\eta}_c^n \hat{\eta}^2} + \theta^2 \right) , \tag{A.30}$$

when all modes in the simulation are non-relativistic and the field values are small.

### A.2.1 Oscillons

Large overdensities right after the QCD phase transition are caused by oscillons. Oscillons are, in contrast to strings and domain walls, not topological defects but arise due to non-linearities in the equation of motions, forming at locations where the the axion self-interaction dominates the Hubble friction. As a result, the first oscillons form at the location of collapsed strings and domain walls, where the axion remains excited and reaches large field values. However, at later times, oscillons are observed forming throughout the simulation box. The dynamics of the oscillons are highly non-trivial, especially as the axion self-interaction increases in strength with the growing axion mass.

Oscillons decrease in size over time following the oscillation wavelength $\sim m_a(T)^{-1}$, as axions in the core are relativistic. Good spatial resolution is therefore needed to resolve them. In order to study their behavior we perform a 2D (two spatial dimensions, one time) simulation using the same simulation setup in the PQ- and QCD-epoch as in 3D. We find that there is no qualitative difference between 2D and 3D simulations regarding oscillons, but going to 2D allows us to increase the spatial resolution to $4096^2$ grid sites and to subsequently increase $\hat{\eta}_c$. Note that in the context of our 2D simulations we have explicitly verified that the final-state fields and, in particular, our central results, such as those concerning the DM density estimates and the non-Gaussian density distributions, are stable to increases in the number of grid sites. For example, we increased our resolution by a factor of two in each direction and found consistent results.

We illustrate the evolution of an oscillon in Fig. A.2. Two scenarios are considered with different truncation points of the mass growth, $\hat{\eta}_c = 4.0$ and $\hat{\eta}_c = 6.0$. Note how the radius of the oscillon decreases as long as $m_a(T)$ is increasing. The circles in A.2 have radius $m_a(T)^{-1}$, and the oscillon cores are seen to track this scale. Subsequently, if the mass growth is truncated at $\hat{\eta}_c = 4.0$, the radius of the oscillon is constant as well. When the mass growth is cut-off, the density contrast at the core of the oscillon slowly decreases over time and the oscillons dissipate, as can be seen in the two lower right panels in Fig. A.2.

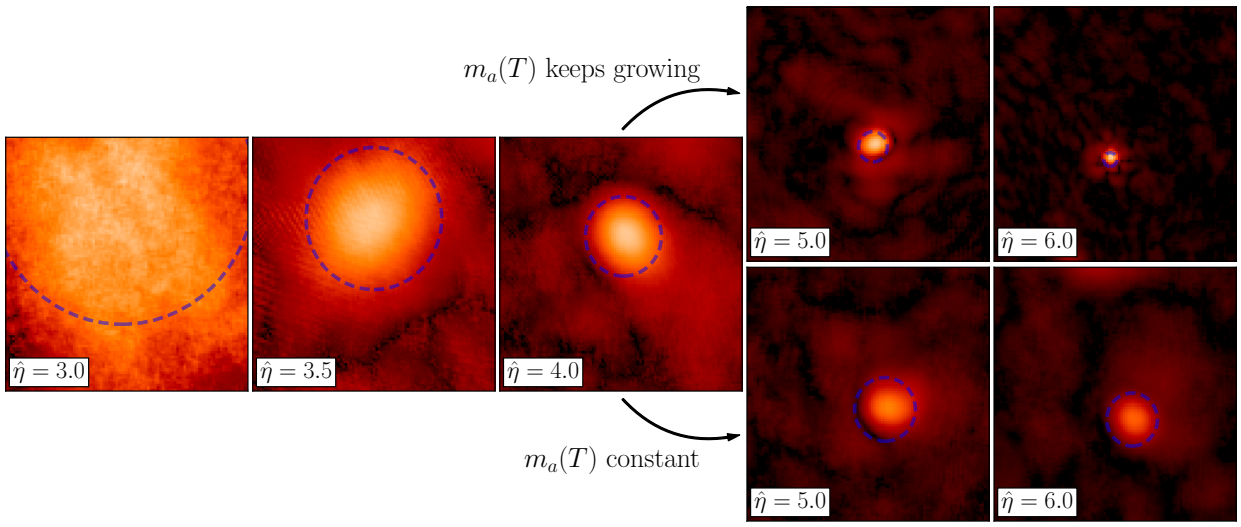



Figure A.2: Illustration of an oscillon ($\log(\rho/\bar{\rho})$) at different times in a 2D simulation. Two scenarios are considered with different truncation points of the mass growth, $\hat{\eta}_c = 4.0$ and $\hat{\eta}_c = 6.0$. The three left panels are identical in both scenarios, while the two top right panels are for $\hat{\eta}_c = 6.0$, and the two bottom panels are for $\hat{\eta}_c = 4.0$. The radius of the oscillon is proportional to the oscillation frequency $\sim m_a(T)^{-1}$ (circles of that radius are shown in dashed blue) and as such is decreasing over time. The oscillon central density slowly dissipates after the mass growth ends, as seen in the bottom right panels for $\hat{\eta}_c = 4.0$.

### A.2.2 Calculating the Axion Relic Abundance

To calculate the axion DM abundance as a function of $m_a$, we first need to understand the relationship between the mass cutoff conformal time $\hat{\eta}_c$ and the decay constant $f_a$. Here we use the relation $T_1/\hat{\eta}_c = T_c$, with $T_c \approx 100$ MeV. This allows us to solve for $f_a$ in terms of $\hat{\eta}_c$. The energy densities are calculated from the axion field and its derivatives according to (A.30) after numerically evolving until $\hat{\eta} = 7$, then analytically evolving until $\hat{\eta}_{\rm MR} = 10^6$, at which point the contribution of the gradient term to the energy density is negligible. As a side note, our definition of $\hat{\eta}_{\rm MR}$ actually puts us at slightly earlier times than global matter-radiation equality. This us because matter-radiation equality is, locally, reached earlier for the largest overdensities and because we want to make sure that gravitational interactions can be neglected. In particular, note that the temperature corresponding to $\hat{\eta}_{\rm MR}$ is given by $T_{\rm MR} = T_c\eta_c/\hat{\eta}_{\rm MR}$. For our most realistic simulation with $\eta_c = 3.6$ this corresponds to $T_{\rm MR} \approx 0.5$ keV. However, if we reinterpret the final state for a more realistic axion with $m_a \approx 25$ $\mu$eV, which has a higher $\eta_c$, then $T_{\rm MR} \approx 4$ keV. In practice, though, the exact value of $T_{\rm MR}$ is not important because by these temperatures the proper motions in the axion field are frozen out and the field is thus not evolving non trivially. As a consequence

our results (both for the DM density and for the spectrum of overdensities) are not sensitive to small (or even relatively large) changes to the exact value of $\hat{\eta}$ that we evolve to.

Note that we present our results in terms of the DM density fraction today $\Omega_a$, which is defined as the ratio of the average energy density today in DM relative to the observed critical energy density. We compute statistical error bars at each value of $f_a$ from the variance as a function of $\tilde{\lambda}$ at fixed $\hat{\eta}_c$. We note that no trend is visible in the data for the dependence of $\Omega_a$ on $\tilde{\lambda}$, as is shown in Fig. A.3. The statistical noise is inferred from the spread in $\Omega_a$ values, which are determined from

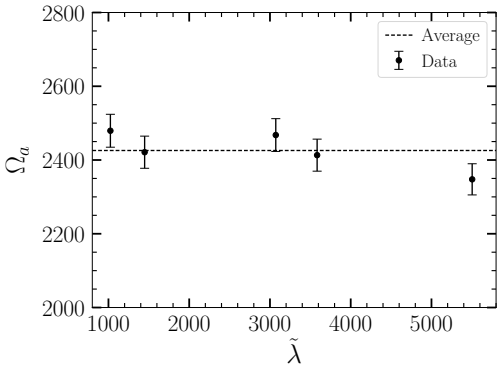



Figure A.3: Our results for the DM density today $\Omega_a$, inferred at $\hat{\eta}_{\rm MR}$, from simulations at different values of $\tilde{\lambda}$ for our most realistic $\hat{\eta}_c$: $\hat{\eta}_c = 3.6$. The uncertainties are the inferred statistical uncertainties arising from the spread in the DM density determinations as a function of $\tilde{\lambda}$. No trend is discernible for the dependence of $\Omega_a$ on $\tilde{\lambda}$, above the statistical noise.

the output at $\eta_{\rm MR}$, between different $\tilde{\lambda}$. The observed variations are consistent with the expected noise from Poisson counting statistics due to having a finite number of overdensities within the simulation box.

In Fig. A.4 we show our results for $\Omega_a$ as a function of $f_a$, compared to earlier predictions in [118] and [112]. For reference, we also include predictions for the relic abundance based on the field value and the time derivative at $\hat{\eta} = 7$. Here it is less straightforward to determine the DM axion abundance, relative to taking the results at $\hat{\eta}_{\rm MR}$, as some of the modes in the simulation are still relativistic. This introduces an additional systematic uncertainty, since the field is not completely red-shifting like radiation at this time. For these reasons it is important to evolve the field until it is completely non-relativistic before measuring the DM density.

Because the ratio of the axion mass density to entropy density is constant after the axions have become non-relativistic and the number of axions is conserved, we can redshift our energy density from our matter-radiation equality $\hat{\eta}_{\rm MR}$ to today. Then, we compare this energy density to the most up-to-date measurement of the average DM density in the Universe today $\rho_{\rm DM} = 33.5 \pm 0.6$ $M_\odot/{\rm kpc}^3$ [128]. Note that we have propagated all cosmological uncertainties other than those on

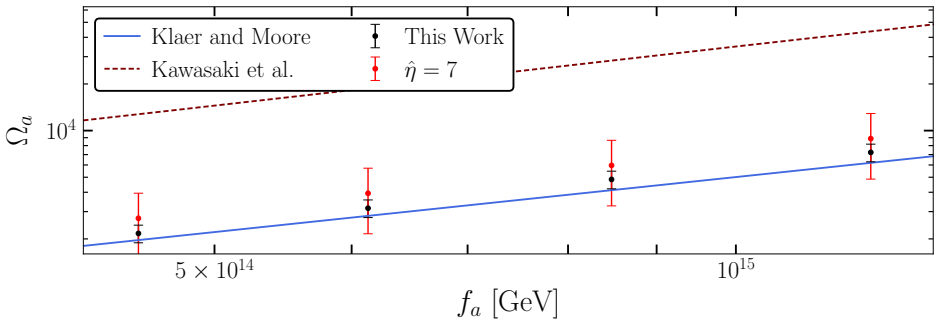



Figure A.4: A comparison of the predictions for the relic abundance of axions dark matter as a function of $f_a$ obtained in [118] (Kawasaki et al.) and [112] (Klaer and Moore) with the simulation results realized in this work. Error bars are combined statistical and correlated systematic errors, with the former dominating at $\hat{\eta} = 7$ due to large field gradients and the latter at $\hat{\eta} = \hat{\eta}_{MR}$.

$N_{\rm eff}$, which we have fixed to the Standard Model value. These cosmological uncertainties introduce an approximately $3\%$ correlated uncertainty across the results of our simulations. We additionally have an approximately $8\%$ uncertainty due to our assumption of fixed $g_*$, which is examined in greater detail later. These uncertainties are the dominant ones in our results, and we emphasize that they have not been typically considered in determinations of the DM axion mass. From the $\Omega_a$ data, for the various $f_a$ values simulated, we may extrapolate to predict the $f_a$ for an axion which produces the observed DM relic abundance by fitting a simple power law relation of the form

$$\Omega_a(f_a) = c_1 \cdot f_a^{\alpha}\,, \tag{A.31}$$

as discussed in the main body of this work. Note that we expect $\alpha = (6+n)/(4+n)$, where $n$ is the index of the axion mass growth. We assume this scaling is valid to make our estimate for the $m_a$ that gives the correct DM abundance. The relation between $\alpha$ and $n$ is expected to arise for the following reason. Let us estimate the axion DM density from an axion with a constant initial misalignment angle $\theta_i$. The present-day axion abundance as produced by the misalignment mechanism can be estimated by

$$\rho_a(T_0) = \rho_a(T_3)\frac{m_a(T_0)}{m_a(T_3)}\frac{g_*(T_0)T_0^3}{g_*(T_3)T_3^3}\,, \tag{A.32}$$

where $T_0$ is the present-day temperature, $T_3$ is the temperature at which the axion began to oscillate ($m_a(T_3) = 3H(T_3)$), and $g_*(T)$ the number of effective degrees of freedom at temperature $T$.

The initial axion abundance $\rho_a(T_3)$ is given

$$\rho_a(T_3) = \frac{m_a(T_3)^2 f_a^2}{2}\theta_i^2\,, \tag{A.33}$$

Anharmonicity factors can be included, but have no temperature or $f_a$ dependence. The temperature $T_3$ depends on $f_a$ through the relation $T_3 \propto f_a^{-2/(4+n)}$. Substituting these relations in and keeping only terms which depend on $f_a$, we have

$$\rho_a(T_0) \propto f_a^{(6+n)/(4+n)} \frac{g_*(T_0)}{g_*(T_3)}\,. \tag{A.34}$$

We thus expect the relic abundance to scale with $f_a$ like $\rho_a \propto f_a^{(6+n)/(4+n)}$. Note that the DM abundance from string and domain wall production is calculated similarly in [118], and although our results are not consistent with those presented in that work, the abundance calculation they present proceeds similarly, yielding string and domain wall production that scale like $f_a^{(6+n)/(4+n)}$ as well.

On the other hand, we may also calculate the the $m_a$ that gives the correct DM abundance by using our fit value for $\alpha$, as defined in (A.31), instead of the theoretical value. Doing so leads to a slightly lower $m_a$ estimate, as described in the main text.

### A.2.3 Tests of the Overdensity Field Gaussianity

In typical cosmological contexts, overdensity fields are treated under the assumption that they are Gaussian random fields. For a real-space Gaussian field, we may Fourier transform the field and find that the squared magnitude of each mode is independently exponentially distributed with mean set by the power-spectrum and with the phase of each mode independently uniformly distributed on $[0, 2\pi)$ [429]. For reference, in Fig. A.5 we show our power spectra $\Delta_k^2$ at fixed $\tilde{\lambda}$ across our various choices for $\hat{\eta}_c$. Note that we construct the power spectra from the fields that have been evolved until $\hat{\eta} = \hat{\eta}_{\rm MR}$. However, as we demonstrate below, the power spectrum fails to accurately describe the overdensity field we realize in our simulations because the field is highly non-Gaussian at small scales. As a result, standard tools for predicting structure formation that rely upon an underlying Gaussian overdensity field, such as the Press-Schechter formalism, cannot be applied to predict the spectrum of structures that form from the overdensities in the axion field, at least on the very smallest scales.

First, we note that the largest field values taken within the overdensity fields at the state realized by the analytic evolution until $\hat{\eta} = \hat{\eta}_{\rm MR}$ are $\mathcal{O}(10)$, whereas the minimum value the overdensity field can take is $-1$ by construction. This is trivially incompatible with the interpretation of the

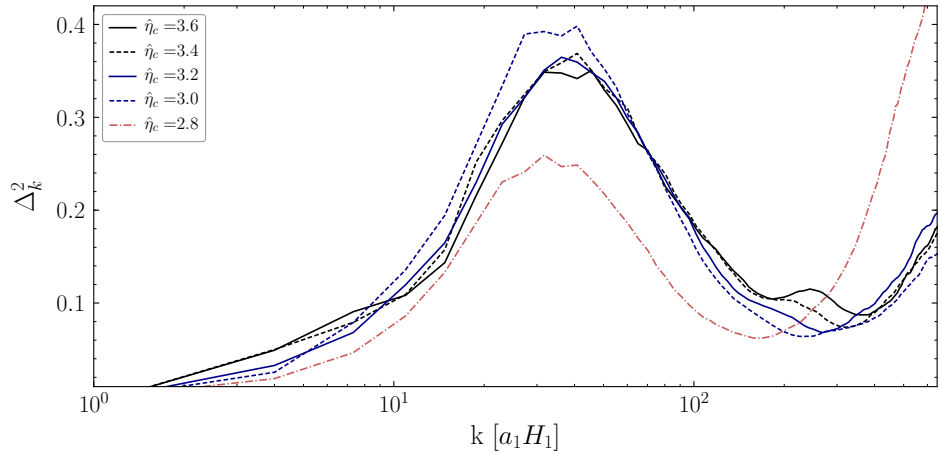



Figure A.5: A comparison of the power spectra realized in simulations for $\tilde{\lambda} = 5504$ for different choices of $\hat{\eta}_c$. New features in the power spectrum emerge as we push to larger values of $\hat{\eta}_c$, and we cannot exclude the possibility that further features would emerge were we to simulate with a greater value for the cutoff. On the other hand, the power-spectrum is highly non-Gaussian at small scales, so the distribution $\Delta_k^2$ alone is not adequate for understanding the small-scale nature of the overdensity field.

overdensity field as a Gaussian random field, which would have symmetric variance about its mean of $0$. For our overdensity fields to realize $\mathcal{O}(10)$ maxima with $-1$ as a construction-imposed minimum, there must exist considerable phase-correlations between Fourier modes, contrary to the uncorrelated phases of a Gaussian random field.

We also may inspect the distribution of power at each mode in the Fourier transformed overdensity field. If the overdensity field were Gaussian, then the power in each mode would be exponentially distributed with mean set by the value of the mean power spectrum. To test this, we plot the probability distribution $dP/dx$ of $x = |\hat{\delta}(\mathbf{k})|^2/\langle|\hat{\delta}(\mathbf{k})|^2\rangle_{|\mathbf{k}|=k}$, with $\hat{\delta}(\mathbf{k})$ the Fourier transformed overdensity field at momentum $\mathbf{k}$, as measured in the final states of our field at $\hat{\eta} = \hat{\eta}_{\rm MR}$. We compare the observed distributions with the expected Gaussian random field assumption of an exponential distribution with unit mean in Fig. A.6. Dramatic deviations from the expected behavior are observed for large $|\mathbf{k}|$. We stress, however, that in addition to these distributions departing from the expected exponential distributions, the real and imaginary components across modes are also highly phase correlated on small scales. We additionally note that this study of the Fourier spectrum of the overdensity field is performed after the field has been undergone an eight-fold down-binning, so that all modes inspected are well above the simulated lattice-spacing scale. Moreover, the oscillons which are obvious indicators of non-Gaussianity, are resolved by several lattice spacings at all times in our simulation. In this way, we can be confident that the

appearance of non-Gaussianity is a physical effect rather than one associated with discretization error. As mentioned in Sec. A.2.1, we have also explicitly verified in the context of our 2D simulations that increasing the number of grid sites leads to consistent results, further indicating that the non-Gaussian density spectra are physical and not related to the finite lattice spacing.

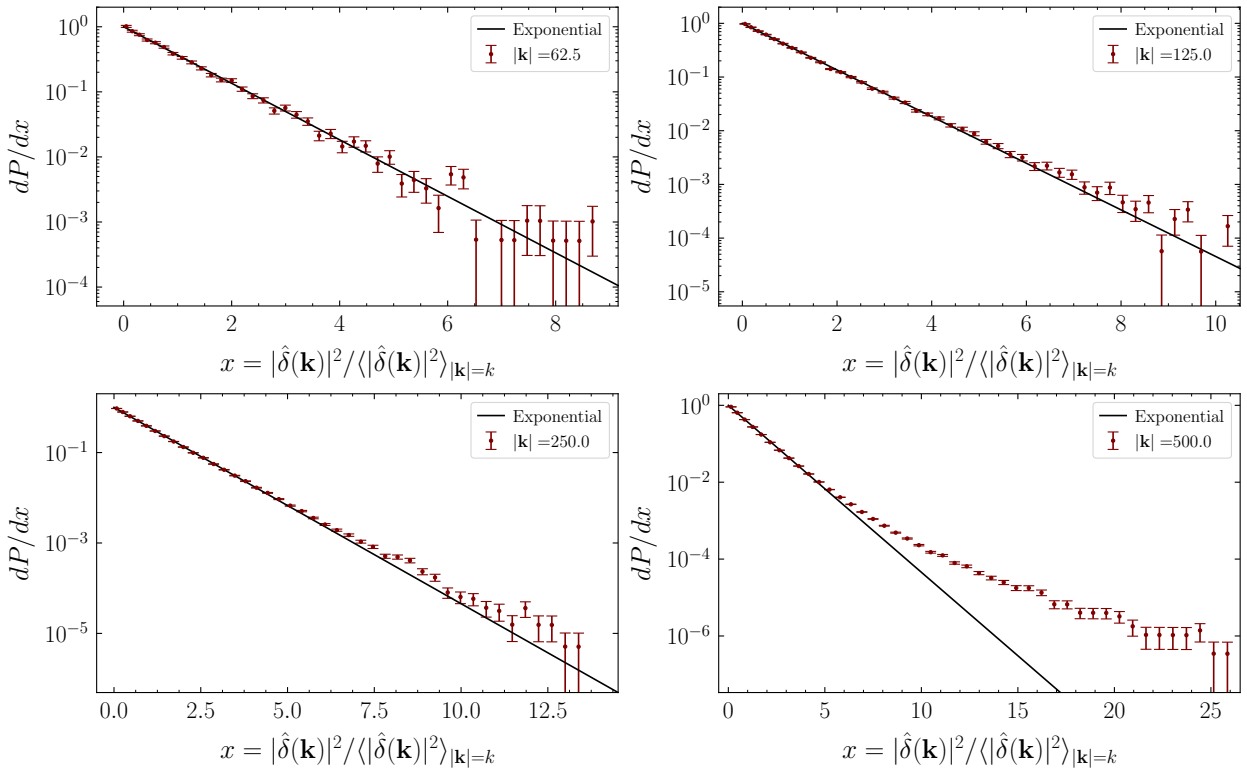



Figure A.6: A comparison of the distribution of the squared magnitudes of Fourier components for four different fixed reference momentum $k$. The expected exponential distribution for a Gaussian field is also indicated. While the distributions are Gaussian at large scales, they become increasingly non-Gaussian at small scales. The momentum mode $|\mathbf{k}| = 500$ corresponds to approximately 6.5 grid sites. These distributions were constructed from our most realistic simulation with $\tilde{\lambda} = 5504$ and $\eta_c = 3.6$.

### A.2.4 Minihalo Mass Spectrum

In this subsection we give additional details and results for the minihalo mass and density spectrum. In addition to the technical difficulties associated with a non-Gaussian overdensity field, computational limitations prevent us from performing realistic simulations of $f_a \sim 10^{11}\,\mathrm{GeV}$ axions, which would require us to simulate until $\hat{\eta}_c \approx 15$. We instead interpret our simulation results at smaller $\hat{\eta}_c$ in appropriate units to rescale these results to the target $f_a \approx 2 \times 10^{11}\,\mathrm{GeV}$. We do so with the following methods. The total axion mass contained within some set of grid-sites in our

simulation can be computed from the Hamiltonian as

$$M_{\rm tot} = a(\hat{\eta})^3 \int d^3x \mathcal{H} \approx a(\hat{\eta}) \sum (\Delta x)^3 \mathcal{H} = \left(\frac{a(\hat{\eta})^3 \Delta \bar{x}}{a_1 H_1}\right)^3 \sum \mathcal{H} = \left(\frac{\hat{\eta} \Delta \bar{x}}{H_1}\right)^3 \sum (1+\delta)\bar{\rho}\,, \tag{A.35}$$

where $\bar{\rho}$ is computed by the average of our Hamiltonian in (A.30) in the simulation box. We calculate $H_1$ from $T_1$ based on our choice of $f_a$, then rescale $\bar{\rho}$ to the value of the axion energy density at the time $\hat{\eta}$ such that the correct relic abundance is realized today. In this manner, we aim to rescale all dimension-full quantities related to $f_a$ to our target $f_a$. In particular, we rescale the DM density $\bar{\rho}$ to give the correct DM density realized in our Universe, and we also rescale the minihalo masses by the factor $\propto (a_1 H_1)^{-3}$ appearing in (A.35) to those for the target $f_a$.

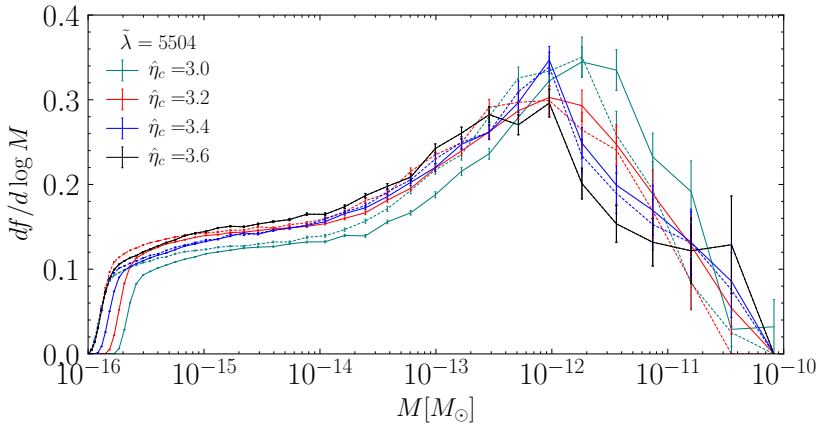



Figure A.7: Comparison between differential mass fractions as a function of the minihalo mass $M$ from our simulations at different $\hat{\eta}_c$. In this plot we have rescaled the minihalo masses such that we achieve the correct DM density $\bar{\rho}$ observed in the Universe, but for the solid curves we have not applied the Hubble volume rescaling factor to reach our target $f_a$. However, the dashed curves do have the Hubble volume rescaling factor included, but here we take our target $f_a$ to be that corresponding to our most realistic simulation with $\hat{\eta}_c = 3.6$. The difference between the dashed mass functions and the solid black mass functions gives a sense of the systematic uncertainty introduced by applying the naive mass rescaling factors instead of simulating with the correct value of $\hat{\eta}_c$ ($f_a$).

We illustrate the rescaling procedure in Fig. A.7. In that figure we show the differential mass distribution of minihalos $df/d\log M$ as a function of minihalo mass $M$. These mass distributions have been rescaled such that $\bar{\rho}$ matches the actual DM density. However, the solid curves do not have the $a_1 H_1$ Hubble volume rescaling included. The dashed curves, on the other hand, apply the Hubble volume rescaling factor but for a target $\hat{\eta}_c$ of $\hat{\eta}_c = 3.6$, which is that corresponding to the black curve. Clearly there are still differences between the rescaled dashed curves and the

black curve, which tells us that there are dynamical effects that arise from changing $\hat{\eta}_c$ that are not captured by the simple rescaling. This should not be too surprising considering that *e.g.* the mass growth affects the oscillon stability, which determines the high-mass part of the distribution. In our work we rescale the mass function to the target $f_a$ as described above, but it is important to keep in mind that this almost certainly results in a systematic uncertainty from the fact that we do not capture the full oscillon dynamical range in doing so. Also note that all of the mass functions abruptly drop off at low halo masses. This is due to our resolution limit on the finite lattice. We also cannot rule out the possibility that the low-mass tail continues down to much smaller masses.

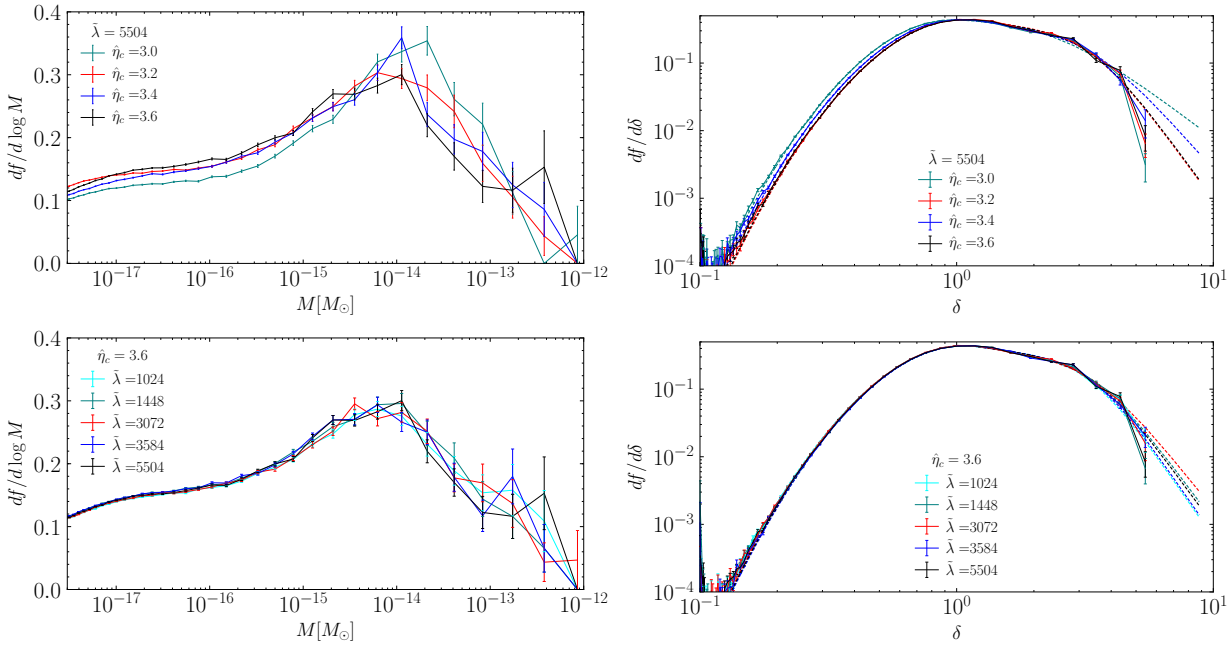



Figure A.8: Comparison between differential mass fractions as a function of the concentration parameter $\delta$ and minihalo mass $M$ for different $\hat{\eta}_c$ and $\tilde{\lambda}$ at $\hat{\eta} = \hat{\eta}_{\rm MR}$. Error bars are statistical. Shown as dotted lines is a fit to the $df/d\delta$ curves as described in the text. We do not extend the $df/d\log M$ curves to lower masses as we are unable to resolve those properly.

We compare the single-differential mass fractions for different values of $\hat{\eta}_c$ and $\tilde{\lambda}$ as a function of $\delta$ and $M$ in Fig. A.8. Note that here we have applied the rescaling factors for the masses to our true target $f_a$, which is that which gives the correct DM density. First of all we note that there is no dependence on $\tilde{\lambda}$ visible in our parameter range within than statistical scatter. As for the differential distribution as a function of $\delta$, there is also no clear dependence on $\hat{\eta}_c$ visible. The only place where a clear dependence on $\hat{\eta}_c$ is visible is in the mass fraction as a function of $M$. Here, the peak values shift to smaller masses upon increasing $\hat{\eta}_c$, even after including the rescaling factors.

It is useful to have an approximate analytic formula for the differential mass fraction. We find that the differential mass fraction as function of $\delta$ can be accurately described by a Crystal Ball

function based on a generalized Gaussian and a power-law high-end tail together with a suppression factor at high-$\delta$:

$$\frac{df}{d\delta} = \frac{A}{1+\left(\frac{\delta}{\delta_F}\right)^S} \begin{cases} e^{-\left[\ln\left(\frac{\delta}{\delta_G}\right)/\sqrt{2}\sigma\right]^d} & \text{for } \ln\left(\frac{\delta}{\delta_G}\right) \leq \sigma\alpha \\ B\left[C + \frac{1}{\sigma}\ln\left(\frac{\delta}{\delta_G}\right)\right]^{-n} & \text{for } \ln\left(\frac{\delta}{\delta_G}\right) > \sigma\alpha \end{cases} . \tag{A.36}$$

The parameters $B$ and $C$ are given by

$$B = e^{-\left(\frac{|\alpha|}{\sqrt{2}}\right)}\left[\left(\frac{\sqrt{2}}{|\alpha|}\right)^d \frac{|\alpha|n}{d}\right]^n , \qquad C = |\alpha|\left[\left(\frac{\sqrt{2}}{|\alpha|}\right)^d \frac{n}{d} - 1\right] , \tag{A.37}$$

and they are chosen such that $df/d\delta$ and its first derivative are continuous. $A$ is not a free parameter as $\int_0^\infty d\delta(df/d\delta) = 1$ must hold. The fit parameters from our most realistic simulation with $\hat{\eta}_c = 3.6$ and $\tilde{\lambda} = 5504$ are given by

$\sigma = 0.448 \pm 0.008$ $\quad n = 115 \pm 8$ $\quad \delta_G = 1.06 \pm 0.02$ $\quad S = 4.7 \pm 1.6$
$d = 1.93 \pm 0.02$ $\quad \alpha = -0.21 \pm 0.07$ $\quad \delta_F = 3.4 \pm 1.2$ .

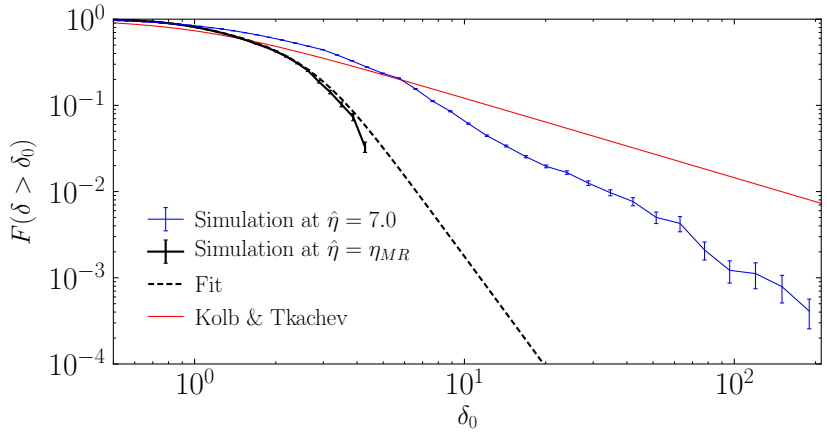



Figure A.9: Comparison between cumulative mass fractions, defined in the text, for our simulation at $\hat{\eta} = 7$ (solid blue) and $\hat{\eta}_{\rm MR}$ (solid black). We use our fit to the differential mass fraction $df/d\delta$ to extrapolate to high $\delta_0$ for our $\hat{\eta}_{\rm MR}$ data (dotted black). Error bars are statistical. We compare our results to those from Kolb and Tkachev [430] obtained at $\hat{\eta} = 4$ by using the fit to their data presented in [99] (red curve).

This fit allows us to make a precise comparison with previous work by Kolb and Tkachev [430].

We present in Fig. A.9 the cumulative mass fraction that is in overdensities larger than $\delta_0$,

$$F(\delta > \delta_0) = \int_{\delta_0}^{\infty} \frac{df}{d\delta} d\delta. \tag{A.38}$$

Unsurprisingly, we find considerably less mass in highly concentrated overdensities relative to [430]. Whereas [430] predicts roughly $10\%$ of the mass is in overdensities with $\delta = 10$ or more, we find a similar result only when using the simulation output at $\hat{\eta} = 7$. Once evolved to matter-radiation equality, that percentage falls to $\sim 0.1\%$.

## A.3 Testing the QCD Potential

The QCD potential may be modified in order to obtain more numerically tractable equations of motion. Our effective Lagrangian

$$\mathcal{L}_{PQ} = \frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}\left(|\Phi|^2 - f_a^2\right)^2 - \frac{\lambda T^2}{6}|\Phi|^2 - m_a(T)^2 f_a^2[1 - \cos \mathrm{Arg}(\Phi)], \tag{A.39}$$

has an apparent singularity when $\Phi \to 0$, which may happen in the string cores, since $\mathrm{Arg}(0)$ is not well defined. An alternate form of the potential has been proposed in the literature (see, *e.g.*, [121]) to mitigate this apparent singularity and make the equations of motion numerically more tractable:

$$\mathcal{L}_{PQ} = \frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}\left(|\Phi|^2 - f_a^2\right)^2 - \frac{\lambda T^2}{6}|\Phi|^2 - m_a(T)^2 f_a^2[1 - |\Phi| \cos \mathrm{Arg}(\Phi)]. \tag{A.40}$$

This leads to equations of motion for early times in the QCD epoch of the form

$$\begin{aligned} \psi_1'' + \frac{2}{\hat{\eta}}\psi_1' - \bar{\nabla}^2\psi_1 + \tilde{\lambda}\hat{\eta}^2\psi_1(\psi_1^2 + \psi_2^2 - 1) - \min(\hat{\eta}, \hat{\eta}_c)^n\hat{\eta}^2 = 0 \\ \psi_2'' + \frac{2}{\hat{\eta}}\psi_2' - \bar{\nabla}^2\psi_2 + \tilde{\lambda}\hat{\eta}^2\psi_2(\psi_1^2 + \psi_2^2 - 1) = 0\,, \end{aligned} \tag{A.41}$$

in analogy to those found in (A.20). The equation of motion for the single axion field appropriate for late times in the QCD epoch are unchanged.

Equipped with these new equations of motion, we re-simulate in three spatial dimensions using our fiducial values of $\eta_c = 3.6$ and $\tilde{\lambda} = 5504$. A comparison of results across metrics, such as the energy density as measured at matter-radiation equality, leads to a $\sim 1\%$ difference between results obtained with the two potentials. We neglect this systematic in our error budget as it is vastly subdominant to other sources of uncertainty.

## A.4 Testing the Impact of the Mass Parametrization

Precise details regarding the temperature dependence of the axion mass remain uncertain. While we have chosen to use the parametrization of [116] with index $n = 6.68$ as done in [117, 118], an more recent result is provided in [123]. In that work, an index of $n \approx 8.2$ is found at high temperatures, though we do note that an increasingly shallow dependence on $T$ is realized at lower temperatures. Motivated by power-law fits to this numerical result and informed by considerations of the changing number of degrees of freedom, recent works have taken an index of $n = 7.6$ in [112] and $n = 7.3$ in [96] to study the axion field. In this section, we use the extreme value $n = 8.2$ to estimate the maximal effect that uncertainties in the mass growth may have on the determination of the DM density.

We perform simulations first in two spatial dimensions, then in the full three spatial dimension using our initial state, now using $n = 8.2$. We fix $\tilde{\lambda} = 5504$ and $f_a \approx 4.8 \times 10^{14}\,\mathrm{GeV}$. This choice of $f_a$ corresponds to $\hat{\eta}_c = 3.6$ in the $n = 6.68$ parametrization. However, the value of $\hat{\eta}_c$ depends on our choice of $n$ and is $\hat{\eta}_c \approx 3.1$ for the choice of $n = 8.2$, since the mass grows faster in that case. We re-simulate with this alternative choice of index until $\hat{\eta} = 7$ and then recompute the present-day axion abundance by analytically transferring the simulation fields to the same late time physical temperature. In both 2D and 3D simulations, we find that there is a $\sim$10% enhancement in the expected relic abundance with $n = 8.2$ versus $n = 6.68$. This is somewhat surprising, considering that the analytic estimate predicts that higher $n$ should result in a lower DM abundance at fixed $f_a$. To understand how this result affects the final determination of the axion mass, we fit the predicted scaling $\Omega_a \sim f_a^{(n+6)/(n+4)}$ for the DM abundance using $n = 8.2$ and find the $m_a$ that gives the correct DM abundance.

The result is that with $n = 8.2$ we find that the $m_a$ that gives the correct DM abundance is enhanced by $\sim$27% compared to the $n = 6.68$ case. We account for this 27% uncertainty as an additional systematic uncertainty in our final determination of the axion mass. We also emphasize the consistency between results obtained for simulations with two spatial dimensions as compared to those with three spatial dimensions. This is an important result which will serve as evidence in the following section examining the impact of the evolution of the number of relativistic degrees of freedom during the QCD epoch.

## A.5 Equations of Motion for Varying Relativistic Degrees of Freedom

In this section we investigate the systematic effect on our results from the assumption of fixed $g_*$. In truth, the value of $g_*$ is not fixed at $g_* \approx 81$ but instead evolves across the temperature range of

interest; it varies from as large as roughly $100$ to as little as roughly $10$ after the QCD phase transition. This does not represent a dire shortcoming of our simulation procedure, however, as varying $g_*$ should only nontrivially affect the dynamics of the axion field during times when axion number density is not a conserved quantity. By $\hat{\eta} \approx 3$ most of the field has become linear, except for the isolated oscillon configurations, which means that the axion number density is mostly conserved at this time and beyond. The variation in $g_*$ before $\hat{\eta} \approx 3$ for our target $f_a$ is relatively minor. To quantify this impact, however, we perform simulations in two spatial dimensions accommodating the varying $g_*$.

With a change of variable we may rewrite the axion equations of motion, in the two-field formalism during the QCD epoch, as

$$\begin{aligned}
\phi_1'' + \left(\frac{R_1\ddot{R}}{\dot{R}^2} + \frac{3}{\eta}\right)\phi_1' - \frac{R_1^2\dot{R}_1^2}{R^2\dot{R}^2}\nabla^2\phi_1 + \frac{\dot{R}_1^2}{\dot{R}^2}\left[\tilde{\lambda}\phi_1\left(\phi_1^2+\phi_2^2-1\right) - \frac{m_a(T)^2\phi_2^2}{H_1^2\left(\phi_1^2+\phi_2^2\right)^{3/2}}\right] = 0 \\
\phi_2'' + \left(\frac{R_1\ddot{R}}{\dot{R}^2} + \frac{3}{\eta}\right)\phi_2' - \frac{R_1^2\dot{R}_1^2}{R^2\dot{R}^2}\nabla^2\phi_2 + \frac{\dot{R}_1^2}{\dot{R}^2}\left[\tilde{\lambda}\phi_2\left(\phi_1^2+\phi_2^2-1\right) + \frac{m_a(T)^2\phi_1\phi_2}{H_1^2\left(\phi_1^2+\phi_2^2\right)^{3/2}}\right] = 0\,,
\end{aligned} \tag{A.42}$$

where we define $\tilde{\lambda} = \lambda f_a^2/H_1^2$ as before. Citing standard references [17], we have

$$H \approx 1.660 g_*(T)^{1/2}\frac{T^2}{m_{Pl}} \tag{A.43}$$

$$t \approx 0.3012 g_*^{-1/2}\frac{m_{Pl}}{T^2} \tag{A.44}$$

$$R \approx 3.699\times 10^{-10} g_*(T)^{-1/3}\frac{\mathrm{MeV}}{T}\,. \tag{A.45}$$

Using these relations, we may compute

$$\frac{R_1\ddot{R}}{\dot{R}^2} = \frac{\left(\frac{t_1}{t}\right)^{1/2}\left(\frac{g(t)}{g(t_1)}\right)^{1/12}\left(13t^2\dot{g}(t)^2 - 12tg(t)\left(t\ddot{g}(t)+\dot{g}(t)\right) - 36g(t)^2\right)}{\left(t\dot{g}(t)-6g(t)\right)^2} \tag{A.46}$$

$$= -\frac{1}{\hat{\eta}}\left[\frac{-13t^2\dot{g}(t)^2 + 12tg(t)\left(t\ddot{g}(t)+\dot{g}(t)\right) + 36g(t)^2}{\left(t\dot{g}(t)-6g(t)\right)^2}\right] \tag{A.47}$$

$$= -\frac{f_1(\hat{\eta})}{\hat{\eta}}\,. \tag{A.48}$$

Above, we have defined

$$f_1(\hat{\eta}) = \frac{-13t^2\dot{g}(t)^2 + 12tg(t)\left(t\ddot{g}(t)+\dot{g}(t)\right) + 36g(t)^2}{\left(t\dot{g}(t)-6g(t)\right)^2}\,, \tag{A.49}$$

where the right hand side is evaluated at the time $t$ corresponding to the conformal time $\hat{\eta}$. Simi-

larly, we evaluate

$$\frac{R_1^2\dot{R}_1^2}{R^2\dot{R}^2} = f_2(\hat{\eta}), \qquad \frac{\dot{R}_1^2}{\dot{R}^2} = \hat{\eta}^2 f_2(\hat{\eta})\,, \tag{A.50}$$

for

$$f_2(\hat{\eta}) = \frac{\left(\frac{g(t)}{g(t_1)}\right)^{7/3}\left(t_1\dot{g}\left(t_1\right) - 6g\left(t_1\right)\right)^2}{\left(t\dot{g}(t) - 6g(t)\right)^2}\,. \tag{A.51}$$

Finally, we define $f_3(\hat{\eta}) = m_a(\hat{\eta}(T))^2/H_1^2$. Combining these results, the equations of motion take the form

$$\begin{aligned}
&\phi_1'' + \left(\frac{3}{\hat{\eta}} - \frac{f_1(\hat{\eta})}{\hat{\eta}}\right)\phi_1' - f_2(\hat{\eta})\nabla^2\phi_1 + \eta^2 f_2(\hat{\eta})\left[\tilde{\lambda}\phi_1\left(\phi_1^2+\phi_2^2-1\right) - \frac{f_3(\hat{\eta})\phi_2^2}{\left(\phi_1^2+\phi_2^2\right)^{3/2}}\right] = 0\\
&\phi_2'' + \left(\frac{3}{\hat{\eta}} - \frac{f_1(\hat{\eta})}{\hat{\eta}}\right)\phi_2' - f_2(\hat{\eta})\nabla^2\phi_2 + \eta^2 f_2(\hat{\eta})\left[\tilde{\lambda}\phi_2\left(\phi_1^2+\phi_2^2-1\right) + \frac{f_3(\hat{\eta})\phi_1\phi_2}{\left(\phi_1^2+\phi_2^2\right)^{3/2}}\right] = 0.
\end{aligned} \tag{A.52}$$

In the single-field formalism, these results are analogously applied to obtain

$$\theta'' + \left(\frac{3}{\hat{\eta}} - \frac{f_1(\hat{\eta})}{\hat{\eta}}\right)\theta' - f_2(\hat{\eta})\bar{\nabla}^2\theta + \eta^2 f_2(\hat{\eta}) f_3(\hat{\eta})\sin\theta = 0\,. \tag{A.53}$$

In Fig. A.10 we show the functions $f_1$, $f_2$, and $f_3$ entering into the equations of motion as functions of $\hat{\eta}$. Note that for $f_3$, we normalize against $\tilde{f}_3(\hat{\eta})$, which we define to be $f_3$ but with a fixed $g_*$. In the absence of a varying $g_*$, all of the curves appearing in Fig. A.10 would be identically one.

To test the impact of an evolving $g_*$, we adopt the parametrization of $g_*$ from [116] and simulate, in two spatial dimensions, for an axion with $f_a = 4.83 \times 10^{15}\,\mathrm{GeV}$. When we assumed $g_*$ was constant, this axion reached its zero-temperature mass at $\hat{\eta} = 3.6$, but accounting for the changing value of $g_*$, the axion now reaches its zero-temperature mass at $\hat{\eta} \approx 5.5$. As before, we conclude our simulation at $\hat{\eta} = 7$, and we calculate a relic abundance that is $7.7\%$ smaller than it is in the fixed $g_*$ case. We note that this scenario represents something of a worst-case scenario for the impact of $g_*$ on the dynamics because $g_*$ varies significantly during the epoch where axion number density is not conserved for this choice of $f_a$, and so we adopt this as a quantification of the systematic error associated with adopting a fixed $g_*$. For our target $f_a$, $g_*$ varies less, relative to the example illustrated, when the axion is in the non-linear regime and so we expect the effect of varying $g_*$ to be less important in this case.

We acknowledge that the estimates above were performed in 2D and not in 3D and that physical results may scale differentially between simulations in different dimensions (for example, oscillons are stable in 1D). (The simulations described in this section would be numerically intensive to perform in 3D.) However, one reason that the 2D results may be a reasonable proxy in this case

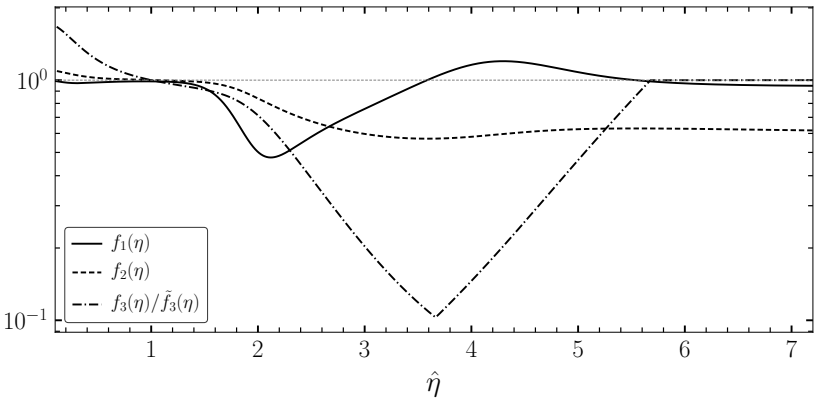



Figure A.10: We depict the variation of $f_1$, $f_2$, and $f_3$ as a function of $\hat{\eta}$ over the relevant range of $\hat{\eta}$ for our simulations accounting for a varying $g_*$. For fixed $g_*$, we would expect $f_1$ and $f_2$ to be constant at value 1. We additionally show the behavior of $f_3$, which describes the evolution of the quantity $m_a(\eta)^2/H_1^2$, normalized to $\tilde{f}_3$, wherein we compute $m_a(\eta)^2/H_1^2$ assuming a fixed $g_*$. Assuming a fixed $g_*$ causes the axion to reach its zero-temperature value earlier in $\hat{\eta}$, but the ratio ultimately reaches unity as the same zero-temperature mass is reached.

for what would happen in 3D is that the evolution of $g_*$ is largely degenerate with the temperature-dependent mass-growth, as pointed out in [96, 112] . In fact, prior works, such as [96, 112], modify their fiducial choice of mass-growth parameter $n$ in order to accommodate the evolution of $g_*$. We previously obtained consistent results across 2D and 3D simulations when studying the impact of changing $n$, which supports the argument that we may use the dimensional extrapolation to evaluate the impact of a varying $g_*$.

## A.6 String and Domain Wall Tensions and Defect Network Collapse

We observe the the defect network collapses early in the QCD epoch. It is important to verify that this collapse does not happen in an unphysical fashion (for example, happen too quickly or too slowly) because of our unphysical value of $\tilde{\lambda}$. Recall that $\tilde{\lambda}$ is proportional to the radio of the radial mode mass squared to the axion mass squared, and so physically this is expected to be an exponentially large number. On the other hand, our most realistic simulations use $\tilde{\lambda} = 5504$. In our simulations, we observe defect network collapse at $\hat{\eta} \approx 2$, indicating that the collapse occurs within a Hubble time. In [125] the criterion $m_s^2/m_a^2 \gtrsim 43$ was established as a necessary condition for defect network metastability, with $m_s$ the mass of the radial mode. If $m_s^2/m_a^2 \lesssim 43$, the QCD potential sufficiently distorts the radial mode potential so as to allow the complex scalar to roll

over the tilted Mexican hat potential, resulting in unphysical domain wall decays which would not be realized in the physical scenario where $(m_s/m_a)^2 \gg 1$. In our simulation with our choice of $\tilde{\lambda}$, we preserve $m_s^2/m_a^2 > 43$ until $\hat{\eta} = 2.3$, which supports the claim that the domain wall decays we observe in our simulations by $\eta = 2$ do not arise from an unphsysical mechanism.

Additionally, our fiducial choice of $\tilde{\lambda}$ preserves the correct timescale for defect network collapse. The linear string tension is given (see, *e.g.*, [125])

$$T \approx \pi f_a^2 \ln(m_s/H). \tag{A.54}$$

Restrictions in our dynamical range prevent us from simulating at large values of the radial model mass $m_s$, and thus, although our interest is in $f_a \approx 10^{11}\,\mathrm{GeV}$ resulting in $\ln(m_s/H) \approx 60 \sim 70$, we can only simulate at $\ln(m_s/H) \approx 4.5$. By comparison, the domain wall surface tension is [125, 431]

$$\sigma \approx 8 m_a f_a^2\,, \tag{A.55}$$

and at small values of $m_s$ a potential concern is that the domain wall tension plays an unphysically large role in driving the dynamics of the defect network collapse. As we will show in a toy model, however, this is likely not the case.

We consider the idealized scenario of a circular, horizon-scale domain wall at $\hat{\eta} = 1$. This is to say that the domain wall has surface area $\pi R^2$ for $R = H(\hat{\eta} = 1)$ in our simplified units. The domain wall will then end on a circular string with total length $2\pi R$. The potential energy of such a configuration is approximated by

$$V(R(\hat{\eta}), \hat{\eta}) = \pi^2 \ln(2\tilde{\lambda}\hat{\eta}^2) R + 8\pi \hat{\eta}^{n/2} R^2\,, \tag{A.56}$$

and the corresponding relativistic Lagrangian takes the form

$$L = V(R(\hat{\eta}), \hat{\eta}) \left[1 - \left(\frac{R'(\hat{\eta})}{\hat{\eta}}\right)^2\right]^{1/2}. \tag{A.57}$$

Using the Euler-Lagrange equations of motion, we numerically solve for $R(\hat{\eta})$ satisfying $R(1) = 1$ and $R'(1) = 0$. Results for the evolution are shown in Fig. A.11, where it can be seen that the evolution of the string-domain wall system results in a collapse time that is highly insensitive to $\tilde{\lambda}$. Explicitly, our simulated value of $\tilde{\lambda}$ gives consistent results with the physically-expected value of $\tilde{\lambda}$ for the collapse of the circular domain wall. Thus, we confirm that our choice of $\tilde{\lambda}$ is likely sufficient to realize the correct dynamics of defect collapse.

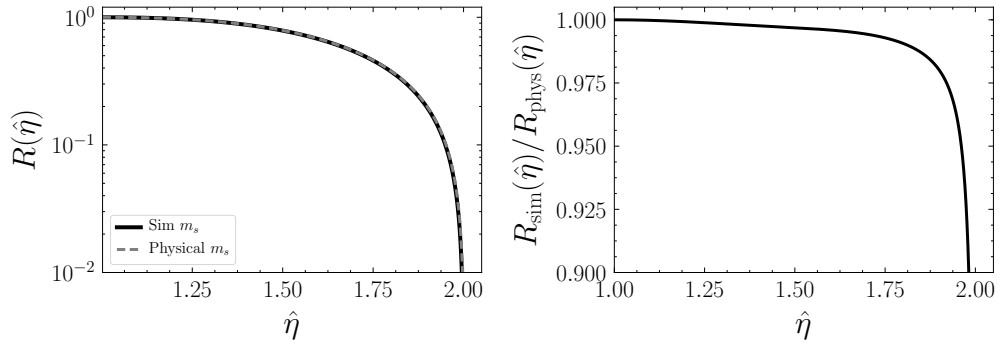



Figure A.11: Numerical evolution of the idealized circular string-domain wall collapse. (*Left*) A comparison of the radius of the circular domain wall as a function of conformal time $\hat{\eta}$ for the simulation parameter $\tilde{\lambda} = 5504$ (solid black) and for a physically-motivated parameter value $\tilde{\lambda} \approx 10^{30}$. The collapse of the domain wall occurs at around $\hat{\eta} = 2$, *i.e.*, after a Hubble time. (*Right*) The ratio of the domain wall radius as a function of $\hat{\eta}$ for the two values of $\tilde{\lambda}$. We see that the collapse rate is largely insensitive to the value of $\tilde{\lambda}$.

## A.7 Testing Deviations from the String Scaling Regime

While our simulation was performed in two stages, it can be understood as a single simulation in which the PQ phase transition and the beginning of the QCD phase transition are separated by approximately an order of magnitude in temperature. By comparison, for a physically motivated hierarchy, we would expect these two epochs in our simulation to be separated by at least 11 orders of magnitude in temperature. As a result, our simulations might be expected to be highly unphysical. However, it has been conjectured that the axion field and associated defect network enters a scaling regime some time after the PQ phase transition (see, *e.g.*, [117]). If this conjecture is true and our field configuration has entered the scaling regime before the axion begins to oscillate, our simulation should be expected to give a good description of the physics of interest despite the abbreviated hierarchy.

Recent work has found evidence for logarithmic deviations to the number of strings per Hubble patch in the scaling regime [124]. In this section we confirm that we also observe such deviations. This implies that we are not fully justified in taking the final state of our PQ-epoch simulation, fast-forwarding through the rest of the radiation dominated epoch to the QCD phase transition, and then restarting our simulation directly before the QCD phase transition. This is because the axion-string network should change logarithmically during the evolution between the phase transitions. Below, we provide evidence for the logarithmic deviation to scaling and then perform simulations to address the impact of this deviation on our determination of the axion mass $m_a$ and the spectrum of DM minihalos.

The average number of strings per Hubble patch is commonly defined by [117, 124, 125]

$$\xi(\tilde{\eta}) = \frac{l(\tilde{\eta})t(\tilde{\eta})^2}{L(\tilde{\eta})^3}\,, \tag{A.58}$$

where $l(\tilde{\eta})$, $L(\tilde{\eta})$, and $t(\tilde{\eta})$ are the physical total string length in the box, the physical box length, and the physical time, each a function of $\tilde{\eta}$, respectively. We measure the string length by first identifying grid sites that are next to a string. This is achieved by forming a loop in each of the three dimensions around a test grid site. The grid site is flagged once at least one change larger than $\pi$ in the axion field between consecutive grid sites is found. In a 2D slice this implies the 4 closest grid sites that surround the string core are tagged, such that we use the number of tagged grid sites divided by 4 as a measure for the string length. Note that this is a rough estimate for the string length and more sophisticated methods exist [117].

We compute $\xi(\tilde{\eta})$ at 13 points in $\tilde{\eta}$ in our PQ simulation, with results illustrated in Fig. A.12. As in [124], we find that $\xi$ depends logarithmically on $\tilde{\eta}$ after the PQ phase transition. Note that the shaded region denotes $\tilde{\eta}$ before the PQ phase transition, where it does not make sense to talk about axion strings. We fit the model

$$\xi = \alpha \log\left(\frac{T}{T_{PQ}}\right) + \beta \tag{A.59}$$

to the $\{\tilde{\eta}, \xi\}$ data, where $T_{PQ}$ denotes the temperature of the PQ phase transition, and we find $\alpha \approx -2.60$ and $\beta \approx 1.27$. Note that our values for $\xi$ at comparable $\tilde{\eta}$ are significantly larger than those found in [124]. Part of this discrepancy could be due to the way in which we measure string length versus in that work, which may introduce an overall rescaling between our two results. We are in good agreement, however, with [124] regarding the distribution of string length in long and short strings. As in that work, we find that approximately $80\%$ of the string length resides in long strings, much larger than a Hubble length, at all times in our simulation as seen in Fig. A.13.

Since we do observe logarithmic scaling violations, it is important to determine the impact of these corrections on the minihalo mass spectrum and the DM relic abundance. In particular, we find that $\xi$ should be around a factor of 15 higher at the QCD phase transition than it is for the final state of our most realistic PQ-epoch simulation. The string density $\xi$ at the beginning of the QCD simulation depends on the simulation box size as $\xi_{\rm QCD} \propto L_{\rm QCD}^{-2}$, where $L_{\rm QCD}$ is the box size in units of $1/(a_1 H_1)$ when $H = m_a$, for a fixed initial state. Thus by performing new simulations with the same initial conditions and run parameters as our fiducial analysis (namely $\hat{\eta}_c = 3.6, \tilde{\lambda} = 5504$, and starting at $\hat{\eta}_i = 0.4$) but changing $L_{\rm QCD}$ from 4 to $L_{\rm QCD} = 3$ and $L_{\rm QCD} = 2$, we may enhance $\xi_{\rm QCD}$ by a factor of $16/9$ and $4$, respectively, compared to our fiducial simulation. While these value still fall short of the physically motivated enhancement $\sim$15, such simulations still allow

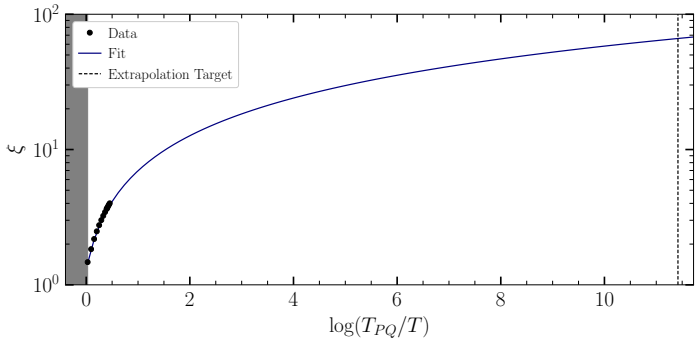



Figure A.12: The string length parameter $\xi$ shown as a function of the ratio of simulation temperature $T$ to the temperature $T_{PQ}$ at which the PQ phase transition occurred, including the results of our fit to the functional form of (A.58). We observe significant evidence for logarithmic deviation from the scaling regime. Extrapolating this result to the QCD phase transition (vertical dashed line) gives the prediction that $\xi$ should be around a factor of 15 higher at the beginning of the QCD phase transition than in the final state of our most realistic simulation.

us to see if there is a trend in how $\xi$ affects observables such as the DM density. We do caution that modifying $L_{\rm QCD}$ in this way is somewhat unphysical as it changes the horizon entry status of modes in the simulation box from the end of the PQ simulation as compared to the beginning of the QCD simulation.

The results of varying $L_{\rm QCD}$ in order to modify $\xi$ are shown in Fig. A.14, where we see no discernible trends in the dependence of the relic abundance on $\xi$. Note that the uncertainties in Fig. A.14 are estimates of the statistical uncertainty. As the box size gets smaller the statistical uncertainty increases. However, we caution that these are estimates only, as we have not run multiple independent simulations for each box size due to computational limitations. It is possible that the true uncertainties at small box sizes are larger than indicated due to the fact that there are a small number of *e.g.* domain walls that form these cases. Still, to be maximally conservative given the available datasets, we estimate the difference between the $L_{\rm QCD} = 2$ and $L_{\rm QCD} = 3$ values for $\Omega_a$ as a systematic uncertainty induced from the deviation to scaling. However, we cannot be sure that this difference is not a result of statistics or from the way in which we artificially mock up initial conditions with higher $\xi$ values. The systematic uncertainty we assign from these tests is $15\%$ correlated between different $f_a$ points.

We show in Fig. A.15 the impact on the minihalo mass spectrum. Again, all simulations largely agree within their error bands (estimated from statistical uncertainties), indicating that an increase in $\xi$ has only a marginal effect on the late-time axion field. Note that computational resources limit us to just these three additional simulations, and we leave a more detailed investigation of the dependence of $\Omega_a$ on $\xi$ to future work.

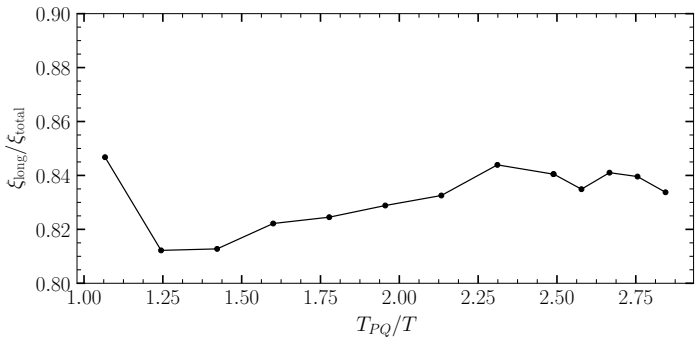



Figure A.13: The fraction of the string length in super-horizon length strings. Like [124], we find roughly $80\%$ of the string length resides within long strings.

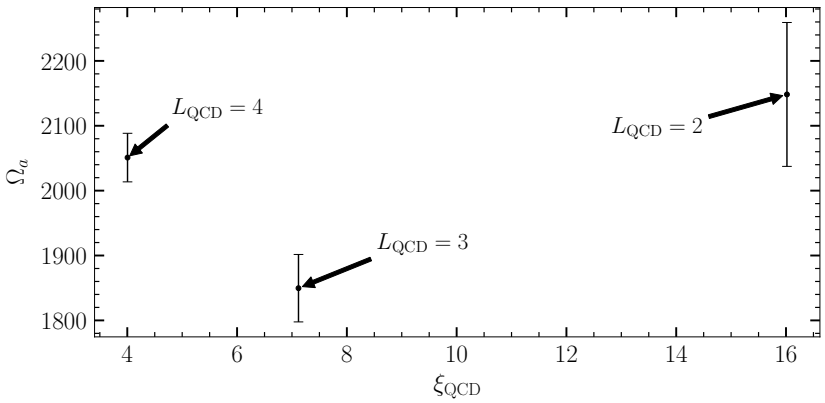



Figure A.14: The present day axion abundance as a function of the string density parameter $\xi$ at the beginning of the QCD simulation at $\hat{\eta}_i = 0.4$. Individual data points are labeled by their box length $L_{\rm QCD}$. The error bars are estimates of the statistical uncertainties, and no clear trend is visible in the data.

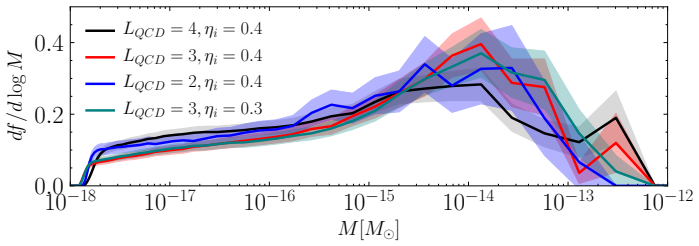



Figure A.15: Differential mass spectrum as a function of the minihalo mass $M$ for different box sizes. Error bands include statistical errors and the uncertainty on the overall normalisation.

# APPENDIX B

# The Statistics of Axion Direct Detection

## B.1 Distribution of the Combined Signal and Background Model

In Sec. 3.2 of the main text we demonstrated that the signal only distribution is exponentially distributed, as given in (3.24). However, we simply asserted that the background only and signal plus background distributions were also exponentially distributed. In this appendix we demonstrate both of these results. We reiterate at the outset that in all cases the correct starting point for determining these distributions is the time-series data, which is where the different contributions are combined. We cannot straightforwardly think about combining distributions at the level of the PSD. To emphasize this, even though the PSD in the background and signal only cases are individually exponentially distributed, the sum of two exponentially distributed numbers is not itself exponentially distributed, and yet the PSD formed from the sum of the background and signal is.

Consider firstly the background only distribution. Imagine we have time-series data collected in the presence of $n_B$ independent background sources, each Gaussian distributed random variables with mean zero and variance $\lambda_B^i/\Delta t$, where $i$ indexes the different backgrounds and the inclusion of $\Delta t$ in the variance is for later convenience. Note that we can choose the backgrounds to have zero mean without loss of generality, because the mean will only impact the $k = 0$ mode of the PSD, which for reasons described below we will not include in our likelihood. In the presence of this noise, the time-series data will take the form

$$\Phi_n = \sum_{j=1}^{n_B} x_n^j \,, \tag{B.1}$$

where $n = 0, 1, \ldots, N-1$ indexes the times at which the measurements were taken and the $x_n^i$

satisfy

$$\langle x_n^i \rangle = 0\,, \quad \langle x_n^j x_m^l \rangle = \delta_{nm}\delta_{jl}\frac{\lambda_B^j}{\Delta t}\,. \tag{B.2}$$

The second relation here follows as we assume our backgrounds are independent, and for a given background the values measured at different times are independent and identically distributed. Moving towards the PSD, consider the discrete Fourier transform of this data:

$$\Phi_k = \sum_{n=0}^{N-1} \Phi_n e^{-i2\pi kn/N} = \sum_{n=0}^{N-1}\sum_{j=1}^{n_B} x_n^j e^{-i2\pi kn/N}\,. \tag{B.3}$$

It is convenient to expand the exponential and analyze the real and imaginary parts of this separately. In detail:

$$\begin{aligned}\Phi_k =& \sum_{n=0}^{N-1}\sum_{j=1}^{n_B} x_n^j \cos\left(\frac{2\pi kn}{N}\right) \\ &-i\sum_{n=0}^{N-1}\sum_{j=1}^{n_B} x_n^j \sin\left(\frac{2\pi kn}{N}\right) \\ \equiv& R_k + iI_k\,.\end{aligned} \tag{B.4}$$

The real and imaginary parts, $R_k$ and $I_k$ respectively, are both Gaussian distributed since they are sums of Gaussian distributed random variables. Accordingly they are completely specified by their means and variances, which we can determine using (B.2). Consider the real part first, as the argument for the imaginary part proceeds in exactly the same fashion. For the mean we have

$$\begin{aligned}\langle R_k \rangle =& \left\langle \sum_{n=0}^{N-1}\sum_{j=1}^{n_B} x_n^j \cos\left(\frac{2\pi kn}{N}\right)\right\rangle \\ =& \sum_{n=0}^{N-1}\sum_{j=1}^{n_B} \langle x_n^j\rangle \cos\left(\frac{2\pi kn}{N}\right) \\ =& 0\,.\end{aligned} \tag{B.5}$$

Similarly

$$\begin{aligned}\langle R_k^2 \rangle =& \sum_{j=1}^{n_B}\frac{\lambda_B^j}{\Delta t}\sum_{n=0}^{N-1}\cos^2\left(\frac{2\pi kn}{N}\right) \\ =& \frac{\lambda_B}{\Delta t}\sum_{n=0}^{N-1}\cos^2\left(\frac{2\pi kn}{N}\right)\,.\end{aligned} \tag{B.6}$$

where we used $\lambda_B \equiv \sum_j \lambda_B^j$ following (3.25). We can evaluate the remaining sum using[1]

$$\sum_{n=0}^{N-1} \cos^2\left(\frac{2\pi k n}{N}\right) = \begin{cases} N & k = 0 \\ N/2 & 0 < k < N \end{cases} . \tag{B.7}$$

Putting these together, we conclude the real part has a variance given by

$$\langle R_k^2 \rangle = \begin{cases} \frac{\lambda_B N}{\Delta t} & k = 0 \\ \frac{\lambda_B N}{2\Delta t} & 0 < k < N \end{cases} . \tag{B.8}$$

The argument for the imaginary part is almost identical, and we find again that $\langle I_k \rangle = 0$, whilst

$$\langle I_k^2 \rangle = \begin{cases} 0 & k = 0 \\ \frac{\lambda_B N}{2\Delta t} & 0 < k < N \end{cases} . \tag{B.9}$$

Knowing how contributions to the Fourier transform are distributed, we now move to the PSD, which will again be a random variable given by:

$$S_{\Phi\Phi}^k = \frac{(\Delta t)^2}{T} |\Phi_k|^2 = \frac{\Delta t}{N} \left(R_k^2 + I_k^2\right) . \tag{B.10}$$

There are many ways to determine the probability density function (pdf) obeyed by $S_{\Phi\Phi}^k$. A particularly straightforward one in this case is to start by determining the cumulative distribution function (cdf), $F[S_{\Phi\Phi}^k]$. We will do this for $N > k > 0$ first, and return to the $k = 0$ case afterwards. To obtain the cdf, we simply integrate the distributions for $R_k$ and $I_k$ over all values up to some $S_{\Phi\Phi}^k$. In detail,

$$\begin{aligned} F[S_{\Phi\Phi}^k] = &\int^{S_{\Phi\Phi}^k} dR_k dI_k \, \frac{\Delta t}{\pi \lambda_B N} \\ &\times \exp\left[-\frac{\Delta t}{\lambda_B N}\left(R_k^2 + I_k^2\right)\right] . \end{aligned} \tag{B.11}$$

To perform this integral it is convenient to change to polar coordinates, $u^2 = R_k^2 + I_k^2$ and $\theta$, so that

$$\begin{aligned} F[S_{\Phi\Phi}^k] = &\int_0^{\sqrt{N S_{\Phi\Phi}^k/\Delta t}} du \, \frac{2\Delta t u}{\lambda_B N} \exp\left[-\frac{\Delta t u^2}{\lambda_B N}\right] \\ = &1 - e^{-S_{\Phi\Phi}^k/\lambda_B} . \end{aligned} \tag{B.12}$$

[1]Note that if $N$ is even, then for the $k = N/2$ mode the sum evaluates to the $k = 0$ result. This extends to (B.8) and (B.9), and indeed when propagated through to the likelihood, implies that this mode will also be gamma and not exponentially distributed.

The pdf is just the derivative of this, so we find

$$P[S^k_{\Phi\Phi}] = \frac{1}{\lambda_B} e^{-S^k_{\Phi\Phi}/\lambda_B} \,, \tag{B.13}$$

demonstrating that for $0 < k < N$ the background is exponentially distributed as claimed in the main body.

Consider next the case for $k = 0$. Utilizing an identical approach, we find firstly that

$$\begin{aligned} F[S^0_{\Phi\Phi}] &= \int_{-\sqrt{NS^k_{\Phi\Phi}/\Delta t}}^{\sqrt{NS^k_{\Phi\Phi}/\Delta t}} \sqrt{\frac{\Delta t}{\pi \lambda_B N}} \exp\left[-\frac{\Delta t}{\lambda_B N} R_0^2\right] \\ &= \mathrm{erf}\left[\sqrt{S^0_{\Phi\Phi}/\lambda_B}\right] \,, \end{aligned} \tag{B.14}$$

implying

$$P[S^0_{\Phi\Phi}] = \frac{1}{\sqrt{\pi \lambda_B S^0_{\Phi\Phi}}} e^{-S^0_{\Phi\Phi}/\lambda_B} \,. \tag{B.15}$$

Clearly the $k = 0$ mode is not exponentially distributed: it is in fact gamma distributed with shape parameter $1/2$ and scale parameter $\lambda_B$. In practice, however, this mode does not contribute to the likelihood function in (3.29) since all of the axions we search for have finite mass and thus finite oscillation frequency. Moreover, the $k = 0$ mode is degenerate with the mean background values that we have chosen to neglect.

Finally we want to show that the combined signal and background dataset is also exponentially distributed for $0 < k < N - 1$. We will show this in a somewhat indirect manner. Firstly, given that the signal is exponentially distributed, as shown in the main text, we will show that the real and imaginary parts of the discrete Fourier transform of such a dataset must be normally distributed. Then we can combine the signal in as if it was just another background in the argument presented above, and it will follow immediately that the full distribution must be exponential. Our starting point is (3.24), where we showed the signal only PSD is exponentially distributed. We repeat this result here for convenience:

$$\begin{aligned} P[S^k_{\Phi\Phi}] &= \frac{1}{\lambda_k} e^{-S^k_{\Phi\Phi}/\lambda_k} \,, \\ \lambda &\equiv A \left.\frac{\pi f(v)}{m_a v}\right|_{v=\sqrt{4\pi k/(m_a T)-2}} \,. \end{aligned} \tag{B.16}$$

As an intermediate step, consider $S^k_{\Phi\Phi} = x + y$, where $x = (\Delta t/N) R_k^2$ and $y = (\Delta t/N) I_k^2$. As the real and imaginary parts are independent and identically distributed for the signal dataset, then so too are $x$ and $y$, and we denote their pdf by $g$. Given that $x, y \geq 0$, we can relate their distributions

to that of the signal PSD via

$$\begin{aligned} P[S^k_{\Phi\Phi}] &= \int_0^\infty dxdy\, g[x]g[y]\delta(S^k_{\Phi\Phi} - x - y) \\ &= \int_0^{S^k_{\Phi\Phi}} dx\, g[x]g[S^k_{\Phi\Phi} - x]\,. \end{aligned} \tag{B.17}$$

To solve this equation for $g$ we take the Laplace transform, denoting transformed quantities with a tilde. This yields

$$\tilde{g}[\tilde{x}] = \frac{1}{\sqrt{1+\tilde{x}\lambda_k}}\,, \tag{B.18}$$

which when inverted becomes

$$g[x] = \frac{1}{\sqrt{\pi\lambda_k x}} e^{-x/\lambda_k}\,. \tag{B.19}$$

From here, to derive the pdf for $R_k$ we can change variables using $x = (\Delta t/N)R_k^2$. In doing so we need to account for the Jacobian and also the fact that whilst $x \in [0,\infty)$, this is only half the domain of possible $R_k$ values. Doing so we find

$$P[R_k] = \frac{1}{\sqrt{\pi N\lambda_k/\Delta t}} \exp\left[-\frac{R_k^2}{N\lambda_k/\Delta t}\right]\,, \tag{B.20}$$

which is exactly a normal distribution with mean zero and variance $N\lambda_k/(2\Delta t)$. The distribution for $I_k$ will be identical, and thus we find the signal is distributed just like a single background but with $\lambda_B^j \to \lambda_k$. If we then repeat the background only argument shown at the start of this appendix with the signal contribution added, we will find the full PSD is again exponentially distributed with mean $\lambda_k + \lambda_B$, completing the required derivation.

## B.2 Comparison to a Bandwidth Average

An alternative analysis strategy to that presented in the main text is to take the average PSD (or power) measured across a given bandwidth range and compare that directly to the average model prediction. This should be contrasted with taking the product of exponential likelihoods across $k$ modes as we introduced in (3.29), and at face value it should have less discriminating power as the information regarding how the axion signal is distributed within the bandwidth has been lost. In this section we quantify this statement by deriving the expected sensitivity of such an approach. As a side point we will also demonstrate how to derive the optimum bandwidth range in performing a bandwidth averaged search.

To begin with, we note that in each frequency bin the PSD formed from the data will still be exponentially distributed. Then, if we are searching in some bandwidth range $\Omega_\omega$, which contains

$n_\omega$ frequency bins, the mean PSD can be formed from a sum of these exponentials and will thus be Erlang distributed. In detail, the likelihood will have the form

$$\mathcal{L}(d|\boldsymbol{\theta}) = \frac{n_\omega^{n_\omega}}{(n_\omega - 1)!} \frac{\left(\bar{S}_{\Phi\Phi}\right)^{n_\omega - 1}}{\bar{\lambda}^{n_\omega}} e^{-n_\omega \bar{S}_{\Phi\Phi}/\bar{\lambda}}\,, \tag{B.21}$$

where we have defined:

$$\bar{S}_{\Phi\Phi} = \frac{1}{n_f} \sum_{k \in \Omega_\omega} S^k_{\Phi\Phi}\,, \tag{B.22}$$

similarly to what we had when discussing the stacked data procedure in Sec. 3.3.2. In the above equation we also introduced the mean model prediction, which assuming we have a frequency independent background will be given by

$$\begin{aligned} \bar{\lambda} =& \bar{\lambda}_S + \lambda_B\,, \\ \bar{\lambda}_S \equiv& \frac{1}{n_\omega} \sum_{k \in \Omega_\omega} A \frac{\pi f(v)}{m_a v}\,. \end{aligned} \tag{B.23}$$

Consider the average signal prediction. This average is taken over some frequency range, or bandwidth, which we denote by $\Delta\omega$, and is equivalent to a range in velocities, $v \in [0, v_{\rm max}]$.[2] Consequently we have

$$\Delta\omega = \frac{1}{2} m_a v_{\rm max}^2\,. \tag{B.24}$$

The bandwidth can also be written as $\Delta\omega = n_\omega d\omega$, where $d\omega$ is the width of an individual frequency bin. Assuming sufficient run time, as $d\omega = 2\pi/T$, then we can also write

$$\Delta\omega = n_\omega m_a v dv\,. \tag{B.25}$$

Taken together, these show that

$$\frac{\Delta\omega}{\Delta\omega} = \frac{2 n_\omega}{v_{\rm max}^2} v dv\,. \tag{B.26}$$

Substituting this into (B.23), we can rewrite the signal prediction as

$$\bar{\lambda}_S = \frac{2A\pi}{m_a v_{\rm max}^2} \int_0^{v_{\rm max}} dv\, f(v)\,. \tag{B.27}$$

To estimate the sensitivity it is most convenient to return to $\Theta$ as introduced in (3.39). This is

[2] In principle the lower velocity could be $v_{\rm min}$ rather than 0, and this value can also be optimized for. Nevertheless as the signal distribution rises sharply from $v = 0$, approximating $v_{\rm min} = 0$ is sufficient for the argument in this appendix.

modified for the averaged PSD likelihood given in (B.21) to

$$\Theta(A) = 2n_\omega \bar{S}_{\Phi\Phi} \left[\frac{1}{\lambda_B} - \frac{1}{\bar{\lambda}}\right] - 2n_\omega \ln \frac{\bar{\lambda}}{\lambda_B}\,, \tag{B.28}$$

where as in Sec. 3.3, we suppress the axion mass dependence. As in the main body, to analytically estimate the sensitivity we can use the Asimov dataset. Here we denote this by $\bar{\lambda}_S^t + \lambda_B$, where $\bar{\lambda}_S^t$ is identical to (B.27), but with the signal strength replaced by its true value: $A \to A_t$. To simplify the resulting form of $\tilde{\Theta}$, we again assume that we are in the limit where the true and modeled average signal strength are subdominant to the average background, such that we obtain

$$\tilde{\Theta}(A) = \frac{2AT\pi}{m_a \lambda_B^2} \left(A_t - \frac{A}{2}\right) \left(\int_0^{v_{\rm max}} dv\, \frac{f(v)}{v_{\rm max}}\right)^2 . \tag{B.29}$$

To compare this directly to results obtained from the analysis in the main body, we need to determine a value for $v_{\rm max}$. A procedure for doing so is to choose the $v_{\rm max}$ that maximizes the significance of any emerging signal, or in detail one that maximizes the test statistic of discovery. Using TS as defined in (3.38), for the present case we have

$$\widetilde{\rm TS} = \frac{A_t^2 T\pi}{m_a \lambda_B^2} \left(\int_0^{v_{\rm max}} dv\, \frac{f(v)}{v_{\rm max}}\right)^2 , \tag{B.30}$$

which we want to maximize as a function of $v_{\rm max}$. The value that does so depends critically on the form of $f(v)$, and so needs to be re-evaluated for each assumption. For example, if we take the simple SHM ansatz as per (3.9), then we find $v_{\rm max} \approx 453$ km/s. Using this value we can then construct the ratio between the TS using our default bin-by-bin approach, denoted $\rm TS^{full}$, to that obtained here, denoted $\rm TS^{av.}$, which is explicitly:

$$\frac{\widetilde{\rm TS}^{\rm full}}{\widetilde{\rm TS}^{\rm av.}} = \left(\frac{1}{2}\int dv \frac{f(v)^2}{v}\right) \left(\int_0^{v_{\rm max}} dv\, \frac{f(v)}{v_{\rm max}}\right)^{-2} \approx 1.14\,, \tag{B.31}$$

where in the final step we again assumed a default SHM form for the speed distribution. Thus as claimed at the outset, even when optimized, this averaging procedure is not as sensitive as our full construction. The optimization is important; if we had instead taken $v_{\rm max} = 300$ (600) km/s, we would have obtained a ratio of 1.87 (1.43) above. Further in the presence of substructure, the averaging approach suffers even further. As a simple estimate of this if we took Maxwellian substructure, with the much smaller velocity dispersion $v_0 = 10$ km/s but the same boost velocity as the SHM, then even at the maximum the ratio is 5.42.

Using this maximum we can also determine the impact on limits. Recalling the definition of the

test statistic for upper limits in (3.36), we find the condition for a 95% limit is determined when

$$\tilde{A}_{95\%} = A_t + \sqrt{2.71 \frac{m_a \lambda_B^2}{T\pi}} \left( \int_0^{v_{\rm max}} dv \, \frac{f(v)}{v_{\rm max}} \right)^{-1} . \tag{B.32}$$

To compare this to case discussed in the main body, we take the simplifying values of $A_t = 0$ and again the default SHM speed distribution. Doing so we find

$$\frac{\tilde{A}^{\rm full}_{95\%}}{\tilde{A}^{\rm av.}_{95\%}} = \left( \int_0^{v_{\rm max}} dv \, \frac{f(v)}{v_{\rm max}} \right) \left( \frac{1}{2} \int dv \, \frac{f(v)^2}{v} \right)^{-1/2} \approx 0.94 \, , \tag{B.33}$$

which corresponds to a ratio of the axion electromagnetic couplings of 0.97 ($A \propto g_{a\gamma\gamma}^2$). This value shows that the full framework sets similar, but slightly stronger, constraints.

Accordingly, in all cases the framework described in the main body outperforms the averaged-power technique described in this appendix. For the case of the SHM, when that technique is optimized the improvements are marginal. Nevertheless in the presence of substructure, or if the optimal signal window is not chosen, then the gain from resolving the individual frequency bins can be much more substantial. Moreover, it is very difficult to constrain aspects of the DM phase-space distribution with the power-averaged technique, since the frequency dependence of the signal is not resolved.

## B.3 Verifying the Asimov Derivation of Upper Limit Bands

Using the Asimov dataset analysis, in Sec. 3.3.3 we were able to calculate the expected 95% limit on the signal strength $A$ at a given $m_a$. We were also able to calculate the 1 and $2\sigma$ containment bands around the expected 95% limit without recourse to Monte Carlo simulations. In this appendix we confirm that these results, presented in (3.54) and (3.56), match those derived using Monte Carlo methods.

For this procedure, we generate 1000 background-only datasets over frequencies in a 22Hz window centered at 550kHz and then scan these PSDs for a bulk SHM model. According to our estimate in (3.62), we expect there to be approximately 55 independent mass points for which we can scan contained within this frequency data. However, for the sake of precision, we will arbitrarily increase our resolution to scan over 150 mass points, between which there may be some degeneracy. At each mass point, we scan over $A$ values between $-5\sigma_A$ and $10\sigma_A$ calculated according to (3.72). We emphasize again that it is necessary that we allow $A$ to take on negative values despite that, by its definition, $A$ must be nonnegative. In practice, this is resolved by imposing a power-constrained limit such that constraints on $A$ are placed no lower than $1\sigma$ below

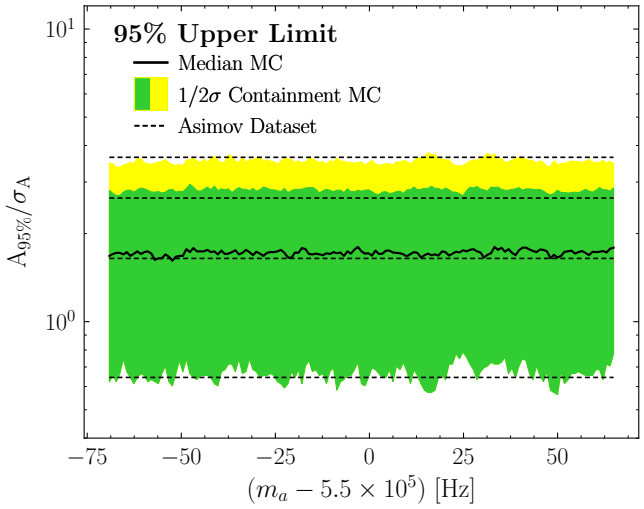



Figure B.1: A comparison between the variation in the 95% upper limit found in Monte Carlo (MC) simulations to that derived analytically with the Asimov dataset. As shown the two are in good agreement.

the expected constraint as calculated by (3.56). In Fig. B.1 we show the median 95% upper limit as well as the 1 (shaded green) and $2\sigma$ (shaded yellow) containment intervals constructed from the ensemble of Monte Carlo simulations. Note that we only show the upper $2\sigma$ region, since we anticipate neglecting fluctuations below $1\sigma$ with the power-constrained method. Additionally, we indicate the same quantities predicted by our Asimov analysis with dashed lines. As the figure demonstrates, the Monte Carlo and Asimov results are generally in good agreement.[3]

## B.4 Asymptotic Distribution for the Discovery Test Statistic

In this appendix we will explicitly calculate, from our likelihood, the survival function for the local TS under the null hypothesis. We will then show that asymptotically the TS is $\chi^2$-distributed, and therefore there is a simple connection with the significance, $Z$, given by $Z = \sqrt{\mathrm{TS}}$. Doing so will verify (3.58), presented in Sec. 3.3.4. Note that this appendix is in many ways an explicit illustration of Wilks' theorem.

To begin with, the situation to keep in mind is that we have a dataset that is drawn from the background only distribution, where in some frequency range there is an upward fluctuation that can be well described by a model including the signal. From this picture, in order to derive our result we will make two simplifying assumptions:

[3]While there may be a small systematic offset, as visible in Fig. B.1, the agreement is likely satisfactory for use at direct detection experiments. However, if required the containment intervals could be further tuned to agree with Monte Carlo simulations like those presented here.

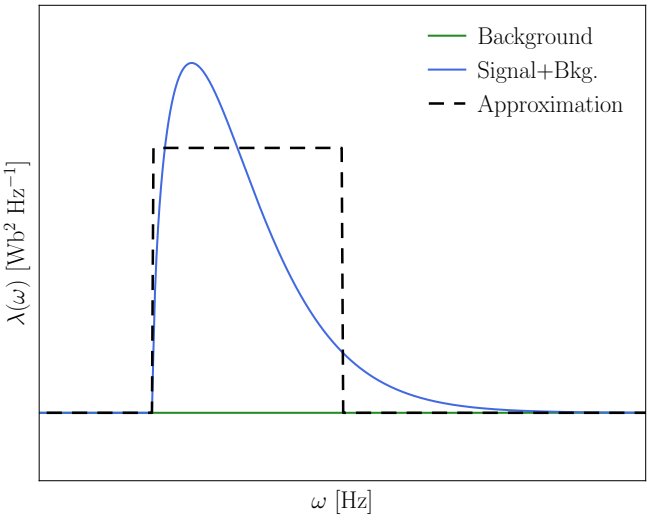



Figure B.2: Schematic depiction of the approximation made to the model used to derive $\mathrm{TS}_{\mathrm{thresh}}$. Specifically we assume that the signal model is non-zero only within a finite frequency range, and equal to the background outside this, and within this range the combined signal and background is flat.

1. that the signal model we use is only non-zero in a set of $n_S$ frequency bins, the set of which we denote $\Omega_S$, and outside this $\lambda_k = \lambda_B$; and

2. that in these $n_S$ bins the background and model predictions are both frequency independent, so to avoid confusion we denote our signal prediction in this range as the $k$-independent $\lambda_S$.

Taken together these assumptions imply we are approximating our model for this upward fluctuation in the background as a step function, similar to what is shown in Fig. B.2. In that figure, which is intended to be schematic, we have shown a flat background model, and added on top of this the signal distribution as expected from (3.24), and also shown the shape of the full model approximation we will use. Note that nothing in our first approximation or the derivation below requires $n_S \ll N$, however for this approximation to be realistic this will usually be the case.

Our aim now is to determine how the discovery test statistic is distributed under these assumptions. Combining these assumptions with the form of $\Theta$ given in (3.39), and then choosing the $A$ that maximizes this quantity, we arrive at:

$$\widetilde{\mathrm{TS}} = \begin{cases} 2n_S \left[ \frac{\bar{S}_{\Phi\Phi}}{\lambda_B} - 1 - \ln \frac{\bar{S}_{\Phi\Phi}}{\lambda_B} \right] & \bar{S}_{\Phi\Phi} > \lambda_B \,, \\ 0 & \bar{S}_{\Phi\Phi} \leq \lambda_B \,, \end{cases} \tag{B.34}$$

where we have defined the average data PSD in this range:

$$\bar{S}_{\Phi\Phi} \equiv \frac{1}{n_S} \sum_{k \in \Omega_S} S^k_{\Phi\Phi} \,. \tag{B.35}$$

Note that this should be distinguished from the subinterval averaged PSD in (3.50). Note also as written this result is independent of $m_a$, so we have suppressed the dependence on the mass.

Now recall that as each of our PSD measurements are exponentially distributed, the average PSD, $\bar{S}_{\Phi\Phi}$, will follow an Erlang distribution. In detail, we have

$$P[\bar{S}_{\Phi\Phi}] = \frac{n_S^{n_S}}{(n_S-1)!} \frac{\left(\bar{S}_{\Phi\Phi}\right)^{n_S-1}}{\lambda_B^{n_S}} e^{-n_S \bar{S}_{\Phi\Phi}/\lambda_B} \,. \tag{B.36}$$

We emphasize again that we are taking the data to follow the background distribution, as in calculating $\mathrm{TS}_{\mathrm{thresh}}$ we are interested in the distribution of the discovery test statistic under the null hypothesis. This explains why the mean in the above distribution is simply $\lambda_B$.

Now we want to use this to derive the distribution for $\widetilde{\mathrm{TS}}$. Before doing so, we need to take a brief aside. Observe that the distribution for the average PSD given in (B.36) is correctly normalized for $\bar{S}_{\Phi\Phi} \in [0, \infty)$. Nevertheless, from (B.34), we see that we only get a non-zero test statistic for $\bar{S}_{\Phi\Phi} > \lambda_B$, thus in the probability distribution for $\widetilde{\mathrm{TS}}$ there will be a pileup of probability at zero accounting for the fact that any time the average PSD is less than the background value, the maximum discovery test statistic will be zero. We can determine the probability of that occurring as:

$$\int_0^{\lambda_B} d\bar{S}_{\Phi\Phi}\, P[\bar{S}_{\Phi\Phi}] = 1 - \frac{\Gamma(n_S, n_S)}{(n_S-1)!} \,, \tag{B.37}$$

where $\Gamma(n_S, n_S)$ is the upper incomplete gamma function. Keeping this additional probability in mind, we determine the distribution for $\widetilde{\mathrm{TS}}$ from our distribution for $\bar{S}_{\Phi\Phi}$ via a change of variables. As an intermediate step, observe that we can invert that equation for $\bar{S}_{\Phi\Phi}$ in terms of TS using

$$\bar{S}_{\Phi\Phi} = -\lambda_B W_{-1}\left(-\exp\left[-1 - \frac{\widetilde{\mathrm{TS}}}{2n_S}\right]\right) \,, \tag{B.38}$$

where $W_{-1}$ is the lower branch of the Lambert $W$ function. This function provides an inverse to equations of the form $y = xe^x$, such that $x = W(y)$. As $W$ is multivalued, we choose the lower branch $W_{-1}$, where $W < -1$, which implies that $\bar{S}_{\Phi\Phi} \geq \lambda_B$. This shows that the change of variables will not cover the situation where the average PSD is less than the background, which

we account for using the result of (B.37). Using this change of variables, we then arrive at

$$P[\widetilde{\mathrm{TS}}] = \frac{n_S^{n_S}}{2n_S!}\frac{w^{n_S}e^{-n_S w}}{w-1} + \left[1 - \frac{\Gamma(n_S, n_S)}{(n_S-1)!}\right]\delta(\widetilde{\mathrm{TS}})\,,$$
$$w \equiv -W_{-1}\left(-\exp\left[-1 - \frac{\widetilde{\mathrm{TS}}}{2n_S}\right]\right)\,. \tag{B.39}$$

At this stage we can move to the asymptotic form of this result. To invoke Wilk's theorem, we need to take the large sample size limit. Here this is controlled by $n_S$, and so we take $n_S \to \infty$, and in particular $n_S \gg \mathrm{TS}$. Taking these limits and keeping just the leading term, we obtain

$$P[\widetilde{\mathrm{TS}}] = \frac{e^{-\widetilde{\mathrm{TS}}/2}}{\sqrt{8\pi\,\widetilde{\mathrm{TS}}}} + \frac{1}{2}\delta(\widetilde{\mathrm{TS}})\,. \tag{B.40}$$

This equation represents the asymptotic form of the discovery test statistic distribution under the background only hypothesis. We can now directly integrate this distribution to get the survival function, in detail to find the probability of a background fluctuation yielding a test statistic greater than some value:

$$\begin{aligned} S[\widetilde{\mathrm{TS}}] \equiv \int_{\widetilde{\mathrm{TS}}}^{\infty} d\widetilde{\mathrm{TS}}'\, P[\widetilde{\mathrm{TS}}'] =& \frac{1}{2}\mathrm{erfc}\left(\sqrt{\frac{\widetilde{\mathrm{TS}}}{2}}\right) \\ =& 1 - \Phi\left(\sqrt{\widetilde{\mathrm{TS}}}\right)\,, \end{aligned} \tag{B.41}$$

where $\mathrm{erfc}$ is the complementary error function and again $\Phi$ is a zero mean, unit variance Gaussian. This result verifies (3.58).

## B.5 Sensitivity Scaling for $T < \tau$

The main results from the Asimov dataset analysis performed in Sec. 3.3 demonstrated that our sensitivity increased with collection time as $T^{1/4}$, which is manifest in both (3.55) and (3.57). Nevertheless in deriving both of these results, we assumed that $T$ was large enough that frequency bins fully resolved variations in the signal; explicitly, we assumed that $T \gg \tau$, where $\tau$ represents the coherence time of the signal. This assumption was used in (3.43) so that we could rewrite the sum over frequency modes as an integral. As commented in Sec. 3.3.5, we would expect that for $T < \tau$ the sensitivity should instead scale as $T^{1/2}$ [58]. In this appendix we repeat our analysis, now assuming the collection time is less than the coherence time, and demonstrate we recover this scaling also.

To do so, we start with $\Theta$, from which we can derive 95% limits and the TS of an excess, as

described in Sec. 3.3. In particular, we begin with (3.41) which is the furthest we advanced in the Asimov analysis of $\Theta$ before invoking the assumption of $T \gg \tau$. Repeating that result for convenience, we have

$$\tilde{\Theta}(A) = 2\sum_{k=1}^{N-1}\left[\lambda_k^t\left(\frac{1}{\lambda_B}-\frac{1}{\lambda_k}\right)-\ln\frac{\lambda_k}{\lambda_B}\right], \tag{B.42}$$

where again $\lambda_k^t$ is the expected signal plus background, but with the signal set to its true value.

In the case where $T < \tau$, where we cannot resolve the signal, we can approximate it as being confined to a single $k$ mode, say $k = k_S$. We are effectively approximating $T \ll \tau$ here, much as we did $T \gg \tau$ in the main body, simply to expose the scaling. This allows us to rewrite the above as

$$\tilde{\Theta}(A) = 2\left[\lambda_{k_S}^t\left(\frac{1}{\lambda_B}-\frac{1}{\lambda_{k_S}}\right)-\ln\frac{\lambda_{k_S}}{\lambda_B}\right], \tag{B.43}$$

as for all other modes $\lambda_k^t = \lambda_k = \lambda_B$, and so the contributions vanish. As in the main body, if we again consider the case of an emerging signal, then we can assume that $A\pi f(v)/(m_a v) \sim A_t\pi f(v)/(m_a v) \ll \lambda_B$, which to lowest order simplifies our result as

$$\tilde{\Theta}(A) = 2A(A_t - A)\left(\frac{\pi f(v)}{m_a v \lambda_B}\right)^2. \tag{B.44}$$

Note the velocity appearing in this result is fixed by the value of $k_S$.

By relating the collection time to the width of our frequency bins and hence velocity, we have again that $1/T = m_a v\Delta v/(2\pi)$, where recall $\Delta v$ is the width with which we can probe in velocity space. Accordingly we arrive at

$$\tilde{\Theta}(A) = \frac{1}{2}T^2A(A_t - A)\left(\frac{f(v)\Delta v}{\lambda_B}\right)^2. \tag{B.45}$$

Importantly, note that as $f(v)$ is a normalized pdf and $\Delta v$ is roughly the range over which it varies, we have $f(v)\Delta v \sim \mathcal{O}(1)$. The exact numerical value is irrelevant: the key observation is that the combination is no longer dependent on $T$. As such we see in this limit $\tilde{\Theta} \propto T^2$, which should be contrasted with (3.44), where $\tilde{\Theta} \propto T$. As $A \propto g_{a\gamma\gamma}^2$, when we use $\tilde{\Theta}$ to derive the TS or 95% limit as we did in the main body we will find they both scale as $T^{-1/2}$, as expected.

# APPENDIX C

# The Statistics of Axion Direct Detection with Multiple Detectors

## C.1 Coherence Length and Time

In this appendix we briefly review the concepts of the coherence length and time, as relevant to wavelike DM. We emphasize that in our work both concepts only arise heuristically. Indeed, the coherence length and time are only defined parametrically, and for all quantitative results we instead rely on the likelihood formalism in [80], which produces not only all parametric scalings, but also the required $\mathcal{O}(1)$ factors.

Consider first the coherence length $\lambda_c \sim (m_{\rm DM} v_0)^{-1}$, the scale over which wavelike DM remains coherent. In discussions of ultralight DM, "coherence length" is often used interchangeably with "de Broglie wavelength." Strictly speaking, though, the de Broglie wavelength $\lambda_{\rm dB} = 2\pi/(m_a v)$ is a property of particles with fixed velocity $v$, while the coherence length describes the dephasing of various plane wave components with different velocities. When $v_0 \sim v$, these two length scales are comparable, but there are relevant situations where the two diverge, such as for cold streams, and then the distinction between the coherence and de Broglie wavelengths becomes important.

The coherence time is then the timescale over which a measured signal of ultralight DM will build up coherently. In real space, this is the time it take for a new spatially coherent packet of the DM wave, which has size $\lambda_c$, to arrive at the instrument. If these packets travel with a mean speed of $\bar{v}$, then the timescale is $\tau \sim \lambda_c/\bar{v} \sim (m_{\rm DM} \bar{v}\, v_0)^{-1}$. The same result can be arrived at from a frequency space consideration. The Fourier transform of an experimental data set collected over a time $T$ will have a frequency resolution of $\Delta\omega = 2\pi/T$. If the entire signal fits within a single frequency bin, the result is associated with a single draw from an exponential distribution, as shown in [80]. Once we resolve the signal, however, we obtain multiple draws which will combine incoherently, partially offsetting the benefit of additional integration time. The coherence time is

therefore dictated by the width of the signal in frequency space, and then as $d\omega = m_{\rm DM} v\, dv$, we again arrive at $\tau \sim (m_{\rm DM} \bar{v}\, v_0)^{-1}$.

## C.2 Demonstrating $\mathbf{d} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$

The goal of this appendix is to demonstrate a fact that was used without proof in the main body: the data set $\mathbf{d}$, given in (4.15), is a random variable drawn from a multivariate normal distribution with zero mean and covariance matrix $\mathbf{\Sigma}$ as given by (4.16). In order to show this we will start from the known statistics of the axion field, as reviewed in the main body, together with a Gaussian background, and show that the mean and variances of the data sets follow the expected normal distribution. We will further confirm this result with a Monte Carlo realization of the axion field. From here, rather than confirm that all higher moments are also consistent with Gaussianity, we will instead confirm numerically that the distribution is normal. Indeed, the diagonal components of $\mathbf{d}$, which govern the statistics of individual detectors, must be Gaussian as proven in [80].

Let us begin by restating (4.6) in a simplified notation. We introduce a single multi-index $d = abc$, and a random variable $f_d = \alpha_d \sqrt{f(\mathbf{v}_d)(\Delta v)^3}$, yielding

$$a(\mathbf{x}, t) = \frac{\sqrt{\rho_{\rm DM}}}{m_a} \sum_d f_d \cos\left(\omega_d t - m_a \mathbf{v}_d \cdot \mathbf{x} + \phi_d\right). \tag{C.1}$$

We now envision collecting a data sensitive to this axion field at each of the $\mathcal{N}$ detectors, located at positions $\mathbf{x}_i$. Specifically, we imagine collecting $N$ measurements at a frequency $f = 1/\Delta t$ at each experiment, so that we have at our disposal $N \times \mathcal{N}$ data points $\{\Phi_n^{(i)}\}$, with

$$\Phi_n^{(i)} = m_a \sqrt{\frac{A_i}{\rho_{\rm DM}}} a_n(\mathbf{x}_i,\, n\Delta t) + x_n^{(i)}\,. \tag{C.2}$$

The second term in this expression captures the background noise. We will assume the noise is Gaussian, which holds for a wide range of sources as described in the main body, and in detail that it satisfies

$$\left\langle x_n^{(i)} \right\rangle = 0\,, \quad \left\langle x_n^{(i)} x_m^{(j)} \right\rangle = \delta_{ij}\delta_{nm} \frac{\lambda_{B,i}}{\Delta t}\,. \tag{C.3}$$

In other words, we assume the noise has zero mean, is uncorrelated between detectors, and has a variance that increases with the measurement frequency $f$. The variance is controlled by the mean power in the background, $\lambda_{B,i}$, and if there are multiple background sources at a single detector, their power can simply be combined.

From this data set, we compute the discrete Fourier transform $\{\Phi_k^{(i)}\}$ using (4.7), and then the associated real and imaginary parts, $R_k^{(i)}$ and $I_k^{(i)}$ from (4.8). These variables are what combine to form the data vector $\mathbf{d}$, and so the goal is to study their statistics. Before proceeding, let's introduce some further notation to keep expressions compact. Firstly, we encapsulate the axion phase into a single term,

$$\varphi_{d,n}^{(i)} = \omega_d n\Delta t - m_a \mathbf{v}_d \cdot \mathbf{x}_i + \phi_d \,. \tag{C.4}$$

To capture the trigonometric sums introduced by the Fourier transforms, we write

$$c_{n,k} = \cos\left(\frac{2\pi kn}{N}\right) = \cos\left(\omega n\Delta t\right), \tag{C.5}$$

and the equivalent expression for sine is denoted $s_{n,k}$. Using this, the real and imaginary parts of the data set can be written

$$\begin{aligned} R_k^{(i)} =& \frac{\Delta t}{\sqrt{T}} \sum_{n=0}^{N-1} \left[\sqrt{A_i}\sum_d f_d \cos\varphi_{d,n}^{(i)} + x_n^{(i)}\right] c_{n,k}, \\ I_k^{(i)} =& -\frac{\Delta t}{\sqrt{T}} \sum_{n=0}^{N-1} \left[\sqrt{A_i}\sum_d f_d \cos\varphi_{d,n}^{(i)} + x_n^{(i)}\right] s_{n,k} \,. \end{aligned} \tag{C.6}$$

From these expressions, we can see immediately that $\langle R_k^{(i)}\rangle = \langle I_k^{(i)}\rangle = 0$. That this holds for the background follows from (C.3), and for the axion signal contribution we have

$$\left\langle f_d \cos\varphi_{d,n}^{(i)}\right\rangle = \langle f_d\rangle \left\langle \cos\varphi_{d,n}^{(i)}\right\rangle = 0 \,. \tag{C.7}$$

The first step follows as the value of $\alpha_d$ (and hence $f_d$) is uncorrelated with $\phi_d$ (and hence $\varphi_{d,n}^{(i)}$), whilst the second utilizes the fact $\langle\cos\varphi\rangle = 0$ when the argument $\varphi$ is a random phase. This establishes that $\langle\mathbf{d}\rangle = \mathbf{0}$.

Next we consider the covariances. In particular, we will compute $\langle R_k^{(i)} R_k^{(j)}\rangle$. The calculation where one or both of the real components is replaced by an imaginary equivalent proceeds simi-

larly, and we will comment on the important differences throughout. In detail, we will compute

$$\begin{aligned}\left\langle R_k^{(i)} R_k^{(j)} \right\rangle &= \frac{(\Delta t)^2}{T} \\ &\times \left\langle \sum_{n=0}^{N-1} \left[ \sqrt{A_i} \sum_d f_d \cos\varphi_{d,n}^{(i)} + x_n^{(i)} \right] c_{n,k} \right. \\ &\times \left. \sum_{m=0}^{N-1} \left[ \sqrt{A_j} \sum_s f_s \cos\varphi_{s,m}^{(j)} + x_m^{(j)} \right] c_{m,k} \right\rangle .\end{aligned} \tag{C.8}$$

Note the effect of sending $R_k^{(j)} \to I_k^{(j)}$ is simply to replace $c_{m,k} \to -s_{m,k}$, and similarly for $R_k^{(i)}$. Continuing with the calculation at hand, expanding out the final two lines, we will have expressions involving only the signal, only the background, and cross terms. As the background value is uncorrelated with the signal, the cross terms will be zero via an almost identical argument to the vanishing of the means. Of the remaining terms, consider the background first.

$$\frac{(\Delta t)^2}{T} \left\langle \sum_{n,m=0}^{N-1} \left(x_n^{(i)} c_{n,k}\right) \left(x_m^{(j)} c_{m,k}\right) \right\rangle = \frac{\delta_{ij}\lambda_{B,i}}{N} \sum_{n=0}^{N-1} (c_{n,k})^2 = \frac{\delta_{ij}\lambda_{B,i}}{2} , \tag{C.9}$$

which holds except for $k = 0$ (or $k = N/2$ for $N$ even). Note if we were evaluating $\langle I_k^{(i)} I_k^{(j)} \rangle$, we would have the same expression but with $c_{n,k} \to s_{n,k}$, and therefore the background contribution would be identical. If we were evaluating $\langle R_k^{(i)} I_k^{(j)} \rangle$, however, the background contribution would vanish as $\sum c_{n,k} s_{n,k} = 0$. Taken together, these results demonstrate the appearance of $\lambda_B$ in (4.16).

Now we turn to the signal contribution, for the moment dropping the overall factor of $\sqrt{A_i A_j}(\Delta t/N)$,

$$\left\langle \sum_{n,m=0}^{N-1} \sum_{d,s} f_d f_s \cos\varphi_{d,n}^{(i)} \cos\varphi_{s,m}^{(j)} c_{n,k} c_{m,k} \right\rangle = \sum_{n,m=0}^{N-1} c_{n,k} c_{m,k} \sum_{d,s} \langle f_d f_s \rangle \langle \cos\varphi_{d,n}^{(i)} \cos\varphi_{s,m}^{(j)} \rangle . \tag{C.10}$$

Again we used the independence of the amplitude and phase of the random walk that emerges in calculating the axion field statistics. The second expectation value in this expression will vanish unless the random phases in the cosines are equal, effectively as

$$\langle e^{i(\phi_d - \phi_s)} \rangle = \delta_d^s . \tag{C.11}$$

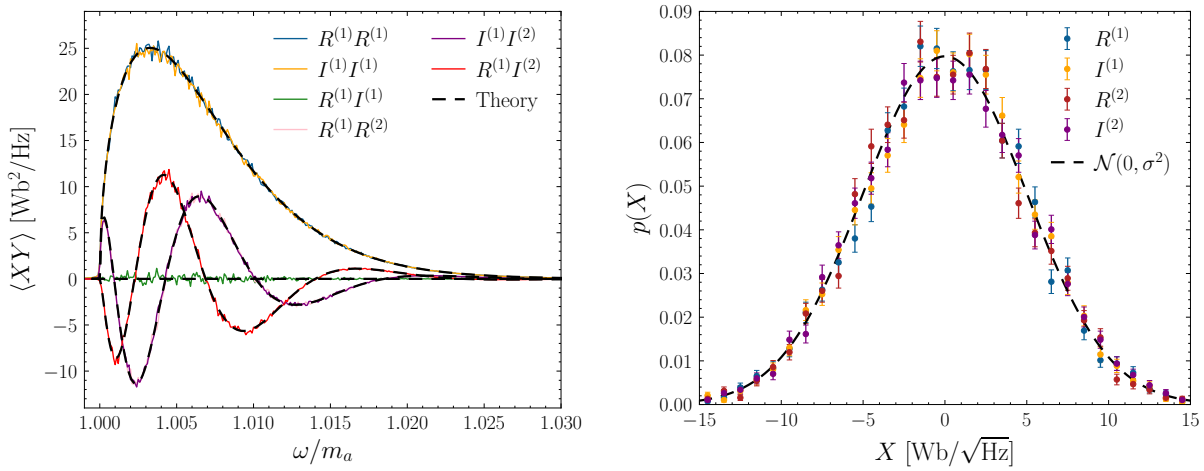



Figure C.1: Monte Carlo validation that the statistics of DM interferometry are as claimed in App. C.2. In the left figure we confirm that the variances of the real and imaginary signal-only data sets, collected for the $\mathcal{N} = 2$ experiments, is as claimed in (4.16). This was proven directly in the text, but in the plot we show that the average of 4,000 Monte Carlo simulations provides a consistent prediction for the variances as a function of frequency in the different cases. On the right figure, for the frequency where $\langle R^{(1)}R^{(1)}\rangle$ achieves its maximum, we show the distribution of values across the simulations. In detail, we see that the real and imaginary components are normally distributed, and consistent with a mean-zero normal distribution, where the variance is given as on the left, here $\sigma^2 \approx 25\,\text{Wb}^2/\text{Hz}$. We found that the distributions were consistent with the Gaussian expectation at the level of $p > .05$ using the D'Agostino and Pearson omnibus normality test [432, 433]. In both cases, each Monte Carlo simulation involves a direct construction of the axion field starting from (C.16) with $N_a = 100,000$, taking $m_a = 2\pi$ Hz, and $A = 1\,\text{Wb}^2$. Further, we take the velocity distribution to follow a variant of the SHM in (4.50), but with $v_0 = 0.07$ and $\mathbf{v}_\odot = (0, 0.08, 0)$, both in natural units. The (unphysically) large velocity helps simplify the computation of the Fourier transform. The detector separation is $\mathbf{x}_{12} = d(0, 1, 0)$, with $d \approx 4.4\lambda_c$.

Further, as $\langle\alpha_d^2\rangle = 2$, we can also evaluate the result as

$$\sum_{d,s}\langle f_d f_s\rangle\langle\cos\varphi_{d,n}^{(i)}\cos\varphi_{s,m}^{(j)}\rangle = 2\sum_d f(\mathbf{v}_d)(\Delta v)^3\langle\cos\varphi_{d,n}^{(i)}\cos\varphi_{d,m}^{(j)}\rangle\,, \tag{C.12}$$

which we can simplify further as,

$$\begin{aligned}
&\langle\cos\varphi_{d,n}^{(i)}\cos\varphi_{d,m}^{(j)}\rangle \qquad\qquad (C.13)\\
=&\frac{1}{2}\left[\cos\left(\omega_d(n-m)\Delta t - m_a\mathbf{v}_d\cdot\mathbf{x}_{ij}\right) + \langle\cos\left(\omega_d(n+m)\Delta t - m_a\mathbf{v}_d\cdot(\mathbf{x}_i+\mathbf{x}_j) + 2\phi_d\right)\rangle\right]\\
=&\frac{1}{2}\cos\left(\omega_d(n-m)\Delta t - m_a\mathbf{v}_d\cdot\mathbf{x}_{ij}\right)\\
=&\frac{1}{2}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\cos\left(m_a\mathbf{v}_d\cdot\mathbf{x}_{ij}\right) + \frac{1}{2}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\sin\left(m_a\mathbf{v}_d\cdot\mathbf{x}_{ij}\right).
\end{aligned}$$

In the final step we see the emergence of the $\mathbf{k}\cdot\mathbf{x}$ type phase factors that separate $\mathcal{F}_{ij}^{c,s}(v)$ defined in (4.18) from $f(v)$. We have broken the calculation into a number of pieces at this stage, let us begin to put things back together. Combining the different expressions above, we have

$$\begin{aligned}\left\langle R_k^{(i)} R_k^{(j)} \right\rangle =& \frac{1}{2}\delta_{ij}\lambda_{B,i}(\omega) + \sqrt{A_i A_j}\frac{\Delta t}{N}\sum_d f(\mathbf{v}_d)(\Delta v)^3 \times \qquad \text{(C.14)}\\ & \left[\cos\left(m_a \mathbf{v}_d\cdot\mathbf{x}_{ij}\right)\sum_{n,m=0}^{N-1} c_{n,k}c_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\right.\\ & \left. + \sin\left(m_a \mathbf{v}_d\cdot\mathbf{x}_{ij}\right)\sum_{n,m=0}^{N-1} c_{n,k}c_{m,k}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\right]\\ =& \frac{1}{2}\delta_{ij}\lambda_{B,i}(\omega) + \frac{\pi\sqrt{A_i A_j}}{2}\sum_d (\Delta v)^3 f(\mathbf{v}_d)\cos\left(m_a\mathbf{v}_d\cdot\mathbf{x}_{ij}\right)\delta(\omega_d-\omega) \qquad \text{(C.15)}\\ =& \frac{1}{2}\delta_{ij}\lambda_{B,i}(\omega) + \frac{\pi\sqrt{A_i A_j}}{2m_a v_\omega}\int d^3\mathbf{v} f(\mathbf{v})\cos\left(m_a\mathbf{v}\cdot\mathbf{x}_{ij}\right)\delta(|\mathbf{v}|-v_\omega)\\ =& \frac{1}{2}\left[c_{ij}(\omega)+\delta_{ij}\lambda_{B,i}(\omega)\right].\end{aligned}$$

The final result is the claimed form of $\langle R_k^{(i)} R_k^{(j)}\rangle$ used in the main body, but let us detail the steps in the calculation, working backwards. In the last step we simply recalled the definitions introduced in (4.18) and (4.17). The penultimate step simply involved approximating the sum over all velocity components $d = abc$ with an equivalent integral. The only non-trivial manipulation occurred when we evaluated the sums over $n$ and $m$. These were performed using a set of discrete Fourier transform double orthogonality relations, which for convenience we have collected in App. C.3. From those relations, we can see that as $\langle R_k^{(i)} R_k^{(j)}\rangle$ involved $c_{n,k}c_{m,k}$, only the cosine of $\mathbf{k}\cdot\mathbf{x}_{ij}$ survived. By analogy, if we were evaluating $\langle I_k^{(i)} I_k^{(j)}\rangle$, we would instead have $s_{n,k}s_{m,k}$ in the sums, which would again isolate the cosine. On the other hand, for $\langle R_k^{(i)} I_k^{(j)}\rangle$ (where the background contribution vanishes as described above), we have $c_{n,k}s_{m,k}$, which instead singles out the sine, implying the above result would have $c_{ij}(\omega) \to s_{ij}(\omega)$. The same argument holds for $\langle I_k^{(i)} R_k^{(j)}\rangle$, up to a sign.

Taken together, the above arguments suffice to demonstrate analytically that the variance of the data set is as claimed in the main body. We can also confirm this result numerically. On the left of Fig. C.1, we show that a direct construction of the axion field as a sum over $N_a$ plane wave components,

$$a(\mathbf{x},t) = \frac{\sqrt{2\rho_{\rm DM}}}{m_a\sqrt{N_a}}\sum_{i=1}^{N_a}\cos\left[\omega_i t - m_a\mathbf{v}_i\cdot\mathbf{x}+\phi_i\right], \qquad \text{(C.16)}$$

where $\mathbf{v}_i$ is drawn from $f(\mathbf{v})$ and $\phi_i$ is drawn uniformly from $[0, 2\pi)$, leads to the exact same results.[1] The detailed parameter choices are described in the figure caption, and the curves represent the average over repeating this procedure 4,000 times. In all cases, there is excellent agreement between this approach and the corresponding theory curves.

On the right of Fig. C.1 we confirm a point that we did not demonstrate directly, namely that the individual real and imaginary components are normally distributed. The distribution is shown amongst the 4,000 simulated data sets for the two components measured at two different detectors. In all cases consistency is observed with the predicted Gaussian distribution. We performed a chi-squared test to determine the goodness of fit and found $p$-values greater than $0.05$. In detail, the $R^{(1)}$, $I^{(1)}$, $R^{(2)}$, and $I^{(2)}$ data sets shown in Fig. C.1, had corresponding $p$-values of $0.06$, $0.12$, $0.97$, and $0.27$.

## C.3 Orthogonality Relations

In App. C.2 we made use of several unstated orthogonality relations. We collect these in the present appendix. Firstly, the following expressions vanish for any $k$

$$\begin{aligned}
&\sum_{n,m=0}^{N-1} c_{n,k}s_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} s_{n,k}c_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} c_{n,k}c_{m,k}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} s_{n,k}s_{m,k}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=0\,.
\end{aligned} \tag{C.17}$$

[1] Binning the velocities leads to (4.6) in the main text, with a Rayleigh-distributed amplitude in each bin.

However, there are four non-zero combinations. In detail, for most values of $k$,

$$\begin{aligned}
&\sum_{n,m=0}^{N-1} c_{n,k}c_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} s_{n,k}s_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} c_{n,k}s_{m,k}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\sum_{n,m=0}^{N-1} s_{n,k}c_{m,k}\sin\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=\left(\frac{N}{2}\right)^2\frac{2\pi}{T}\delta(\omega_d-\omega).
\end{aligned} \tag{C.18}$$

The exception to the above is if $k=0$, or $k=N/2$ for $N$ even. For those values, only one of the above three sums is non-zero, in detail

$$\begin{aligned}
&\sum_{n,m=0}^{N-1} c_{n,k}c_{m,k}\cos\left(\frac{\omega_d}{\omega}\frac{2\pi k(n-m)}{N}\right)\\
&=N^2\frac{2\pi}{T}\delta(\omega_d-\omega).
\end{aligned} \tag{C.19}$$

However, recall that we usually exclude these exceptional $k$ values from our likelihood.

The non-zero results above were written in terms of Dirac $\delta$-functions, however this is an approximation. Recall all results are obtained through the discrete Fourier transform, within which the frequency can be interpreted as $\omega=(2\pi/T)k$, with $k=0,1,\ldots,N-1$. In truth, if we define $k_d=\lfloor\omega_d T/2\pi\rfloor$, then what appears in the above sums is the Kronecker-delta $\delta^k_{k_d}$. However, in the spirit of assuming our frequency resolution is sufficient enough to approximate $\omega$ as a continuous variable, we take

$$\delta^k_{k_d}=\delta(k_d-k)=\frac{2\pi}{T}\delta(\omega_d-\omega), \tag{C.20}$$

which is the form it appears in (C.18) and (C.19).

## C.4 Data Stacking Procedure

In practical situations it is usually neither feasible nor necessary to save the entire time-series data to disk and then construct the Fourier transform of the full data set. The frequency resolution of this complete Fourier transform would be $\Delta\omega = 2\pi/T$, and potentially much smaller than the scale of any expected features induced by the signal due to $f(\mathbf{v})$. As a specific example, the ABRACADABRA-10 cm experiment [72, 82] recorded the PSD data over short time periods and then stacked the PSD data over the time subintervals to construct the average PSD data. The advantage of this averaging procedure is that it requires less storage and is easier to deal with computationally, since there are less frequencies involved than would be in the full data set without time sub-binning.

With this in mind, it is useful to understand how we may stack the Fourier transform data over multiple experiments in such a way that we preserve the full power of the likelihood in (4.19) but that allows us to reduce the data volume needed to be saved to disk. (An optimized procedure for stacking the data from a single experiment is presented in [80].) Let us imagine that we record time-series data in $N_T$ equal time subintervals of time $\Delta T = T/N_T$, and that in each subinterval the frequency spacing of the $\Delta N = N/N_T$ Fourier components is sufficient to resolve the axion signal by multiple frequency bins, *i.e.* we retain sufficient frequency resolution that our signal remains well resolved. We then denote the full data set by $d = \{\mathbf{d}_k^\ell\}$, indexed now by both $k = 1, \ldots, \Delta N - 1$, denoting the Fourier component, and $\ell = 1, \ldots, N_T$, the data subinterval. The appropriate likelihood is then simply the product of the likelihood in (4.19), but now also over all values of $N_T$. However as $N_T \times \Delta N = N$, the number of frequency bins in the Fourier transform of the full data, at this stage we have not reduced the size or complexity of the data or likelihood evaluations at all. In order to do so, we can combine the data into the following average data matrix, which can be computed prior to any evaluation likelihood,

$$[\bar{\mathbf{d}}_k]^{ij} = \frac{1}{N_T}\sum_{\ell=1}^{N_T} d^i_{k,\ell} d^j_{k,\ell}\,. \tag{C.21}$$

Here, the indices $i$ and $j$ run over the $2\mathcal{N}$ entries of the data vector in (4.15), $k$ indexes the discrete Fourier transform, and $\ell$ specifies the appropriate subintervals. In terms of the average data matrix, the likelihood can be written as

$$\mathcal{L}(d|\mathcal{M},\boldsymbol{\theta}) = \prod_{k=1}^{\Delta N-1} \frac{\exp\left[-\frac{N_T}{2}\mathrm{Tr}(\bar{\mathbf{d}}_k \cdot \boldsymbol{\Sigma}_k^{-1})\right]}{\left[(2\pi)^{2\mathcal{N}}|\boldsymbol{\Sigma}_k|\right]^{N_T/2}}\,, \tag{C.22}$$

where we have left the $\boldsymbol{\theta}$ dependence of $\boldsymbol{\Sigma}$ implicit. We can now compare how much data needs to be stored for this stacking procedure compared to the full data set. Again, we have $N_T$ subintervals,

each with $\Delta N$ Fourier components, and for each we have $2\mathcal{N}$ components in our data vector. As (C.21) is a real symmetric matrix, we need $\mathcal{N}(2\mathcal{N}+1)$ components to specify it for each $k$ value. Thus in total, we need to store $\mathcal{N} \times (2\mathcal{N}+1) \times \Delta N$ entries to disk, although if $\mathbf{\Sigma}_k^{-1}$ has a number of zeros (associated with experiments well within or outside the coherence length $\lambda_c$), fewer points may be required. This number should be contrasted with the $2\mathcal{N} \times N = 2\mathcal{N} \times \Delta N \times N_T$ values that would be needed in the absence of a data stacking procedure. Thus, as long as $N_T \gg \mathcal{N}$, a significant reduction in the data set can be achieved. For simplicity, in the main body of the paper we assume that no data stacking has been performed, though it is important to keep in mind that all results we derive may also be applied to the stacked data likelihood. An important caveat is that care should be taken when accounting for daily modulation to make sure the data is stacked with other data taken at a similar time of day, otherwise the effect can be washed out.

Finally, we briefly demonstrate using the Asimov procedure that as long as the subintervals retain sufficient frequency resolution that the signal remains well resolved, the stacked and full likelihoods are equally sensitive. If the signal prediction remains unchanged in each subinterval, then the averaged data set defined in (C.21) has the following expected value,

$$\langle [\bar{\mathbf{d}}_k]^{ij} \rangle = \frac{1}{N_T} \sum_{\ell=1}^{N_T} \langle d^i_{k,\ell} d^j_{k,\ell} \rangle = \mathbf{\Sigma}^t \,. \tag{C.23}$$

It is straightforward to then evaluate the equivalent Asimov $\Theta$, and one finds a result enhanced by $N_T$, but with $T \to \Delta T$ when replacing the sum over Fourier components with an integral over speed. For instance, the equivalent of (4.29) has $T \to N_T \Delta T$. Yet as $N_T \Delta T = T$, by definition, the test statistic is identical, and therefore the stacking procedure is optimal as claimed.

# APPENDIX D

# First Results from ABRACADABRA-10 cm

In recent years, the absence of a compelling direct detection of DM in accelerator and Weakly Interacting Massive Particle (WIMP) searches has reignited the search for ADM. The coincidence of new developments in quantum sensors and quantum information technology has stoked this reawakened interest, and the past few years have seen a wealth of new experimental ideas and approaches that are beginning to revolutionize the field [187, 434]. While most ADM searches have traditionally focused on a narrow mass range from $10 \lesssim m_a \lesssim 100\,\mu$eV, recent theoretical work has made a compelling case for ADM in the mass range $m_a \lesssim 1\,\mu$eV [51, 52, 169, 251, 252, 264–267, 269, 435].

The ABRACADABRA-10 cm experiment has recently released results of the first direct detection search for ADM below $1\,\mu$eV [72]. The design of the experiment was motivated by the proposal of [71], and is based on measuring the coupling of ADM to electromagnetism – similar to experiments probing different mass regimes like the long-running ADMX [67, 258] and HAYSTAC [219]. In this lower mass range, the axion field $a$, in the presence of a large magnetic field can be thought of as an induced effective current

$$\mathbf{J}_{\text{eff}} = g_{a\gamma\gamma}\frac{\partial a}{\partial t}\mathbf{B}\,, \tag{D.1}$$

where $g_{a\gamma\gamma}$ is the axion-photon coupling. This current sources a small AC magnetic field that can be measured with a sensitive enough magnetometer.

ABRACADABRA-10 cm is a prototype detector for a new search approach, and its implementation contains novel elements that have not previously been used in ultralight dark matter searches:

- A toroidal magnet geometry, with the detection element placed in the near-zero-field region – the first operational non-microwave cavity sub-eV ADM search;
- A broadband readout mode involving continuous-stream data-taking for $\sim 10^6$ seconds (roughly 1 month), and several compression techniques to mitigate the total data storage requirements while preserving the desired signal bandwidth;

- A calibration technique where a signal is injected through current in a calibration loop in a similar geometry as the expected axion signal;
- A data analysis pipeline tailored to the expected statistics of the axion field in the quasistatic regime, where the signal is best described as a flux power spectral density rather than photon-counting with the added constraint that "rescanning" is prohibitively time-consuming.

In this Appendix, we provide context and additional details for each of these novel elements and their specific implementation in ABRACADABRA-10 cm. In Section D.1, we describe the design and construction of the toroidal ABRACADABRA-10 cm detector. We describe the data collection approach used for the broadband readout in Sec D.2, and describe the calibration of the detector in Sec. D.3. In Sec. D.4, we describe the data analysis and limit extraction approach used for our broadband search. We conclude by commenting on the improvements and modifications necessary for scaling up the ABRACADABRA-10 cm prototype to an experiment capable of probing QCD axion couplings.

# D.1 Detector Design and Construction

The ABRACADABRA-10 cm detector and setup can be split into six separate systems: the toroidal magnet, the magnet support infrastructure and shielding, the pickup loop circuit, the SQUID electronics, the calibration circuit, and finally the cryostat and detector support infrastructure. In this section, we discuss the design and construction of each.

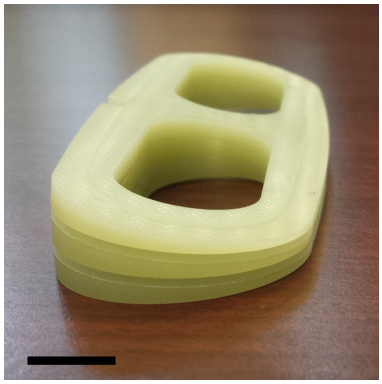
The expected signal-to-noise ratio (SNR) of ABRACADABRA-10 cm can be written approximately as

$$\mathrm{SNR} = g_{a\gamma\gamma}\sqrt{\rho_{\mathrm{DM}}}\mathcal{G}VB_{\mathrm{max}}\left(\frac{M_{\mathrm{in}}}{L_T}\right)\frac{(\tau t)^{\frac{1}{4}}}{S_{\Phi\Phi}^{1/2}}\,, \tag{D.2}$$

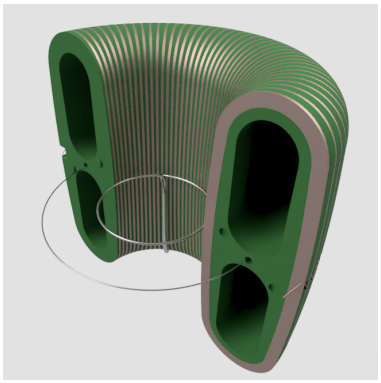
where $V$ is the volume of the toroid, $\mathcal{G}$ is a geometric factor, $B_{\mathrm{max}}$ is the max field inside the toroid, $M_{\mathrm{in}}$ is the inductive coupling of the SQUID, and $L_T$ is the total inductance of the readout circuit. Here, we assume that the integration time $t$, exceeds the axion coherence time, $\tau$. The final parameter of importance is the flux noise level, $S_{\Phi\Phi}^{1/2}$, typically measured in $\mu\Phi_0/\sqrt{\mathrm{Hz}}$. The relevant parameters are summarized in Table D.1.

## D.1.1 The Toroidal Magnet

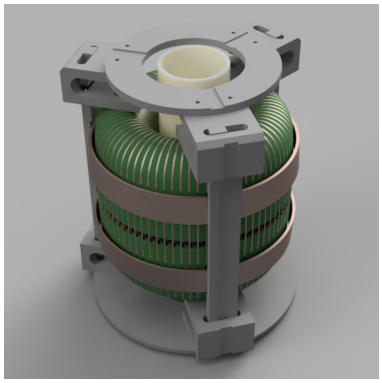
The magnet structure is built around 80 identical Delrin wedges, (see Fig. D.1a). When glued together, they create a toroidal frame with an inner radius of 3 cm and an outer radius of 6 cm, with a total height of 12 cm. The total volume of magnetic field is $V \approx 890\,\mathrm{cm}^3$.



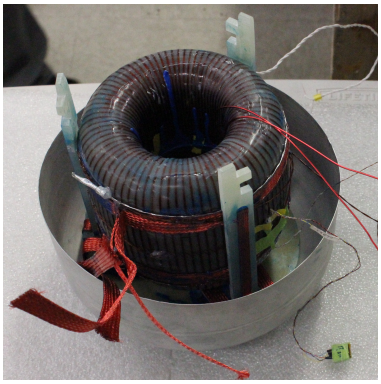
(a)



(b)



(c)



(d)

Figure D.1: (a) Three of the 80 Delrin wedges that form the toroid structure stacked together. The black bar indicates a $\approx$ 1 cm scale. (b) Cutaway rendering of the toroid with the 1 mm diameter wire pickup loop in the center. A 0.5 mm diameter wire runs through the center of field region to form the calibration loop. Toroid height is $\approx$ 12 cm. (c) Rendering of the ABRACADABRA-10 cm support structure. The pickup loop is supported by a PTFE (white) tube through the center. The magnet is supported by an outer G10 support structure and thermalized with two copper bands. (d) Photo of the assembled ABRACADABRA-10 cm, with the top of the superconducting shield and support structure removed.

The magnet current is carried by a NbTi(CuNi) wire which is wound 1,280 times around the magnet. Between each pair of wedges is a groove that has 16 winds of wire, laid down in pairs 8 layers deep. The wire is held in place with epoxy. To cancel the azimuthal current, the toroid is counterwound.

The toroid was wound by Superconducting Systems, Inc [270] in three separate pieces, with three separate lengths of NbTi(CuNi) wire. The pieces are then glued together and the wires are connected together with two superconducting crimps. These crimps are then attached to the outside of the toroid. These joints could create small stray fields which contribute to the backgrounds for the axion search, but unfortunately could not be avoided in the construction.

The toroid is mounted in a dilution refrigerator (described below) and cooled to $\lesssim 1\,\mathrm{K}$. The NbTi(CuNi) wire superconducts below $\lesssim 9\,\mathrm{K}$. We charge the magnet by injecting a 121 A current into the toroid. Once charged, we turn off heat to a superconducting switch (located away from the magnet) which then locks the current into the magnet. The current source is disconnected from the charging leads on top of the refrigerator.

When fully charged, the maximum field in the magnet is $B_{\mathrm{max}} = 1\,\mathrm{T}$. This was confirmed with a Hall probe to a precision of $\sim 1\%$, with the uncertainty coming from uncertainty in position of the probe in the field. Once in persistent mode, we observed no decay in the field to a precision of $\lesssim 0.1\%$ on the scale of 1 week. The Hall sensor was removed before normal data taking.

### D.1.2 Support Infrastructure

The toroid is mechanically supported by a G10 frame held together with nylon bolts (see Fig. D.1c). The goal of this structure was to rigidly mount the toroid in place, while minimizing the amount of non-superconducting metal near the magnet. The one exception to this is the copper straps which wrap around the outside of the toroid that provide the required thermalization to cool the magnet. These straps undoubtedly contribute some level of noise for our axion search, though in the current setup it is not the dominant noise source. In the future we will search for alternative thermalization approaches.

The entire toroid and support structure are mounted inside the external shield (see Fig. D.1d). The shield consists of a spun copper can that has been coated inside and out with a 25-75 $\mu$m layer of tin, for a total thickness of $\approx$1 mm. The copper provides good thermal conductivity to minimize thermal gradients across the shield. It also provides the thermal conductivity to the copper straps which cool the magnet. Once below 3.7 K, the tin becomes superconducting and expels environmental magnetic fields and acts as a shield against electromagnetic interference. Optimizing and characterizing this external shielding will be the subject of future work.

The external shield is built in two hemispheres (top and bottom) which have $\approx 12\,\mathrm{mm}$ of vertical

Table D.1: Summary of the ABRACADABRA-10 cm detector design parameters.

| | | |
|---|---|---|
| Pickup Loop Radius | $R_p$ | 20.1 mm |
| Pickup Loop Wire Diameter | $r_p$ | 1.0 mm |
| Magnet Inner Radius | $R_{\text{in}}$ | 30 mm |
| Magnet Outer Radius | $R_{\text{out}}$ | 60 mm |
| Magnet Height | $h$ | 120 mm |
| Magnet Max Field | $B_{\text{max}}$ | 1.0 T |
| Geometric Factor | $\mathcal{G}_V$ | 0.027 |
| Pickup Loop Inductance | $L_p$ | 95.5 nH |
| SQUID Input Inductance | $L_{\text{in}}$ | 150 nH |
| SQUID Inductive Coupling | $M_{\text{in}}$ | 2.5 nH |

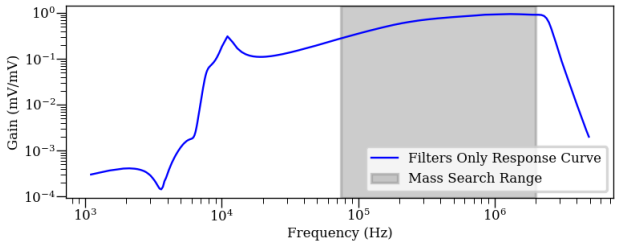



Figure D.2: Gain of the combined high-pass and anti-aliasing filters. All spectra are corrected for this response function – unless otherwise noted. Measured in-situ, using injected signals at different frequencies. This also defines the usable range of data. For our search we use the range 75 kHz – 2 MHz.

overlap when assembled. There is a small gap in one location between the inner and outer shield through which the magnet wires, pickup loop wires and calibration loop wires pass as three sets of twisted pairs. The shield halves are connected with a layer of solder and epoxy to ensure both electrical and mechanical connection.

A 12 mm thick aluminum top plate is epoxied to the top of the top shield and acts as the contact point for the thermalization to the rest of the cryostat and mechanical mounting point to the vibration isolation system (see below). The aluminum plate is electrically isolated from the shield to minimize grounding loops. A thermometer is epoxied to the outside of the bottom shield which monitors the temperature of the farthest point from the thermalization. However, during data taking this thermometer is not active.

### D.1.3 Pickup Loop Circuit

The pickup loop measures the magnetic field in the center of the toroid – a region that should have zero field in the absence of an axion signal. The time-averaged magnitude of the flux through the

pickup loop due to $\mathbf{J}_{\rm eff}$ can be written as

$$|\Phi_a|^2 = g_{a\gamma\gamma}^2 \rho_{\rm DM} V^2 \mathcal{G}_V^2 B_{\rm max}^2 \equiv A\,, \tag{D.3}$$

where $\rho_{\rm DM}$ is the DM density, $V$ is the volume of the magnet, $B_{\rm max}$ is the maximum field in the magnet and $\mathcal{G}$ is a geometric factor. The pickup loop itself consists of a 1 mm diameter solid NbTi wire wrapped around the outside of a 5.5 cm diameter polytetrafluoroethylene (PTFE) tube that is 18.1 cm tall. The geometric factor $\mathcal{G}_V$ weights the effective current in Eqn. D.1, by the contribution to the flux through the pickup loop. This can be written as

$$\mathcal{G}_V \equiv \frac{1}{B_{\rm max} V} \left| \int_{\rm Loop} dA \int_{\rm Toroid} dV' \frac{\mathbf{B}(\mathbf{r}') \times (\mathbf{r}' - \mathbf{r})}{|\mathbf{r}' - \mathbf{r}|^3} \cdot \hat{\mathbf{n}} \right| \tag{D.4}$$

where $\hat{\mathbf{n}}$ is the normal to the plane of the pickup loop, and the integrals are taken over the area enclosed by the pickup loop and the volume of the toroid. The integrand is reminiscent of the Biot-Savart law, with the current taken to be the axion-induced effective current $\mathbf{J}_{\rm eff}$ which follows lines of $\mathbf{B}$ [71].[1] For the ABRACADABRA-10 cm geometry, we calculate this using a COMSOL [277] simulation to be $\mathcal{G}_V = 0.027$. The two wire leads from the pickup loop, which consists of the same wire as the loop, are twisted into a twisted pair configuration and run out under the bottom of the toroid through the gap in the shield. Once outside of the shield, the wires run $\approx 15$ cm inside a stainless steel mesh sleeve. At this point, the 1 mm wires are joined to 75 $\mu$m twisted-pair PFA-insulated wire with superconducting crimped solder. The 75 $\mu$m wires run for ≈1 m inside hollow superconducting solder capillaries [271] to the input of the SQUIDs mounted on the 700 mK (Still) stage of the cryostat. The SQUIDs have an input inductance of 150 nH to match the calculated inductance of the pickup loop of $L_p = 95.5$ nH. Including the inductance of the wires, the total design inductance of the pickup loop circuit is ≈550 nH. However, measuring the inductance of the circuit yielded a value closer to $L_T \approx 3.3\,\mu$H, we discuss this further in Sec. D.3. The data presented in [72] was taken in a broadband readout configuration with no resonant amplifier.

### D.1.4 SQUID Setup And Readout

The first stage was readout with a Magnicon two stage SQUID current sensor [278, 436]. The SQUID is operated at a temperature of 870 mK and has typical flux noise floor of $0.6\,\mu\Phi_0^2/\text{Hz}$. The inductive coupling between the input coil of the SQUID and the SQUID is $M_{\rm in} = 2.52$ nH. The SQUID is operated with the Magnicon XXF-1 electronics in FLL mode with a SQUID flux to voltage conversion of $\partial V/\partial\Phi_S = 1.29\,\text{V}/\Phi_0$. In FLL mode, the response of the SQUID is linear over the dynamic range of the amplifier $\pm 11$ V or $\approx \pm 8.5\,\Phi_0$, however, this comes at the cost of limiting the bandwidth of the system to $\approx 6$ MHz.

[1] In the notation of [71], $\mathcal{G}_V = V_B/V$.

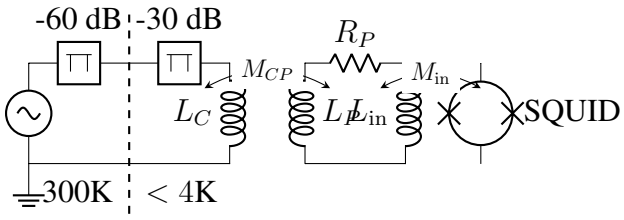



Figure D.3: A conceptual diagram of the ABRACADABRA-10 cm calibration circuit. The calibration loop, $L_C \approx 300$ nH, is concentric with the pickup loop, $L_P = 95.5$ nH. The circuit is plugged into the SQUID with input inductance $L_{\rm in} \approx 150$ nH. The parasitic resistance in the circuit is measured as $R_P \approx 13\,\mu\Omega$.

The output voltage of the SQUID was recorded with an AlazarTech 9870 8-bit digitizer [437]. To achieve the needed voltage precision we use the smallest available input range of $\pm 40$ mV which leads to a typical digitizer noise floor of $3.5 \times 10^{-9}$ mV$^2$/Hz. However, due to the large background below $\sim 20$ kHz (see Fig. D.4), we must first pass the signal through a 10 kHz high-pass filter to prevent railing the digitizer. Additionally, we use a 2.5 MHz anti-aliasing filter. The frequency response of these two filters is shown in Fig. D.2. These filters define the usable range of data for our axion search of 50 kHz – 3 MHz. It is worth pointing out that even though the gain is less than unity over the majority of our search range, both the signal and dominant noise is scaled by this gain, so the SNR is unchanged.

The digitizer is clocked to a Stanford Research System FS725 Rb frequency standard, with a ten-second Allan variance of $< 10^{-11}$.

### D.1.5 Calibration Circuit

We measure the end-to-end gain using a calibration system. It consists of the 0.5 mm diameter NbTi wire passing through the body of the toroid (i.e. in the magnetic field region), creating a 9 cm diameter loop concentric with and in the same plane as the pickup loop (see Fig. D.1b). This wire runs out of the detector shield as a twisted pair and then into an RG196 coaxial cable. This cable is connected to a 30 dB attenuator at the 4 K stage and then continues up to the top of the cryostat and through a BNC feed-through out of the vacuum region. During data taking, this BNC is left unplugged, and the attenuator contributes noise from a 50 $\Omega$ resistor at 4 K, which is well below our current noise level. We calculate the mutual inductance between the calibration and pickup loops both with an analytic calculation based on the geometry, as well as with a COMSOL [277] simulation. These values agree and predict a mutual inductance of $M_{CP} = 19.3$ nH.

While calibrating, we typically add an additional 60 dB of warm attenuation for a total of 90 dB of attenuation to get the signal to reasonable size. All attenuators are impedance matched at 50 $\Omega$.

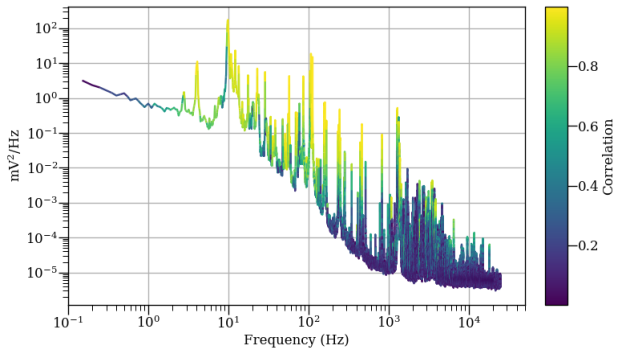



Figure D.4: Low frequency SQUID spectra from ABRACADABRA-10 cm taken with an accelerometer attached the 300 K plate. The spectrum is that of the SQUID output, with the degree of correlation with the accelerometer indicated by color (i.e. the correlation coefficient). The accelerometer begins to lose sensitivity above a few kHz, so it is not clear from this measurement how far up the correlation continues. These data were taken with a larger dynamic range on the digitizer, so have a relatively high ADC noise floor of $\sim 5 \times 10^{-6}$ mV$^2$/Hz. (Data taken without signal shaping filters.)

However, the output of the final attenuator is shorted by the calibration loop, which has an inductance of ≈300 nH; for frequencies below $\approx$ 30 MHz this causes it to behave as a current source driving a current through the calibration loop with amplitude independent of frequency. A wiring schematic of the calibration circuit is shown in Fig. D.3.

### D.1.6 Cryostat and Detector Suspension

The ABRACADABRA-10 cm detector is mounted inside an Oxford Instruments Triton 400 dilution refrigerator. It is mechanically supported by the detector suspension system. This consists of a 0.038" Kevlar thread which attaches to a vented bolt screwed into the center hole of the top aluminum mounting plate on ABRACADABRA-10 cm. The thread runs $\sim$ 1.5 m up through the various cold stages of the cryostat to a steel spring which supports the weight of the detector. The steel spring has a spring constant of $k \approx 20.4$ N/m and connects to a hook which is mounted about 1 m above the 300 K plate of the cryostat. The hook is at the top of a 1 m long vacuum tube which is rigidly connected to the 300 K plate. Due to the poor thermal conductivity of Kevlar, the spring and top end of the thread stay at 300 K while the bottom of the thread is cooled to $\lesssim$1 K with the detector.

The detector is thermalized to the coldest stage of the cryostat through four 10 mm wide 75 $\mu$m thick copper ribbons. Specifically, they are attached to the Mixing Chamber plate of the cryostat and then to the aluminum top plate of the detector. They are mounted with significant slack to minimize vibration through these ribbons. They are electrically isolated from the top of the detector

by using thin Kapton pads between the copper ribbon and the aluminum plate and connected with Nylon bolts.

The detector suspension system is designed to act like a pendulum which rolls off lateral vibration above frequencies of $f \approx 0.4$ Hz and in the vertical direction above frequencies of $f \approx 0.3$ Hz. The operating frequency of the pulse tube is 1.4 Hz and creates one of the main vibrational noise sources in the ABRACADABRA-10 cm data. It is clear from Fig. D.4, that even with this suspension system, vibrational noise still plays a significant role in the ABRACADABRA-10 cm backgrounds, and is a future path for potential improvement.

To improve the magnetic shielding of the detector, we wrapped the cryostat in MuMetal shielding. As MuMetal performs best at room temperature, we only wrapped the outermost vacuum vessel. The vertical walls of the vessel were lined both inside and out with a $200\,\mu$m thick layer. The bottom of the vessel was covered with a single layer on the inside. The top of the vacuum vessel and cryostat were not covered with MuMetal due to all the instrumentation and cryostat infrastructure. We measured the DC magnetic field attenuation ex-situ to be a factor of $\sim 5 - 10$.

## D.2 Data Collection Procedure

### D.2.1 Axion Search Data

In [71], the original proposal for a broadband search involved collecting time series data at a high sampling frequency continuously for months to years. However, this runs into practical disk space limitations. For example, one month sampled at 10 MS/s would fill ≈26 TB of disk space. This is maneagable, but would not scale well to a 1 GS/s sampling rate for a full year. However, this sort of sampling is not necessary for resolving ADM signals, where the expected signal width is given by $\Delta f/f \sim 10^{-6}$. Instead, we take an approach that maintains the required spectral resolution, while minimizing the required disk space.

For ABRACADABRA-10 cm, we sample continuously at 10 MS/s. Once samples are pulled from the digitizer, the data follows two processing paths: transforming and downsampling. First, the samples are accumulated into a 10 s buffer (of $10^8$ samples), which is then transformed via DFT [272] into a PSD. Once the next 10 s is available, it is transformed and its PSD is then averaged with the first, and so on. This builds up an averaged PSD, called $\bar{\mathcal{F}}_{10\mathrm{M}}$, which has Nyquist frequency of 5 MHz and frequency resolution of 100 mHz. This spectrum would be able to resolve axion signals down to $m_a \sim 100$ kHz with at least one bin width. After 80 averages (or 800 s), the average spectrum is written to disk and the averaging is reset. This level of averaging was chosen as a balance between storage space and being able to resolve time variation of background noise.

In parallel with this, the 10 MS/s time series is decimated by a factor of 10, to a 1 MS/s time

series. This data is then accumulated into a 100 s buffer – again of $10^8$ samples – then transformed with a DFT and converted into a PSD. In this way, we build up a second averaged PSD called, $\bar{\mathcal{F}}_{1\mathrm{M}}$, with a Nyquist frequency of 500 kHz and a frequency resolution of 10 mHz. This spectrum would be able to resolve signals down to $m_a \sim 10$ kHz with at least one bin width. After 16 averages (or 1600 s), the average spectrum is written to disk and the averaging is reset.

The data are then decimated by another factor of 10 and written directly to disk at a sampling rate of 100 kS/s. Offline, we take the time series data and transform it as one $2.45 \times 10^{11}$ sample long DFT to form a final spectrum, $\mathcal{F}_{100\mathrm{k}}$. Unlike the other spectra, $\mathcal{F}_{100\mathrm{k}}$ is not averaged over multiple integration periods, but is instead a single PSD with Nyquist frequency 50 kHz and frequency resolution of $\approx 408$ nHz. This spectrum could be used for searches for axion signals down to below 1 Hz, however, it is not used in the present analysis.

Each decimation step is done by first applying a top-hat filter with a 10-bin width, and then down-sampling by keeping every $10^{\mathrm{th}}$ filtered sample. This approach was chosen because it is fast computationally, though it is not quite optimal. We collected data from July 16, 2018 through August 14, 2018, accumulating a total exposure of $T = 2.45 \times 10^6$ s or $24.5 \times 10^{12}$ samples. In total, the data consist of 3065 independent $\bar{\mathcal{F}}_{10\mathrm{M}}$ spectra and 1532 $\bar{\mathcal{F}}_{1\mathrm{M}}$ spectra as well as a $2.45 \times 10^{11}$ continuous samples of 100 kS/s data. The total data footprint was about 3.8 TB for an average write rate of 12.4 Mbps – both of which are easily handled by a desktop PC.

### D.2.2 Magnet Off and Digitizer Noise Data

We also perform a Magnet Off measurement to understand backgrounds that are not correlated with the magnet. This data was collected with the exact same procedure and hardware configuration as the Magnet On data. Neither the cryostat, nor the SQUIDs were stopped in between measurements. We started collecting Magnet Off data within a few days of stopping the Magnet On run. We collected Magnet Off data from August 18, 2018 through August 27, 2018, for a total of $8.00 \times 10^5$ s of data.

We also collected $\approx 16$ h of digitizer noise data, with nothing plugged in to measure the noise level inherent to the ADC and computer.

### D.2.3 Data Quality

Figure D.5 shows examples of spectra for magnet on and magnet off data. There are a few features of these spectra worth discussing. The region below 100 kHz shows large noise spikes and a baseline increasing towards lower frequency. These spikes are generally too broad to be identified as ADM, but instead are incoherent noise backgrounds. We also observe that this noise is significantly reduced when the magnet is turned off. We interpret this noise as due to vibration of the

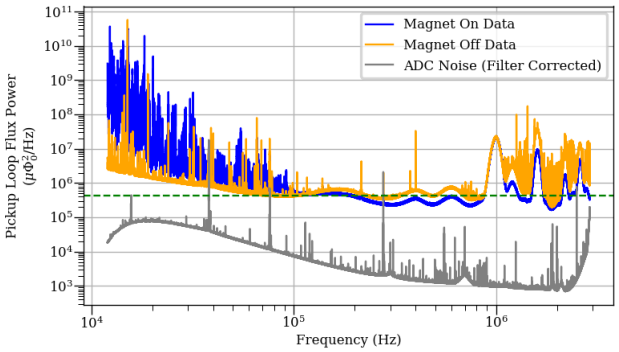



Figure D.5: Example $\bar{\mathcal{F}}_{10\mathrm{M}}$ SQUID spectra with magnet on (*blue*) and off (*orange*), along with the digitizer noise floor (*gray*). SQUID spectra are averaged over ≈9 h, digitizer data averaged over ≈16 h. The typical SQUID noise floor is shown in green dashed line. Note: The spectra were collected at different times and some of the transient noise peaks are not seen in all spectra.

detector. In particular, this appears to be the high frequency tail of the noise in Fig. D.4. The fact that it is reduced when the magnet is turned off implies that stray fields from the toroidal magnet are being seen by the pickup loop. We see that for $f > 100\,\mathrm{kHz}$ this noise becomes sub-dominant; however, it is likely that it continues to higher frequency. This will present a major challenge for future detector configurations, including those with resonator readouts, that hope to lower the noise floor by many orders of magnitude.

In the region from $100\,\mathrm{kHz} \lesssim f \lesssim 850\,\mathrm{kHz}$, the noise is mostly flat with a few small broad bumps and is approximately consistent with the expected SQUID flux noise floor. We see a slow variation in this noise level over the month of data taking, associated with variations in the noise level of the SQUIDs.

The region above $\sim 850\,\mathrm{kHz}$ shows two effects: broad bumps with $\sim 100\mathrm{kHz}$ widths and a forest of very narrow transient peaks. The broad peaks are due to an unknown and incoherent noise source that decreases our sensitivity in that frequency range. The origin of this noise will be the subject of future investigation, but for now we tolerate the decreased sensitivity. The forest of narrow transient peaks, on the other hand, present a larger problem. These peaks are $\lesssim$100 mHz wide and actually narrower than we expect for an ADM signal in this range. They are transient in time and appear to be correlated with working in the lab and with working hours. They were present for a portion of the time that we collected Magnet On data and all the time that we collected Magnet Off data. The transient nature and narrow width of these peaks imply that their source is likely from digital electronics turning on and off somewhere in, or near, the lab.

This transient noise was observed to be either present as a forest of many lines or completely absent. For our ADM search, the easiest approach was to use this fact to tag and eliminate the effected periods of time. This could be done reliably by eye, but we use a more quantitative

approach described in Sec. D.4.2. Though the lines only appeared at frequencies above $\sim 850$ kHz, we excluded all data from the tagged time intervals. In total, these cuts eliminated $\sim$30% of the exposure.

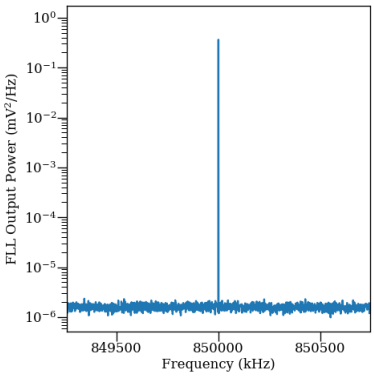
## D.3 Calibration

We quantify our detector response to a potential ADM signal, by performing a set of calibration measurements. Each measurement involves injecting a series of AC signals with known amplitude and frequency into the calibration system described in Sec. D.1.5. We compare the power measured by our readout circuit to the power expected from the flux through the pickup loop generated by the calibration loop.

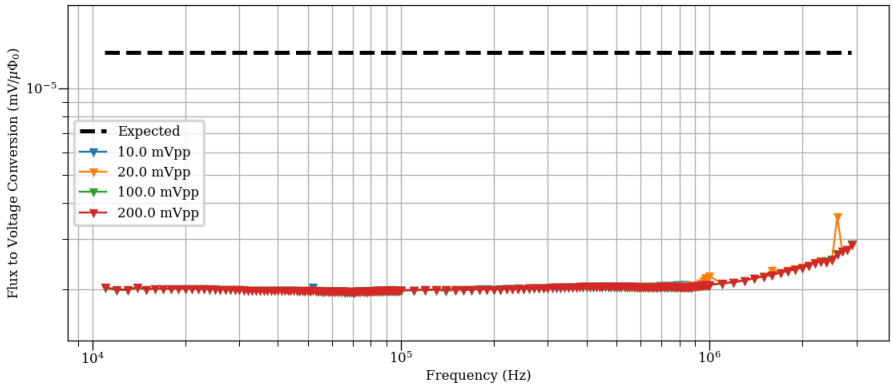
The input signal is generated by a Stanford Research Systems SG380 signal generator, locked to the same Rb frequency standard as the digitizer. The SG380 has very low phase noise and is able to output a tone with very long coherence time (longer than our measurement time), such that the resulting peak in the PSD was less than one frequency bin wide. For each amplitude and frequency, we perform a similar data collection to our axion search. We collect, transform and average 1 s buffers to form an averaged PSD. A zoomed example calibration line is shown in Fig. D.6a. The resulting peaks are typically only one bin wide. We measure the power in each calibration peak and compare this to the expected flux power generated by the calibration loop.

We perform this procedure for between 120 and 200 frequencies from 10 kHz to 3 MHz, and for four different input amplitudes: 10 mVpp, 20 mVpp, 100 mVpp and 200 mVpp. The resulting gain spectrum is shown in Fig. D.6b, and shows good agreement between the different input amplitudes. We perform the calibration measurement before the ADM search run, between the Magnet On run and the Magnet Off run, and again after the Magnet Off run. The resulting calibration curves were very consistent in time and did not depend on whether the magnet was on or off. For the present search, we determine our final calibration by interpolating the 200 mVpp data taken with the magnet on (red curve in Fig. D.6b).

From Fig. D.6b, we see that our measured gain is a factor of $\approx 6.5$ below what is expected based on the calculated circuit inductance. We tested each element of the calibration circuit and determined that the discrepancy came from the pickup loop side of the measurement. We determined the factor of $\approx$6.5 to most likely come from parasitic inductance in the readout circuit. This degrades the overall sensitivity of our axion search, and is the focus of future upgrades.




(a)




(b)

Figure D.6: (a) Example calibration peak at 850 kHz with 10 mVpp excitation and 90 dB of attenuation. Bin width is 1 Hz wide and all power is contained within a single bin. Output voltage is measured at the output of the amplifier electronics. (b) Measured detector response for four different input amplitudes taken with the magnet on. The measured gain is a factor of $\approx 6.5$ below the expected response (dashed line at the top). The outlier in the 20 mVpp spectrum is the result of a background line contributing power to the measured peak.

## D.4 Axion Search and Limit Extraction

For the present analysis, we restrict our axion search to the frequency range $75\,\mathrm{kHz} < f < 2\,\mathrm{MHz}$ or axion mass range $0.31 < m_a < 8.3\,\mathrm{neV}$. We therefore do not include the $\mathcal{F}_{100\mathrm{k}}$ spectrum in the rest of this analysis, as it has a Nyquist frequency of 50 kHz. We use the $\bar{\mathcal{F}}_{10\mathrm{M}}$ spectra to search the range from $500\,\mathrm{kHz} < f < 2\,\mathrm{MHz}$, and the $\bar{\mathcal{F}}_{1\mathrm{M}}$ spectra to search the range $75\,\mathrm{kHz} < f < 500\,\mathrm{kHz}$. In this way, a potential signal would be covered by at least 10 frequency bins at all frequencies. We further average the averaged spectra $\bar{\mathcal{F}}_{10\mathrm{M}}$ and $\bar{\mathcal{F}}_{1\mathrm{M}}$ to contain 3200 and 480 averages respectively. This decreases our ability to resolve time variations in our background noise to ≈9 h and ≈18 h, respectively. This step is not necessary for our analysis and is purely to decrease the computational resources required by a factor of $\approx 40$. After this, our 1 month of data collection is spanned by 75 $\bar{\mathcal{F}}_{10\mathrm{M}}$ spectra and 37 $\bar{\mathcal{F}}_{1\mathrm{M}}$ spectra.[2]

Our data analysis procedure closely follows the method introduced in [80]. Our expected signal is a narrow peak in the pickup loop PSD above the noise background, with a width $\Delta f/f \sim 10^{-6}$ arising from the ADM velocity dispersion. The challenge in a broadband search such as this is that we are scanning a large number of mass points without the benefit of being able to efficiently 'rescan' mass points with possible signal detections. As such, we need to be thorough with our statistical modeling, as at least some points are likely to populate the tails of any distribution. In this section, we describe the statistical modeling of our expected signal and background, as well as a data quality cut for tagging the periods of time when transient noise causes the data to look neither like signal nor background.

### D.4.1 Likelihood Analysis

The local ADM field can be thought of as arising from a partially coherent sum over a very large number of individual axion particles, where the phases of each particle are randomly distributed. As a result, the expected signal power in each frequency bin is drawn from an exponential distribution. When averaged over $N_{\mathrm{avg}}$ independent PSDs, the signal in each frequency bin $k$ will follow an Erlang distribution. When combined with background noise that is incoherent and Gaussian distributed in the time domain, the resulting PSD data is still Erlang-distributed [80]. Accordingly, for a single averaged PSD, our combined signal-plus-background model prediction in each bin is an Erlang distribution,

$$P(\bar{\mathcal{F}}_k; N_{\mathrm{avg}}, \lambda_k) = \frac{N_{\mathrm{avg}}^{N_{\mathrm{avg}}}}{(N_{\mathrm{avg}}-1)!} \frac{(\bar{\mathcal{F}}_k)^{N_{\mathrm{avg}}-1}}{\lambda_k^{N_{\mathrm{avg}}}} e^{-\frac{N_{\mathrm{avg}}\bar{\mathcal{F}}_k}{\lambda_k}}\,, \tag{D.5}$$

[2] We recycle the notation because we have only changed the number of spectra contributing to the average, but otherwise, they are conceptually equivalent.

with shape parameter $N_{\rm avg}$ and mean $\lambda_k = s_k + b$, where

$$s_k = \begin{cases} A \left.\frac{\pi f(v)}{m_a v}\right|_{v=\sqrt{4\pi f_k/m_a - 2}} & f_k > m_a/2\pi\,, \\ 0 & f_k \leq m_a/2\pi\,, \end{cases} \tag{D.6}$$

and $b$ is the expected background power. Here, $A$ denotes the combination of parameters that control the signal strength, defined in Eqn. D.3. We assume $f(v)$ is given by the Standard Halo Model (SHM), with velocity dispersion $v_0 = 220\,$km/s, and $v_{\rm obs} = 232\,$km/s the DM velocity in the Earth frame [273] and $\rho_{\rm DM} = 0.4\,$GeV/cm$^3$ [248, 249].

We build an analysis over a set of $\mathcal{N}$ averaged spectra $\bar{\mathcal{F}}_j$, each one an average over $N_{\rm avg}$ individual PSD. For example, the analysis of the $\bar{\mathcal{F}}_{\rm 10M}$ spectra, we have $\mathcal{N} = 75$ averaged spectra, where each averaged spectrum $\bar{\mathcal{F}}_j$ is an average over $N_{\rm avg} = 3200$ PSDs (with the possible exception of the final averaged spectrum which usually has fewer PSDs contributing). We search for an axion signal at mass $m_a = f/(2\pi)$, by restricting our search to a window containing frequency bins from $f_{k_i(m_a)} = m_a/(2\pi)$ to $f_{k_f(m_a)} = (1 + 4v_0^2) f_{k_i(m_a)}$ – approximately 8 times the width of the expected signal. Since incoherent background noise varies on frequency scales much larger than this, we can approximate the background noise level in this window as independent of frequency. We tested that our final results were insensitive to the precise choice of this window width. We account for long term variability in our noise levels by allowing the expected background level to vary from one averaged spectrum $\bar{\mathcal{F}}_j$ to the next; we denote $\mathbf{b} = \{b_1, b_2, \ldots, b_{\mathcal{N}}\}$ to be these background values, which we treat as nuisance parameters. The expected axion signal strength $A$ is constant across our data taking period and thus is the same for each $\bar{\mathcal{F}}_j$. For a given axion mass point $m_a$, signal strength $A$ and background values $\mathbf{b}$, we calculate the likelihood of our data $\mathbf{d}$:

$$\mathcal{L}(\mathbf{d}_{m_a}|A, \mathbf{b}) = \prod_{j=1}^{\mathcal{N}} \prod_{k=k_i(m_a)}^{k_f(m_a)} P(\bar{\mathcal{F}}_{j,k}; N_{{\rm avg},j}, \lambda_{j,k}), \tag{D.7}$$

where $k$ indexes the (windowed) frequency bins and $j$ indexes the different spectra. We allow for the generic possibility that each spectrum $\bar{\mathcal{F}}_j$ has a different number of averages, $N_{{\rm avg},j}$. This accounts for the spectra collected at the very end of the data taking period which have a different number of averages.

With the likelihood in Eq. (D.7), we perform a likelihood ratio test to search for a possible axion signal. To claim a detection, we place a $5\sigma$ threshold on the profiled likelihood ratio between the signal-plus-background and background-only hypotheses. We define a TS for discovery as

$$\mathrm{TS}(m_a) = 2\ln\left[\frac{\mathcal{L}(\mathbf{d}_{m_a}|\hat{A}, \hat{\mathbf{b}})}{\mathcal{L}(\mathbf{d}_{m_a}|A = 0, \hat{\mathbf{b}}_{A=0})}\right], \tag{D.8}$$

where $\hat{A}$ and $\hat{\mathbf{b}}$ are the values of $A$ and $\mathbf{b}$ which achieve the global maximum of the likelihood, and $\hat{\mathbf{b}}_{A=0}$ is the value which achieves the constrained maximization with $A = 0$. The maximization of $A$ is performed over a range including positive and negative parameter values, accommodating that a negative parameter value may provide the optimal fit to the data. If $\hat{A} < 0$, it is understood that the corresponding best-fit axion coupling is 0, as no value of $g_{a\gamma\gamma}$ could lead to negative-valued $\hat{A}$. The $5\sigma$ condition for discovery at a given $m_a$ is $\mathrm{TS}(m_a) > \mathrm{TS}_{\mathrm{thresh}}$, where

$$\mathrm{TS}_{\mathrm{thresh}} = \left[\Phi^{-1}\left(1 - \frac{2.87 \times 10^{-7}}{N_{m_a}}\right)\right]^2 \tag{D.9}$$

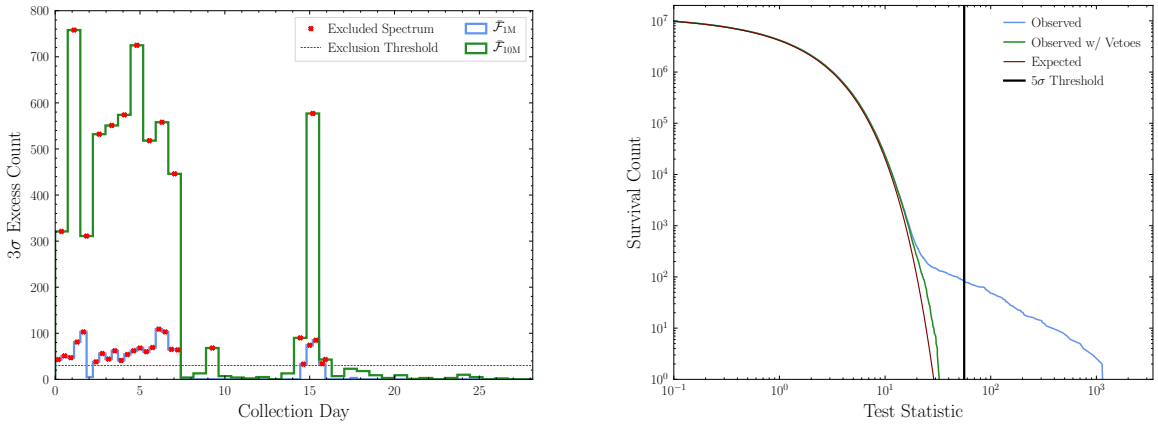
accounts for the local significance as well as the LEE for the $N_{m_a}$ independent masses in the analysis (here $\Phi$ is the cumulative distribution function for the normal distribution with zero mean and unit variance) [80]. For this analysis, $N_{m_a} \approx 8.1 \times 10^6$ between 75 kHz and 2 MHz (see below), and $\mathrm{TS}_{\mathrm{thresh}} = 56.1$.

Where we do not see a detection, we set a 95% C.L. limit, $A_{95\%}$, with a similar profiled likelihood ratio. To do so, we use the following test statistic for upper limits

$$t(m_a, A) = \begin{cases} 2\ln\left[\frac{\mathcal{L}(\mathbf{d}_{m_a}|\hat{A},\hat{\mathbf{b}})}{\mathcal{L}(\mathbf{d}_{m_a}|A,\hat{\mathbf{b}}_A)}\right] & A \geq \hat{A}\,, \\ 0 & \text{Otherwise}\,. \end{cases} \tag{D.10}$$

Here, $\hat{\mathbf{b}}_A$ is the background values that maximizes the likelihood for a given $A$. Using $t$, we can establish the 95% C.L. limit $A_{95\%}$ where $t(m_a, A_{95\%}) = 2.71$. In this limit setting procedure, it is necessary that $\hat{A}$ was allowed to be negative-valued if this provided the best fit in order to make an accurate calculation of $A_{95\%}$. In addition, we implement one-sided power-constrained limits [195], which in practice means that we do not allow ourselves to set a limit stronger than the $1\sigma$ lower level of the expected sensitivity band. This ensure that our constraints are statistically conservative while also addressing the possibility that $A_{95\%}$ is negative-valued.

Finally, we discuss the set of mass points over which we scan. In principle, we can search for an axion signal at any value between $75\,\mathrm{kHz} < m_a/(2\pi) < 2\,\mathrm{MHz}$. In our data, this range is spanned by $57.5 \times 10^6$ frequency bins and it might seem natural to search for an axion signal centered at each frequency bin. However, since each axion signal model is resolved by between 10 and 100 frequency bins, neighboring frequency bins would produce very strongly correlated results. Alternatively, a log-spaced set of $N$ masses such that $m_a^{i+1}/m_a^i = 1 + 4v_0^2$, would achieve a minimal coverage such that every frequency bin belongs to exactly one fit window. However, the spacing that achieves a set of statistically independent axion mass points is given by $m_a^{i+1}/m_a^i \approx 1 + 3v_0^2/4$ [80], which yields $N_{m_a} \approx 8.1 \times 10^6$ independent axion masses within our frequency range. For our search procedure, we therefore increase the granularity of the search and




(a) (b)

Figure D.7: (a) The number of $3\sigma$ excesses accounting for the look-elsewhere effect in each spectrum after vetoing the excesses that are present in the corresponding Magnet Off data. (b) The distribution of local TS values in the full month of analyzed data after removing periods of transient noise. In blue is the observed distribution of local TS values prior to vetoing the Magnet Off excesses. In green, the observed distribution of TS values after the Magnet Off veto. In red, the expected distribution under the null hypothesis. We see that after applying vetoes, there is excellent agreement down to very low survival counts, with no remaining $5\sigma$ excesses.

produce a set of log-spaced masses that obey $m_a^{i+1}/m_a^i = 1+v_0^2/2$. This eight-fold enhancement in the resolution of our tested masses, as compared to the minimal coverage set, results in overlapping signal windows of masses studied in our analysis and allows us to over-resolve a potential axion signal by a factor of two. This yields a total of $13.0\times10^6$ mass points to test, which is appropriately larger than the estimated number of independent mass points. We emphasize that this choice of mass points is not a fundamental limit on our mass resolution but is instead imposed merely by computational resources. In the event of an observed excess, we could fit a region around it with the mass floating in the fit.

For each mass studied, we also compute the expected sensitivity bands from the null-hypothesis models using the Asimov dataset procedure [172], following implementation outlined in [80].

### D.4.2 Quality Cuts

We can also use the analysis infrastructure described in the previous section to veto mass points where the condition of Gaussian-distributed incoherent noise does not hold as well as to create a quantitative data quality cut to identify periods of time with excess transient noise.

In order to tag periods of time with increased transient noise, we leverage the fact that the transient noise does not appear as a single peak, but instead as a forest of many correlated peaks.

Under the null hypothesis of a flat background, the survival function for the test statistic $t$ is given by

$$S(t) = 2\left(1 - \Phi\left(\sqrt{t}\right)\right) . \tag{D.11}$$

The presence of a true axion signal, would yield a small number of mass points that deviate from this distribution. But a violation of the null hypothesis of a flat background – e.g. due to a forest of correlated transient noise – would lead to a much larger deviation from this distribution.

For each $\bar{\mathcal{F}}_j$, we calculate the number of mass points with at least a $3\sigma$ excess within the time period covered by that PSD, accounting for the LEE. We find that the number of $3\sigma$ excesses, follows a clear bi-modal distribution with an obvious time correlation, see Fig. D.7a. This allows us to place a quantitative cut by requiring that an averaged PSD, $\bar{\mathcal{F}}_j$, have fewer than 30 mass points with a $3\sigma$ excess. This effectively eliminates periods of time with transient noise. We emphasize two points here: first, by placing this cut on a statistic which is calculated across a broad ranges of frequency, we do not produce a bias at any one mass point or range. Second, as we describe in the next section, a single axion signal would not be expected to create 30 mass points with $3\sigma$ excesses or larger. So, while exotic models with multiple axion could be affected by this cut, it would not present a problem for our baseline model of a single axion. This cut removes $\approx 30\%$ of our exposure.

Once we have removed periods of time with high transient noise, we remove individual mass points that have non-transient noise peaks – or are in other ways inconsistent with our null hypothesis of a flat background. We perform our axion discovery analysis on the Magnet Off data – where we expect no axion signal to be present. Any mass points that show LEE-corrected excess beyond $5\sigma$ in the Magnet Off data are vetoed. We consider these mass points to have poorly understood backgrounds where we do not have sensitivity to ADM. Out of $13.0 \times 10^6$ mass points, this requirement vetoed 18,733(6,651) points in the range $500\,\mathrm{kHz} < f < 2\,\mathrm{MHz}$ ($75\,\mathrm{kHz} < f < 500\,\mathrm{kHz}$) and implies a decrease in our signal recovery efficiency of 0.2%.

The axion search data, collected with the magnet on, showed 83(0) excesses with LEE corrected significance $\geq 5\sigma$, however they were all vetoed by cutting against the Magnet Off data. It is worth pointing out that the number of $5\sigma$ excesses in the axion search data was much smaller than in the Magnet Off data, due to the lower transient noise levels seen during that run. In Fig. D.7b, we show the distribution of TS values before and after the Magnet Off veto, compared against the distribution expected under the null hypothesis. The strong agreement between the theoretical expectation and the observed distribution after vetoes are applied demonstrates that the experimental backgrounds are well-modeled by the null hypothesis and ABRACADABRA-10 cm has strong discovery power under deviations from this theoretical expectation. In particular, in the results presented in [72], we find no significant excesses after vetoes are applied.

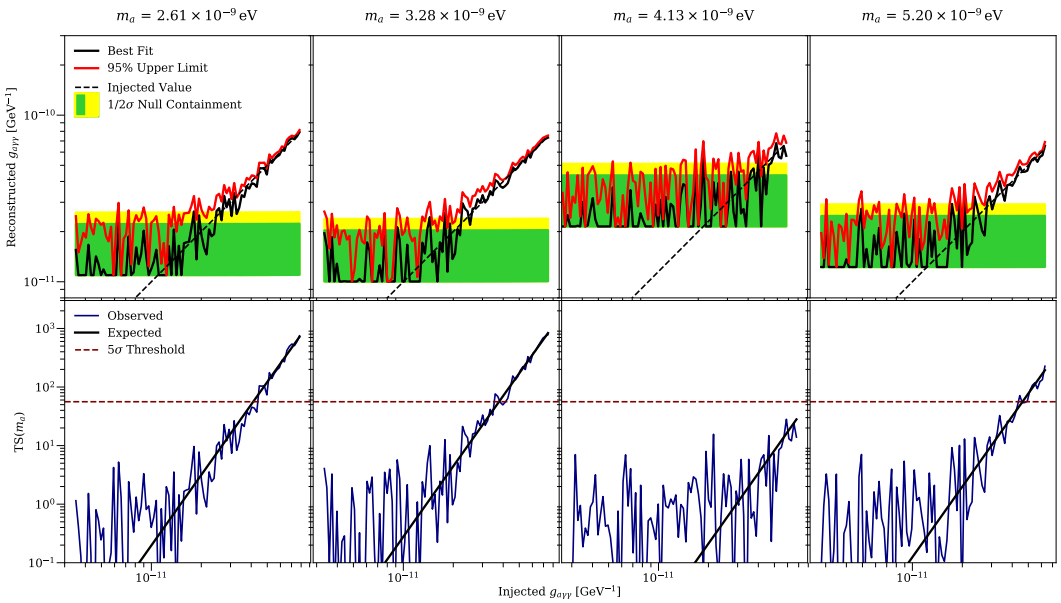



Figure D.8: (*Top row*) The recovered signal parameters as a function of the injected signal parameters in four Monte Carlo realizations with identical mean background levels. Green and yellow bands indicate the expected 1 and $2\sigma$ containment for the upper $95\%$ limit on the axion coupling under the hypothesis of no axion signal. (*Bottom row*) The observed and expected test statistic for discovery as a function of the injected signal strength. The dashed red line indicates the threshold for a discovery at $5\sigma$ significance accounting for the LEE, while the dashed black line indicates the upper $95\%$ limit on the observed test statistic under the null hypothesis.

### D.4.3 Recovering an Injected Signal

As a final test of our analysis procedures, we test that we are able to recover an injected signal and discover an axion at the claimed significance. This is crucial because as axion searches achieve greater sensitivity, there will be an inevitable trade-off between broadband and narrow-band coverage, and a claimed exclusion at a given $g_{a\gamma\gamma}$ will be used as justification to avoid re-scanning parameter space that has already been tested.

To test the discovery power of our analysis procedure, we generate Monte Carlo (MC) spectra characterized by a mean background level $\hat{b}$ and $\{N_{\mathrm{avg},j}\}$ averaged PSDs identical to that of the observed data. We then add an artificial axion signal with signal template set by the SHM, for a range of axion coupling strengths $A$. We perform our quality cuts on the individual spectra, then perform the joint analysis on the surviving spectra. We then evaluate the best-fit axion coupling and the $95^{th}$ percentile limit on that coupling as a function of the "true" axion coupling of the injected signal for each MC realization. Figure D.8 shows the resulting distributions of reconstructed $g_{a\gamma\gamma}$ and $t$-values for six axion masses.

Critically, the ability of our analysis procedure to accurately recover the correct axion parameters when allowable by the background level is unaffected by the quality cuts. This is most clearly seen in the bottom panels of Fig. D.8, which show the $t$-value as a function of the injected signal strength, as compared to the expected $t$-value as a function of the injected signal. We see strong agreement between the expected and observed test statistic when the signal is strong enough that we expect to be able to discover it (i.e., when the injected coupling lies above the null model containment bands). These examples also demonstrate that when our signal is not significant enough to be discovered, our limit-setting procedure is unaffected by the quality cuts.

## D.5 Conclusions

The successful run of the ABRACADABRA-10 cm experiment [72] introduced and validated several new techniques useful for constraining axion dark matter, including a toroidal magnet geometry sensitive to ADM at $m_a \lesssim 1\,\mu$eV, a broadband readout technique capable of handling the data-load required to study millions of axion masses simultaneously, a signal injection through a calibration loop to characterize this type of detector, and the first implementation of the broadband data analysis technique proposed in [80]. In this paper, we have described in detail the implementation and validation of these techniques, which lend additional to support to the results presented in [72].

Of greatest practical concern for the first results is the identification of the mismatch between expected and measured end-to-end gain, which we aim to rectify in the next data-taking run, and the mitigation of vibrational noise. We have also emphasized the statistical analysis employed to extract the first results. The goal of this analysis is to establish a sure footing for the presented statistical limits with a robust understanding of the exclusion limits. This is important as next generation experiments reach for ever higher sensitivities and re-scanning regions of parameter space becomes prohibitively time-consuming. In addition, the excellent performance of our data quality cuts will allow use of a blind analysis pipeline, which we expect to use in future runs.

ABRACADABRA-10 cm represents the first step in an experimental search program, which aims to ultimately be sensitive to ADM in the coupling range preferred by QCD axions. Future phases of ABRACADABRA will require larger magnets with higher fields, improved shielding, and strong mitigation of mechanical vibration. Augmenting the techniques described here with a resonant amplification readout and scan strategy will also greatly improve the sensitivity of a future full scale ABRACADABRA detector [71, 221, 275]. We have already begun engineering studies towards designing and building such a detector and ABRACADABRA-10 cm creates a strong foundation for this ongoing work.

# APPENDIX E

# Second Results from ABRACADABRA-10 cm

## E.1 Detector upgrade and electromagnetic simulations

The sensitivity of ABRACADABRA-10 cm to ADM is set by the coupling strength between the axion induced current $\mathbf{J}_{\rm eff}$ and the readout SQUIDs. This coupling can be conceptually split into two parts: the coupling between $\mathbf{J}_{\rm eff}$ and the pickup, and the coupling between the pickup and the SQUIDs. Before Run 2, the ABRACADABRA-10 cm detector was upgraded in two ways to increase each of these two coupling strengths.

The first step of the upgrade was the installation of the superconducting pickup cylinder. The pickup cylinder geometry more effectively cancels the flux induced by $\mathbf{J}_{\rm eff}$ and thus couples more strongly to it. The cylinder was constructed out of a 150 $\mu$m thick sheet of Nb wrapped around a PTFE tube, secured with Kapton tape. The resulting cylindrical pickup was 10 cm tall with a 5.1 cm diameter and centered vertically in the magnet bore. This is close to the maximum diameter that could practically fit. A 1 mm gap was left in the wrapping of the Nb sheet to prevent electrical contact and the formation of a complete loop. The PTFE tube was glued and clamped onto the magnet support structure inside the superconducting shielding can. From experience in Run 1, a strong mounting was critical to reducing relative motion between the pickup and magnet.

The second step of the upgrade was a replacement of the wiring between the pickup cylinder and the SQUID readouts. The new wiring – along with the new pickup cylinder – reduced the total inductance of the readout circuit, resulting in more current in the SQUIDs. The new wiring consists of 75 $\mu$m superconducting solid NbTi twisted-pair wires that are spot welded to two corners of the Nb sheet. Spot welding ensures a superconducting connection between the Nb sheet and the NbTi wires. We used a series of four spot welds on each corner for redundancy, in case of breakage during handling or due to differential thermal contraction. These wires were then taped to the PTFE cylinder with Kapton in order to minimize stress on the connections. The wires run $\sim$57.5 cm to the SQUID input. The wires are shielded in a superconducting capillary which extends from about 1 cm from the Nb sheet up to the point where the wires enter the SQUID shielding can. In addition

to providing electric shielding, this capillary also reduced the inductance per unit length. The inductance of the new cylinder and the readout wires are calculated using simulations in COMSOL Multiphysics to be 20 nH and 288 nH, respectively. This decrease in the readout inductance was confirmed with calibrations up to a factor $\sim$1.8.

## E.2 Active Feedback System

A major challenge encountered in Run 1 was low frequency vibrations converting stray fields from the magnet into low frequency noise. Since this noise is generally below the frequency ranges of interest, we were able to simply filter it and ignore it. In Runs 2 and 3, the higher gain of the upgraded detector amplified this vibrational noise enough to rail the SQUID amplifier. Because most of this noise was relatively slow – below $\sim$ 1 kHz – we installed an active feedback system to cancel it. We fed part of the output signal into the input of a Stanford Research Systems SIM960 analog PID controller. The output was filtered through a 1 kHz low-pass filter (LPF) and fed into the calibration loop via a 10 dB warm attenuator followed by 40 dB of cold attenuation, as sketched in Fig. E.1. The LPF guaranteed that the feedback system could not interfere with signals in our ROI, while the warm 10 dB attenuator reduced the power dissipated on the cold stages of the fridge.

Before Run 3, we added power combiners and power splitters to the active feedback circuit in order to better impedance match and isolate the various parts of the circuit. This improved our in situ calibration, as described below.

## E.3 Detector calibration

At a basic level, the ABRACADABRA-10 cm readout converts the flux through the pickup cylinder $\Phi_p$ to an output voltage from the SQUID amplifier, $V_{\rm SQUID}$. The detector calibration provides an end-to-end measurement of the detector response $\partial V_{\rm SQUID}/\partial \Phi_p$ to an axion-like signal across the full range of frequencies being searched. A schematic of the detector configuration for the Run 3 calibration can be seen in Fig E.1. We generate a fixed frequency signal of known amplitude using an Stanford Research Systems SG380 signal generator. This signal is attenuated by 93 dB before passing into the calibration loop in the detector. The current in the calibration loop generates a flux through the pickup cylinder, inducing a current and response in the readout circuit in the same way that an axion signal would, up to geometric factors.

The response of the system to a calibration signal can be written as

$$\frac{\partial V_{\rm ADC}}{\partial V_{\rm Sig}} = \frac{\partial V_{\rm ADC}}{\partial V_{\rm SQUID}} \frac{\partial V_{\rm SQUID}}{\partial \Phi_p} \frac{\partial \Phi_p}{\partial I_C} \frac{\partial I_C}{\partial V_{\rm Sig}} \tag{E.1}$$

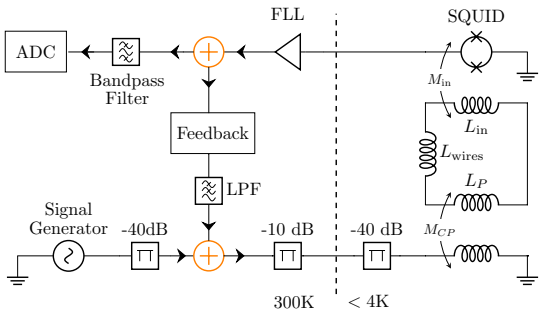



Figure E.1: ABRACADABRA-10 cm Run 3 calibration circuit diagram. A fake axion signal generated in the signal generator is attenuated by 93 dB (including 3dB loss in the combiner) before being coupled into the pickup cylinder analogously to an axion signal. The resulting power excess is readout on the SQUID and measured in the ADC digitizer. In Run 3, calibration is performed with the magnet turned on and the active feedback circuit running. During data taking, the signal generator is replaced with a 50 $\Omega$ terminator. The FLL feedback resistor and inductor are omitted for clarity.

where $V_{\rm ADC}$ is the RMS voltage measured by the digitizer, $I_C$ is the RMS current entering the calibration loop, and $V_{\rm Sig}$ is the peak-to-peak voltage output by the signal generator. The first and last terms in this conversion are determined by the warm electronics and cold attenuators, and can be measured directly, while the third term is the mutual inductance between the calibration loop and pickup cylinder, which is modeled in COMSOL. By dividing the measured end-to-end calibration by these three terms, we are left with the resulting flux to voltage conversion of the ABRACADABRA-10 cm readout circuit $\partial V_{\rm SQUID}/\partial \Phi_p$.

During Run 3, the calibration was performed in an identical configuration to data taking, namely with the magnet on and the AFS active. The resulting calibration can be seen in Fig. 6.2, and agreed very well with our calculated signal gain. The rolloff above $\sim$1 MHz corresponds to the finite bandwidth of the SQUID electronics. The Run 3 calibration circuit diagram is provided in Fig. E.1. During Run 2, we were not yet able to accurately calibrate the detector with the AFS system in place, which led to the decision to not present this data in our limits. Instead, the uncalibrated Run 2 data was used to tune our analysis procedure.

The flux noise determines the lower limit of our sensitivity. In particular, the 95% upper limit on $g_{a\gamma\gamma}$ under the null hypothesis scales like the square root of the flux noise, which is shown in Fig. E.2. In that figure we illustrate three different noise levels through the SQUID: (i) the measured magnet on flux, which is the relevant flux for the axion signal analysis; (ii) the magnet off flux; (iii) the flux measured in a similar SQUID that is not connected to the pickup loop circuit (labeled "open input"). The increased noise level in the magnet off data relative to the open input SQUID is likely the result of imperfect shielding, with environmental noise magnified by the pickup loop. On the other hand, when the magnet is on increased noise is apparent at low frequencies, which is the

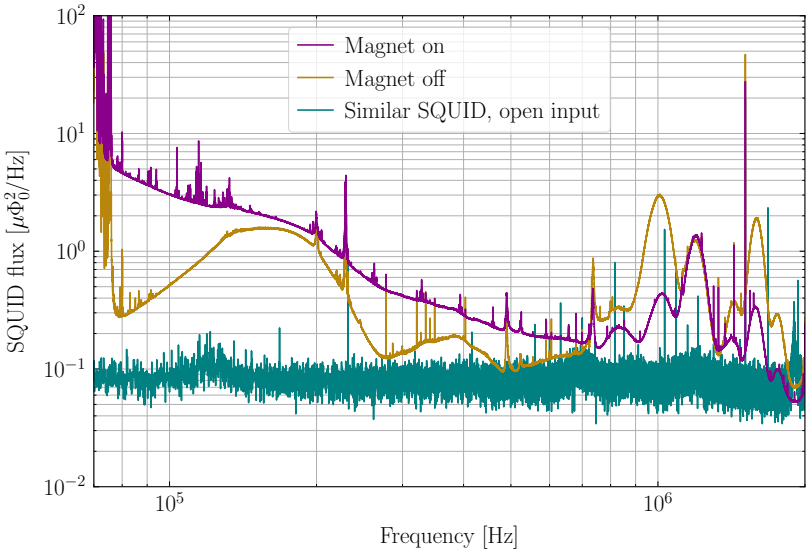



Figure E.2: The SQUID flux for Run 3 over the 70 kHz to 2 MHz frequency range at which we collect data. The magnet on noise level (magenta) is elevated compared to data taken with the magnet off (gold) primarily due to vibrating fringe magnetic fields. For comparison, the noise level from a similar SQUID without anything plugged into its input is plotted in teal.

result of the magnetic fringe fields giving frequency-dependent flux noise because of vibrations. Increasing the quality of the shielding and decreasing either the magnitude of the fringe fields or their vibrational coupling to the pickup loop would improve the sensitivity.

## E.4 Likelihood analysis

In this section, we describe the implementation of the analysis framework used to produce upper limits on $g_{a\gamma\gamma}$ and determine detection significances for potential excesses. We first define the profiled Gaussian likelihood used herein, followed by our procedure for cleaning the data to enable the removal of spurious excesses and confounding backgrounds. We then detail our treatment of a nuisance hyperparameter used to address potential systematics in the data, describe our results in terms of survival functions and upper limits on $g_{a\gamma\gamma}$, and demonstrate the efficacy of our analysis pipeline with injected signal tests. Though uncalibrated and not presented here, the Run 2 data was used while constructing the likelihood analysis framework. As such, we include it in the discussion below.

### E.4.1 Likelihood for Axion Signal Detection

The likelihood analysis utilized in this work is performed using the signal modeling formalism developed in [80], which was also used in studying the Run 1 results [72, 82]. Our starting point is a series of $N$ samples of the flux in the pickup loop $\{\Phi_n\}$, made over a collection time $T$ and at a sampling frequency $f = 1/\Delta t$ (such that $N\Delta t = T$). In the presence of an axion, this flux will receive a contribution from both the DM signal and any background. The mean expectation for the PSD at a frequency $f_k = k/T$ is

$$\Phi_k = A(g_{a\gamma\gamma})s_k(m_a) + \mu_k \,, \tag{E.2}$$

with $\mu_k$ is the mean expected background at this frequency. The signal strength parameter $A$ is given in (6.2) and is controlled by the unknown $g_{a\gamma\gamma}$, while $s_k$ is the signal template for a specific axion mass:

$$s_k(m_a) = \begin{cases} \frac{\pi f(v_\omega)}{m_a v_\omega} & f_k > m_a/2\pi \,, \\ 0 & f_k \leq m_a/2\pi \,. \end{cases} \tag{E.3}$$

Here, $v_\omega = \sqrt{4\pi f_k/m_a - 2}$ and $f(v)$ is the local axion speed distribution, which we take to be the Standard Halo Model with boost velocity $v_\odot = 232$ km/s and velocity dispersion $v_0 = 220$ km/s. To the extent the background is Gaussian in the time domain, the PSD formed from this data will be exponentially distributed, and the sum of multiple PSDs formed during data stacking will be Erlang-distributed. Nevertheless, in the limit of a large number of stackings, the Erlang-distribution becomes normally-distributed. For this reason we are justified in analyzing the data using a Gaussian likelihood.

In detail, the likelihood used is given by

$$\begin{aligned} &\mathcal{L}(d|m_a, A;\, \mathbf{a}, \sigma) \\ &= \prod_k \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(d_k - As_k - \mu_k(\mathbf{a}))^2}{2\sigma^2}\right], \end{aligned} \tag{E.4}$$

where $d_k$ is the average stacked data, $A$ and $s_k$ determine the axion signal as described above, $\mu_k$ is the background model (specified by parameters $\mathbf{a}$), and $\sigma$ is the standard deviation which we will treat as a nuisance parameter, and therefore estimate directly from the data. For a given axion mass $m_a$, the signal only has support over a narrow frequency range, and therefore we truncate the likelihood to $k$ values between $m_a(1 - (v_\odot + v_0)^2/2)/2\pi$ and $m_a(1 + 2(v_\odot + v_0)^2)/2\pi$. Over this narrow range, we find the background is adequately described by a first order polynomial, defined by the two-component vector $\mathbf{a}$ (c.f. Run 1 where the background in each signal window

was described by a flat white-noise spectrum). In summary, our likelihood is a function of five parameters: $m_a$ and $A$, which define the location and normalization of the signal, and nuisance parameters $\mathbf{a}$ and $\sigma$, which describe the mean size, slope, and fluctuations of the background.

Our goal is to use the likelihood in (E.4) to search for deviations from the background only distribution indicative of the presence of an axion. To do so we define the following test statistic (TS), which is a log-likelihood ratio of the signal and null models,

$$t(m_a, A) = 2\ln\left[\frac{\mathcal{L}(d|m_a, A;\, \hat{\mathbf{a}}, \hat{\sigma})}{\mathcal{L}(d|m_a, A=0;\, \hat{\mathbf{a}}, \hat{\sigma})}\right]. \tag{E.5}$$

Hatted background quantities are fixed to the value at which the likelihood attains its maximum value, given the specified signal values (i.e. for $A \neq 0$, $\hat{\mathbf{a}}$ and $\hat{\sigma}$ will in general take different values in the numerator and denominator). In other words, in defining this TS, we profile over the background nuisance parameters. The above test statistic is defined for any $m_a$ and $A$. For a given $m_a$, we then define the discovery TS as

$$\mathrm{TS}(m_a) = \max_{A} t(m_a, A). \tag{E.6}$$

The maximization of $A$ is initially performed over a range including positive and negative values, which is critical for the valid interpretation of TS as a $\chi^2$-distributed quantity under Wilks' theorem; intuitively, background fluctuations below the mean are just as likely as those above. However, as the presence of an actual axion signal will only result in positive spectral excesses, we take $\mathrm{TS}(m_a) = 0$ when the test statistic is maximized with $A < 0$. Accordingly, the discovery test statistic is expected to have the following asymptotic distribution

$$p(\mathrm{TS}) = \frac{1}{2}\left[\delta(\mathrm{TS}) + \chi^2_{k=1}(\mathrm{TS})\right], \tag{E.7}$$

which is expressed in terms of $\chi^2_{k=1}$, the probability density function for the $\chi^2$-distribution with one degree of freedom, and a Dirac $\delta$ function.

Using the test statistic in (E.5), we search for evidence of ADM with masses $m_a$ such that the signal would appear within the frequency range $f_{\min} = 100$ kHz to $f_{\max} = 2$ MHz. The local significance of any excess can be quantified by inverting the distribution in (E.7). In order to cover our entire frequency between $f_{\min}$ and $f_{\max}$, this search is performed in 11.1 million signal windows. As such, the local significance can be misleading and we should instead interpret results after accounting for the look-elsewhere effect (LEE). In doing so, we also need to account for the fact that due to the finite extent of the axion signal template $s_k$, nearby windows are correlated. We account for this self-consistently using the formalism developed in [80], from which we compute

the $N\sigma$ detection threshold, accounting for the LEE by

$$\mathrm{TS}_{\mathrm{thresh}}(N) = \left[\Phi^{-1}\left(1 - \frac{4v_0^2\,\Phi(N)}{3\ln(f_{\max}/f_{\min})}\right)\right]^2, \tag{E.8}$$

expressed in terms of $\Phi$, the cumulative density function of the zero mean and unit standard deviation normal distribution, with $\Phi^{-1}$ its inverse. Using this formalism, we find that the $5\sigma$ detection threshold accounting for the LEE is $\mathrm{TS}_{\mathrm{thresh}} \approx 55$.

A direct application of the formalism outlined thus far to the Run 2 and 3 data sets would result in a number of excesses at moderate or even high significance. Rather than interpreting this result as the discovery as a large number of ADM signatures, we interpret these as false positives sourced by coherent backgrounds that are not adequately captured by our null model. We employ two strategies for improving the background model in light of this. Firstly, we apply a data cleaning procedure in order to remove excesses inconsistent with that expected for ADM, for example transient spectral features or features that appear also when the magnet is off. Secondly, after applying our data cleaning pipeline, we modify our likelihood with a nuisance parameter tuned against the ensemble of observed significance values in the clean dataset; a data driven method for ensuring the quoted significance is consistent with the distributions observed directly in data. After applying both corrections factors, we find no significant excesses remain in our combined dataset. We now describe each of these strategies for improving our background treatment in more detail. We emphasize that all data cleaning and analysis procedures were developed and tested on the 10% of the Run 2 data which was unblinded and was then applied identically to the Run 3 data. The 10% was obtained by subdividing the full Run 2 dataset into 1,000 frequency subsets of equal size, and then taking the first 10% of each subset in frequency.

### E.4.2 Data Cleaning Procedure

Many of the excesses present in the uncleaned data are characterized by a narrow spectral feature, often present in a single frequency bin. The features often drift throughout our collection time, and appear at regular frequency intervals. Although such features are inconsistent with the axion signal expectation, which should be distributed over several frequency bins, such narrow features are far more consistent with our signal model than our linear background model, and therefore result in high-significance $\mathrm{TS}$ values.

A notable example is the background resulting from AM radio broadcasts: these manifest as large excesses at uniform 10 kHz intervals from 560 kHz to 1.60 MHz. We identify these AM radio signals in our data and remove them with a mask of width 15 Hz centered on the radio signal peak, beyond which the radio signal falls below our noise floor. In other cases the origin of the features

is unclear, although the fact that many appear at 50 Hz intervals suggests a universal environmental origin. Regardless, we remain agnostic to their origin and instead remove them using a data-driven procedure we now outline.

For each frequency bin, we identify a single-bin excess as follows. We determine the mean and standard deviation of the data on either side of the bin of interest; in particular, we use the 10 bins on both sides, ignoring the immediately adjacent frequencies. We then use these results to calculate the significance of the data observed in the bin of interest. We repeat this procedure for each frequency bin in each independently collected dataset, i.e. the data before it is stacked. If the bin of interest attains a significance $\delta\chi^2 > 100$ in any one dataset, then it is flagged for masking. If the bin is not flagged by this procedure, we then stack the dataset and repeat this procedure once more. If after stacking, the bin now has $\delta\chi^2 > 35$, then it is again flagged. For all flagged bins, we mask the 21 frequencies centered on the bin of interest. The motivation for considering the individual and stacked datasets is to identify both excesses that drift with time and also those that are only significant in the stacked data where we perform our fiducial analysis.

After the single-bin spikes have been identified and removed, we perform an initial analysis of the data. We analyze the $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) data providing a frequency resolution of $0.1$ $(0.01)$ Hz for axions which would produce a signal in the 500 kHz - 2 MHz (50 - 500 kHz) frequency range. We stack the Run 2 $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) data, which consists of 1460 (700) spectra, into 20 subintervals, each of which are initially analyzed independently. For each axion mass, each subinterval is analyzed independently, with the 50% of subintervals which realize the smallest values of the TS for discovery and any additional subintervals which have $\mathrm{TS} < 9$ accepted. The accepted subintervals are then stacked into a single dataset and analyzed. An analogous procedure is applied to the Run 3 data, where the $\bar{\mathcal{F}}_{\rm 10M}$ ($\bar{\mathcal{F}}_{\rm 1M}$) data, consisting of 1364 (682) spectra, are divided into 22 subintervals. This TS filtering procedure was implemented in order to mitigate the impact of transient excesses that imitate an axion signal in some of the subintervals and might produce a spurious excess if included in the stacked data. With this exclusion criteria, under the null, each spectrum is expected to be excluded with probability $0.1\%$, making this a relatively conservative exclusion criterion, although it does have the effect of somewhat weakening the detection sensitivity of the analysis. The binned data acceptances for the complete Run 3 analysis are shown in the left panel of Fig. E.4, which show that all data is accepted into the stacked analysis data for the overwhelming majority of mass points.

Finally, we perform a series of vetoes aimed at removing any remaining excess which have their origins in unmodeled backgrounds or instrumental effects. We directly analyze stacked, unfiltered $\bar{\mathcal{F}}_{\rm 10M}$ and $\bar{\mathcal{F}}_{\rm 1M}$ data which are collected with the magnet off. Since no axion detection can be made with the magnet off, any mass points which are excesses at $\mathrm{TS} > 16$ in both the magnet-on and magnet-off data are vetoed.

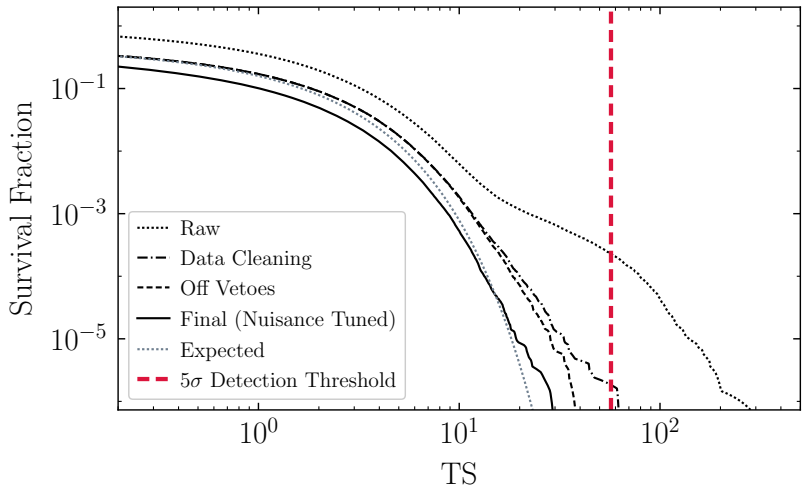



Figure E.3: As in Fig. 6.3, but evaluated on the 10% of unblinded Run 2 data against which we calibrated our analysis procedure.

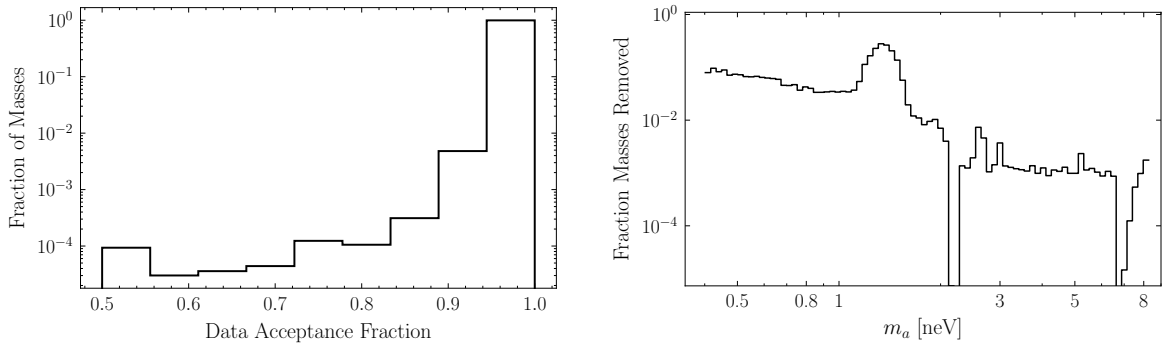



Figure E.4: (*Left*) The histogrammed data acceptance fraction under the data filtering over all masses analyzed in Run 3 data. (*Right*) The fraction of masses removed by magnet off vetoes as a function of frequency in Run 3 data. The acceptance fraction is determined within 100 log-spaced bins between the minimum and maximum axion masses within our analysis range. Note that while we display the Run 2 results, those were used only to develop our analysis protocols and not in the physics analysis.

### E.4.3 Nuisance Parameter Correction

After applying both our individual bin flagging and vetoing procedures, the remaining dataset is designated clean. Nevertheless, the distribution of TS values remains inconsistent with that expected for the asymptotic one-sided $\chi^2$ distribution given in (E.7), indicative of further background mismodeling. To resolve this, we implement an additional nuisance parameter correction to our likelihood.

In detail, we modify the likelihood and TS with additional nuisance parameters $A_m$ and $\sigma_{A_m}$ as follows,

$$\mathcal{L}(d|m_a, A;\, \mathbf{a}, \sigma, A_m, \sigma_{A_m}) = \mathcal{N}(A_m|0, \sigma_{A_m}) \prod_k \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(d_k - (A + A_m)s_k - \mu_k(\mathbf{a}))^2}{2\sigma^2}\right] \tag{E.9}$$

$$\mathrm{TS}(m_a|\sigma_{A_n}) = 2\ln\left[\frac{\max_A \mathcal{L}(d|m_a, A;\, \hat{\mathbf{a}}, \hat{\sigma}, \hat{A}_m, \sigma_{A_m})}{\mathcal{L}(d|m_a, A=0;\, \hat{\mathbf{a}}, \hat{\sigma}, \hat{A}_m, \sigma_{A_m})}\right]. \tag{E.10}$$

The index $m$ indicates that the nuisance parameters depend on the signal window under consideration.

By construction, the additional nuisance parameter is – up to a penalty factor – fully degenerate with the signal. This allows the background model the flexibility to fit signal-like excesses, but at the cost of a Gaussian penalty factor given by $\mathcal{N}(A_m|0, \sigma_{A_m})$, which is a zero mean normal distribution of width $\sigma_{A_m}$ evaluated at $A_m$. The magnitude of this penalty is controlled by the hyperparameter $\sigma_{A_m}$, which can be chosen to ensure the above TS has the expected asymptotic distribution. To be specific, we determine $\sigma_{A_m}$ for each mass (indexed by $k$) by tuning the observed distribution $\mathrm{TS}(m_a|\sigma_{A_m})$ against the expected distribution in the vicinity of the mass point of interest. We consider the ensemble of the discovery test statistics belong to the nearest 94,723 mass points, not including: the mass point of interest; the five nearest mass points above and below the mass point of interest; or any mass points that are vetoed by comparison with the magnet off data. We then tune the value of $\sigma_{A_m}$ to its minimum value such that there are only three discovery test statistics in excess of 16 within the ensemble, which would be expected if the discovery test statistics were half-chi-square distributed. The nuisance hyperparameter $\sigma_{A_m}$ translated into an effective nuisance parameter $g^{\mathrm{nuis}}_{a\gamma\gamma}$ is presented in Fig. E.5, and can be understood as an effective floor for our limit-setting power that competes with the statistical noise floor set by the background strength.

We note that many of the procedures required to fix the hyperparameter can be performed analytically. As the log-likelihood is approximately quadratic around its maximum, $\hat{A}$, near the

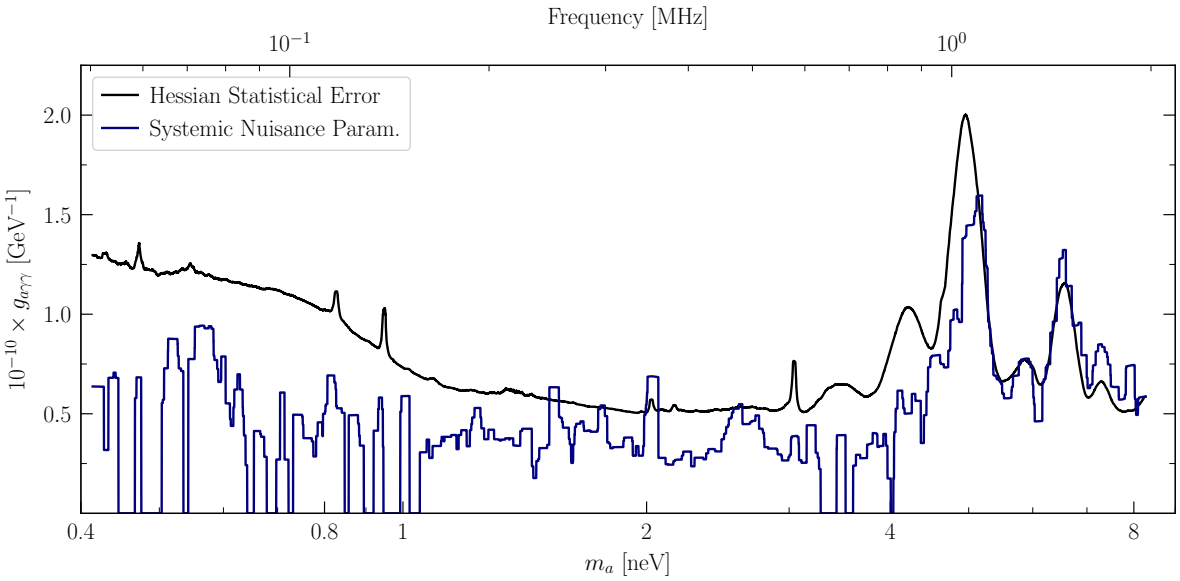



Figure E.5: The hyperparameter, $\sigma_A$, converted to the units of $g_{a\gamma\gamma}$, for the systematic nuisance parameter, $g^{\text{nuis}}_{a\gamma\gamma}$, as a function of axion mass (labeled Systematic Nuisance Param.). We compare the systematic nuisance hyperparameter to the statistical uncertainties (labeled Hessian Statistical Error), which are computed from the Hessian for the log-likelihood without systematic uncertainties about the best-fit axion coupling, $\hat{g}_{a\gamma\gamma}$.

maximum we have

$$t(m_a, A) = \text{TS}(m_a)\left[1 - \left(\frac{A - \hat{A}}{\hat{A}}\right)^2\right], \tag{E.11}$$

where at this stage we do not yet zero out test statistics associated with negative best-fit signal amplitudes. When including the correcting nuisance parameter, the distribution becomes

$$\begin{aligned} &t(m_a, A, A_m|\sigma_{A_m}) \\ =&\text{TS}(m_a)\left[1 - \left(\frac{A + A_m - \hat{A}}{\hat{A}}\right)^2\right] - \left(\frac{A_m}{\sigma_{A_m}}\right)^2, \end{aligned} \tag{E.12}$$

with the final term arising from the Gaussian penalty. We can now define a new test statistic for discovery including the background signal nuisance parameter as

$$\begin{aligned} \text{TS}(m_a|\sigma_{A_m}) =&\text{max}_{A,A_m} t(m_a, A, A_n|\sigma_{A_m}) \\ &-\text{max}_{A_m} t(m_a, A = 0, A_n|\sigma_{A_m}). \end{aligned} \tag{E.13}$$

Using this, for a given $\sigma_{A_m}$, the new test statistic for discovery with the nuisance background signal can be directly constructed from the test statistic without the nuisance background signal.

In particular, since (E.13) involves only maximizations of a quadratic function, the result is given by

$$\mathrm{TS}(m_a|\sigma_{A_n}) = \frac{\mathrm{TS}(m_a)\hat{A}^2}{\hat{A}^2 + \mathrm{TS}(m_a)\sigma_{A_n}^2}, \tag{E.14}$$

which has the effect of decreasing the computed TS. As before, we then zero out $\mathrm{TS}(m_a|\sigma_{A_n})$ when the best fit signal strength parameter $\hat{A}$ is negative.

### E.4.4 Survival Functions, Unvetoed Excesses, and Limits

The analysis procedure was tuned on 10% of the Run 2 data. Once validated, to the full Run 3 dataset, which had remained blinded. The survival function evaluated at various stages of our analysis procedure, realized on the 10% of Run 2 data used for tuning, is shown in Fig. E.3. Approximately 10% of masses in Run 2 and 5% of masses in Run 3 are removed by the peak exclusion and vetoing procedure, with the fraction of masses removed as a function of frequency shown in Fig. E.4.

Even after the nuisance parameter tuning, there remain some discrepancies between the observed and expected survival functions at moderate values ($\mathrm{TS} > 16$) of the test statistic. In particular, there are a small number of mass points which have TS in excess of our $5\sigma$ LEE threshold in Run 2 data. All of these excesses occur in nearby frequencies, associated with a transient, and relatively broad, spectral feature which is shown in Fig. E.6. Further, the mass points which are high significance excesses in the Run 2 data are not significant in the Run 3 data. Accordingly, we do not consider these excesses to represent credible detections.

The independent Run 3 limits are shown with and without the tuned nuisance parameter in Fig. E.7.

### E.4.5 Injected Signal Tests

To further validate the robustness of our analysis framework, we can inject a synthetic signal into the data and confirm that: (1) we are able to recover the signal strength, when expected; and (2) our limits will not exclude an injected signal. To perform this test, we select five representative mass points and inspect the real data in the vicinity of the expected location of an injected signal. We generate independently drawn axion signals at a range of axion couplings strengths which we add on top of each of the spectra collected in Run 3. We then apply our analysis framework, adopting the tuned value of the nuisance parameter that was previously determined from the real data in the vicinity of the injected signal location, and evaluate the best-fit axion coupling, the $95^{\mathrm{th}}$ percentile upper limit on that coupling, and the detection significance as a function of the true axion coupling of the injected signal. As a further test of the performance of our analysis framework, for each of

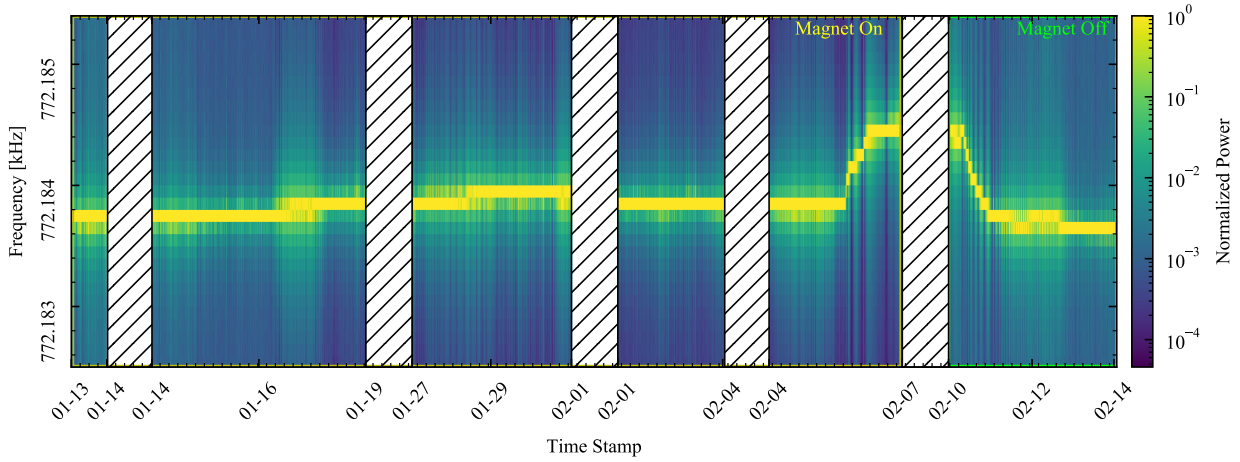



Figure E.6: The time evolution of the broad excess that is associated with the putative signal candidate in the Run 2 data that survived all analysis cuts. The excess persists after the magnet is turned off and evolves in frequency, indicative of a background source. The magnet off veto did not anticipate this level of time evolution and so did not remove these excesses. Since this feature was found after unblinding, we report it here but do not consider it to be a credible axion detection.

the five mass points, we fit the Run 3 data under the null model. We then generate Monte Carlo data under null model fits and repeat our procedure of injecting and analyzing, allowing us to compare the analysis of signal injections on real data with expected performance of the analysis framework under the null model. With the exception of the tuned nuisance parameter, which we continue to keep fixed at its value determined from the real data, this represents an entirely self-contained test of the analysis procedure.

The results of these tests are shown in Fig. E.8. Critically, our analysis procedure is able to place a robust $95^{\text{th}}$ percentile upper limit which does not exclude the true coupling strength at which the signal is injected more often than would be expected and accurately recovers the correct axion parameters at a detection significance within the simulated expectations. We also briefly comment on the somewhat jagged nature of the detection significance in the real data as a function of the injected signal strength. These features are a product of the filtering included in our analysis procedure which removes at most 50% of the spectra in $\sim$5% subintervals if those subintervals have a detection significance in excess of $3\sigma$. This has the effect of somewhat weakening the detection significance in discretized steps and also slightly biases the $95^{\text{th}}$ percentile limit to slightly lower values. This bias is removed using a TS-dependent correction of at most 8% that is incorporated in our limits.

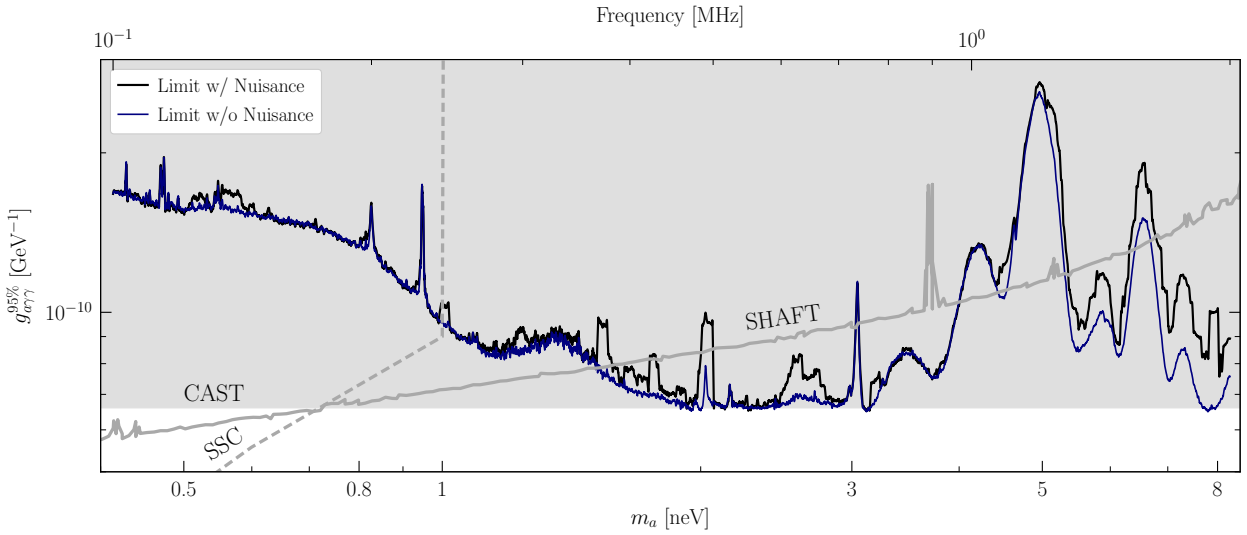



Figure E.7: A comparison of our fiducial limits that include a nuisance hyperparameter correction (black) and those without any correction (blue). Limits set with the nuisance hyperparameter are slightly weaker, but the features and limit-setting power are broadly similar. The figure is smoothed for clarity.

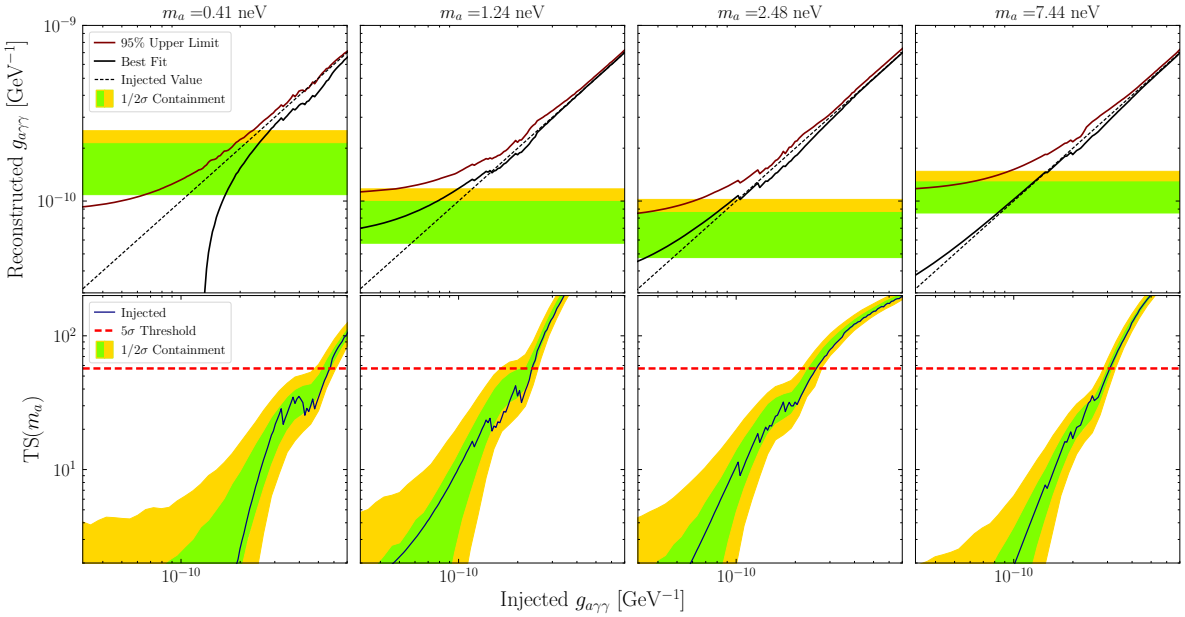



Figure E.8: (*Top row*) The best fit and 95% upper limit on the recovered signal strength as a function of the injected signal strength at five mass points evaluated on the real Run 3 data. The results are compared to the $1\sigma$ and $2\sigma$ expectations for the $95^{\text{th}}$ percentile upper limit under the hypothesis of no axion signal as determined by 2560 Monte Carlo (MC) realizations of the null model fits to the real data at each injected signal strength. (*Bottom row*) In black, the recovered detection test statistic for the signal injected in the real data as a function of injected signal strength. The dashed red line indicates the threshold for a $5\sigma$ detection significance account for the look-elsewhere effect while the green and yellow bands indicate the $1\sigma$ and $2\sigma$ expectations for the detection significance determined from 2560 MC realizations of the null model combined with appropriately varied injected signal strength.

# APPENDIX F

# Radio Searches for Axion Dark Matter

## F.1 GBT Results for the Galactic Center, M31, and M54

In this section we present results for analyses of GBT data for the following observation targets: the GC, the Andromeda galaxy (M31), and the globular cluster M54. The M54 and M31 are complementary targets to the nearby INS targets and the GC. Like the GC observations, the M54 and M31 observations also search for axion-photon conversion from a population of NSs in DM-dense environments. M31 has the advantage of being a large and nearby galaxy, similar to our own Milky Way. M54, on the other hand, is observed to lie at the center of the Sagittarius dwarf galaxy. It is also relatively nearby at a distance around 25 kpc from the Sun. Note that the velocity dispersions of the DM in Sagittarius and the NSs in M54 are $\sim$10 km/s, which increases the gravitational capture cross section for the NSs pulling in ambient DM. On the other hand, the DM density profile of Sagittarius is especially uncertain given its tidal interactions with the Milky Way (see [104] for details).

The parameters of the GBT observations of these targets are summarized in Tab. F.1. These observations proceeded analogously to the INS observations, with a few minor differences. The GC and M31 observations were performed with one VEGAS spectrometer across the L-band (mode 2), while the M54 observation used three VEGAS spectrometers (mode 4). For the GC observations, ON and OFF locations were separated by $2.5^\circ$ since the signal is expected to be spatially extended in this case. Otherwise, the ON and OFF locations were separated by $1.25^\circ$. Note that we also performed an off-center GC observation, in addition to an on-center observation, because of the possibility of a higher signal-to-noise ratio by looking slightly away from the very bright center of the galaxy. The angular positions of all of the targets are well known at the accuracy needed for the GBT at these wavelengths, but for the GC we chose Galactic coordinate $(\ell, b) = (359.9443^\circ, -0.04614^\circ)$ for the center observation and $(\ell, b) = (359.9996^\circ, 0.9958^\circ)$ for the off-center observation. The data were saved in exposures of 0.5 seconds for observations of all other targets in Tab. F.1.

We note that the GBT observations took place in two sessions, and during each session calibrator observations were also performed. The calibrators were 3C286 for the first session, which included the GC observations and M54, and 3C48 for the second session, which included M31 and the INS targets discussed in the main text. The calibrators were observed for approximately two minutes in each of the modes used in that session.

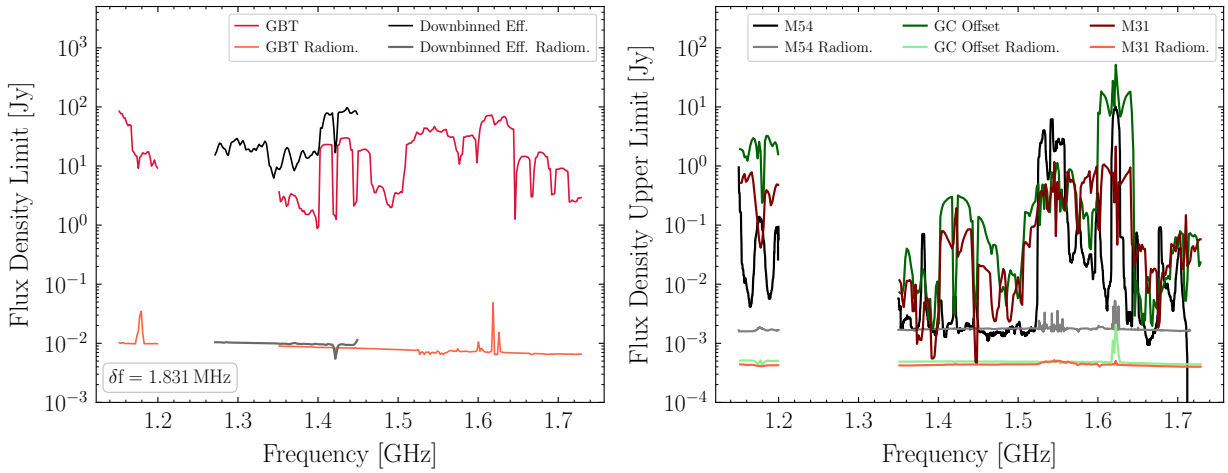



Figure F.1: (*Left*) A comparison of the 95% upper limits of the flux density spectra measured with our windowed analysis for the GBT and Effelsberg observations of the Galactic Center and radiometer expectations. Data are analyzed at an approximately 1.831 MHz frequency resolution corresponding to the fiducial resolution for the GBT analysis. Although the Effelsberg data is consistent with the radiometer expectations at its original resolution, when down-binned to the GBT resolution, it demonstrates similar incompatibility with the radiometer expectations. (*Right*) The 95% upper limits on the signal flux for the indicated sources from the GBT observations. These signal flux limits are compared to the expected flux density limit appropriately computed from the radiometer equation. The analysis is performed at the fiducial analysis bandwidth, see Tab. F.1.

For the population analyses with GBT data we down-bin the data to the level $\delta f/f \sim 10^{-3}$ to account for the fact that we are searching for emission from the whole population of NSs, each of which is Doppler-shifted with respect to the intrinsic frequency [104]. The M54 data is down-binned slightly less to account for the smaller velocity dispersion in the globular cluster compared to the GC and M31 [104]. This is a qualitatively different approach to that taken with the Effelsberg data, which is at higher frequency resolution, because with the GBT data from the populations we do not have high enough frequency resolution to search for the brightest converting NS. Instead, we search for the broader emission from the whole population of NSs. The frequency bins $\delta f_{\rm fid}$ used in the analyses are given in Tab. F.1.

The 95% one-sided upper limits on the line flux densities from these GBT observations are

| Target | $t_{\rm exp}$ [min] | $\delta f_{\rm obs}$ [kHz] | $\delta f_{\rm fid}$ [kHz] | type |
|---|---|---|---|---|
| GC | 30.0 | 92.0 | 1831.0 | pop. |
| GC (off-center) | 30.0 | 92.0 | 1831.0 | pop. |
| M31 | 30.0 | 92.0 | 1831.0 | pop. |
| M54 | 30.0 | 5.7 | 114.4 | pop. |

Table F.1: As in Tab. 7.1 but for the GBT observations of the GC, M31, and M54.

shown in Fig. F.1. Note that these limits used the frequency-down-binned data, with bin widths $\delta f_{\rm fid}$. Just as in Fig. 7.1, we also show the expectations for each observation from the radiometer equation, which only accounts for statistical uncertainties.

In the left panel of Fig. F.1, we show the GBT GC flux density limits. These limits are significantly higher than the radiometer expectations, indicating that the uncertainties, which are determined in a data-driven fashion by fitting the null model to the sideband data, are predominantly systematic and not statistical. This may seem surprising, when considering that in Fig. 7.1 the Effelsberg constraints were comparable to the radiometer expectations. However, it is important to remember that the bandwidths $\delta f_{\rm fid}$ for the Effelsberg analyses in Fig. 7.1 are significantly narrower than those that go into the left panel of Fig. F.1 for the GBT (7.32 kHz for L-band Effelsberg in Fig. 7.1 versus 1831 kHz for the GBT in Fig. F.1). Indeed, in Fig. F.1 we also show what happens if we down-bin the L-band Effelsberg data to 1831 kHz before performing the analysis; in this case, the results are similar to those from the GBT. We thus see that as the bandwidths are increased, the statistical uncertainties become less important relative to the systematic uncertainties. This is consistent with the expectation that the statistical uncertainties in the individual frequency bins shrink with more data, which is acquired by increasing the bin widths, while the systematic uncertainties need not change with bin width.

In the right panel of Fig. F.1 we show the flux density upper limits from the GC offset, M31, and M54 GBT observations, compared to the radiometer equation expectations. Note that these analyses use the $\delta f_{\rm fid}$ bin widths given in Tab. F.1. In a narrow frequency range around around 1.35 - 1.5 GHz, all of the upper limits become comparable to the radiometer expectation, while above 1.5 GHz, all of the upper limits become significantly weaker than expected from statistical uncertainties only. However, this is precisely what is expected from the known sources of RFI that affect GBT.[1] Thus, we conclude that the loss of sensitivity above $\sim$1.5 GHz is likely due to sources of RFI that induce variance in the null-hypothesis model.

For the purposes of additional illustration, we depict the raw, uncalibrated Effelsberg L-Band data between 1.3-1.45 GHz and the raw, uncalibrated GBT data from the INSs RX J0720 and RX

[1] See, *e.g.*, publicly available RFI archives

J0806, with associated flux density limits, in Fig. F.2. The data and the limits are presented without down-binning or smoothing in order to better demonstrate the impact of RFI. Note that the RFI may lead to decreased sensitivity at isolated mass points, but on the other hand with the Off vetoing procedure we find that RFI does not induce any false-positive axion signals.

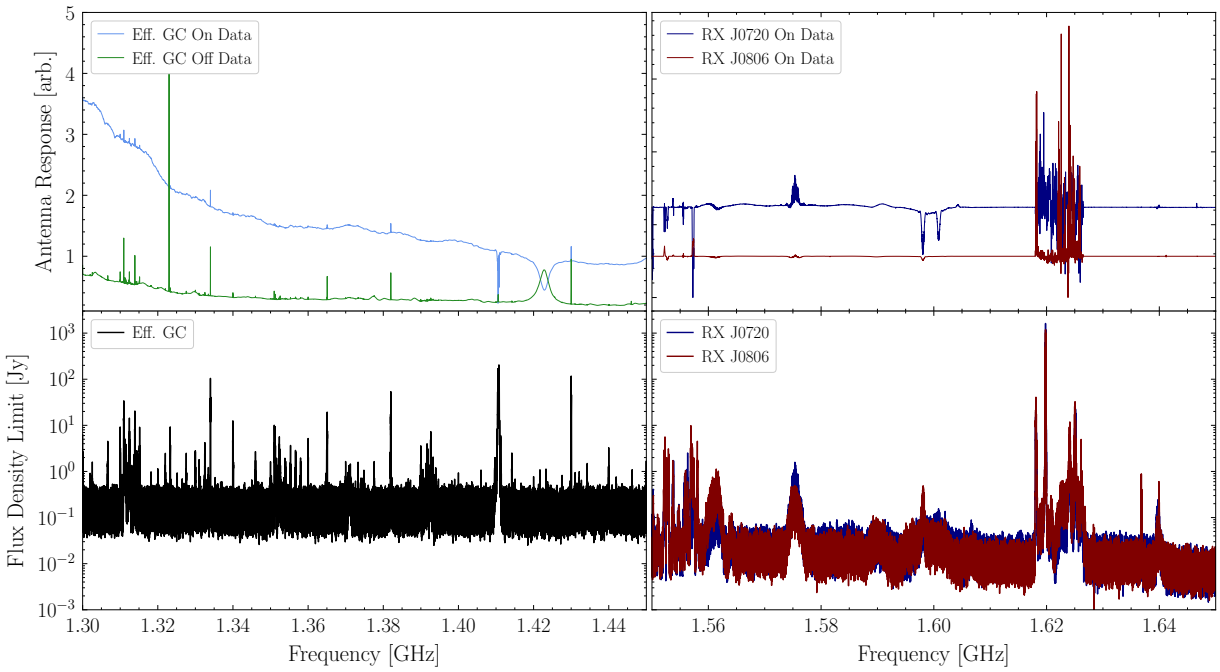



Figure F.2: (*Upper Left*) The raw, uncalibrated Effelsberg data collected in the L-Band at the GC and in the Off Position at frequencies between 1.3-1.45 GHz. Detector features much larger than the expected width of an axion signal and coincident RFI lines in On and Off data are clearly visible. (*Upper Right*) The raw, uncalibrated GBT data collected from the INSs RX J0720 and RX J0806. For visual clarity, we do not show the corresponding Off Position data. Thickets of RFI are especially visible around 1.575 GHz and 1.62 GHz. (*Lower Left*) The flux density limits as a function of frequency obtained from our analysis of the Effelsberg GC data. With the exception of locations of narrow RFI, the limits are flat and characterized by the expected statistical variations from channel to channel. (*Lower Right*) The flux density limits as a function of frequency obtained from our analysis of the GBT INS data.

The interpretations of the additional GBT observations in terms of the axion model are given later in this SM.

## F.2 GBT Data Processing

In this section, we describe the procedure by which we filter the time-series antenna data collected at the GBT and the modified implementation of `GBTIDL` used to process the time-series data

leading to a frequency-dependent stacked antenna temperature [294]. This antenna temperature is then translated to a flux density measurement through comparison to a reference calibration source.

## F.2.1 Determining the Antenna Temperature with Data Filtering

### F.2.1.1 System temperature calibration

For each observation (the procedure that we describe in this section applies to all GBT observations, including those only discussed in the SM), data was collected simultaneously for XX and YY polarizations across a range of frequency channels. For the INSs, data was recorded for $0.2097\,\mathrm{s}$ exposures, while for all other targets, data was recorded for $0.5\,\mathrm{s}$ exposures. In alternating exposures, a calibration noise diode with known effective temperature in each polarization is turned on and off, so that the first antenna measurements in each polarization do not include contributions from the noise diode, the second measurements do, the third measurements do not, and so on. Periodically, the observing position was alternated between ON and OFF positions. We will denote the $i^{\mathrm{th}}$ antenna measurement at the $j^{\mathrm{th}}$ frequency channel in the $ZZ$ polarization (where $ZZ$ stands for either $XX$ or $YY$) for the ON data at by $d_{i,j}^{ZZ}$ so that the noise diode is on and off for even and odd $i$, respectively. Likewise, we will denote the $i^{\mathrm{th}}$ antenna measurement at the $j^{\mathrm{th}}$ frequency in the $ZZ$ polarization for the OFF data by $\tilde{d}_{i,j}^{ZZ}$. Note that except for a handful of locations due to switching the observation position, the data $d_{i+1,j}^{ZZ}$ ($\tilde{d}_{i+1,j}^{ZZ}$) is collected immediately after $d_{i,j}^{ZZ}$ ($\tilde{d}_{i,j}^{ZZ}$). However, for a given $i, j$, $d_{i,j}^{ZZ}$ and $\tilde{d}_{i,j}^{ZZ}$ are *not* collected simultaneously.

The *system temperature* can be measured at each frequency channel in each polarization using the OFF-position data. The noise diode is alternated on at even $i$ and off at odd $i$, enabling us to obtain the system temperature by

$$T_{sys,j}^{ZZ} = T_{cal}^{ZZ} \left[ \frac{\sum_{\mathrm{odd}\,i} \tilde{d}_{i,j}^{ZZ}}{\left(\sum_{\mathrm{even}\,i} \tilde{d}_{i,j}^{ZZ}\right) - \left(\sum_{\mathrm{odd}\,i} \tilde{d}_{i,j}^{ZZ}\right)} + \frac{1}{2} \right] . \tag{F.1}$$

This frequency-dependent treatment is recommended [294] but not implemented by the standard `GBTIDL` procedure, in which the sums are extended to be additionally performed over the inner 80% of all frequency channels so that the observation is characterized by a single system temperature. We found the frequency-dependent procedure to result in a cleaner calibration. Next, using the system temperature as determined by the OFF-position data, we can determine the antenna temperature associated with the ON-position data by

$$T_{a,j}^{ZZ} = T_{sys,j}^{ZZ} \times \frac{\sum_i d_{i,j}^{ZZ} - \tilde{d}_{i,j}^{ZZ}}{\sum_i \tilde{d}_{i,j}^{ZZ}}. \tag{F.2}$$

#### F.2.1.2 Data cleaning

Because the antenna temperature and therefore the flux density spectrum are directly determined by time averages over the ON-position and OFF-position data, improved calibration precision and a cleaner flux density spectrum, less affected by spurious features, can be achieved by filtering out transient noise features that appear in the data. Here, we accomplish this by filtering the data independently at each frequency channel and in each polarization, in both ON and OFF positions, using the Kolmogorov-Smirnov (KS) test. This is a novel procedure we have devised and implemented in order to obtain clean flux density measurements for the high-precision axion search. Alternate non-parametric $k$-sample tests, such as the Anderson-Darling test, may also be appropriate but are more difficult to implement efficiently. We note also that while these cleaning procedures do help reduce the noise in the data, they are not strictly necessary and we find qualitatively similar results without data cleaning.

The filtering is performed in the following way. For a particular observation position, frequency channel, and polarization, we have time-series data which we divide into equally-sized intervals of the shortest length possible satisfying the following conditions:

- the interval represents at least 15 seconds of continuous exposure;
- the interval contains an even number of antenna measurements;
- the interval contains at least 60 independent antenna measurements.

Denoting the $j^{\text{th}}$ interval $y_j$, we determine the $y_j$ with the smallest mean and designate this as our reference interval, which we denote by $\hat{y}$. The reference interval is accepted by default. We then compare each $y_j$ with $\hat{y}$ under the KS test when each dataset is unaltered and when each dataset has its mean independently subtracted off. Datasets $y_j$ which are not found incompatible at the $5\sigma$ level with $\hat{y}$ under both tests are accepted, with the remaining excluded.

The requirement that the intervals represent at least 15 seconds of exposure is motivated by choosing a duration such that our filtering procedure will never select for intervals which do not contain flux density from an axion signal, which would lead to a spurious exclusion of a signal if it were present. The axion signal from a given NS will have a periodicity set by the NS period, and the time-averaged axion signal flux is expected to be constant over time intervals corresponding to many such periods. Therefore, since NS spin periods are expected to be less than 15 seconds, a potential signal will be present in all intervals. We additionally require at least 60 independent antenna measurements so that there is a sufficiently large sample for the KS test to have discriminating power. Choosing the shortest interval possible satisfying these criteria then maximizes the resolution of our filtering process. Moreover, performing the test with and without the channel mean removed is intended to improve the discriminating power of the KS test. Requiring that each

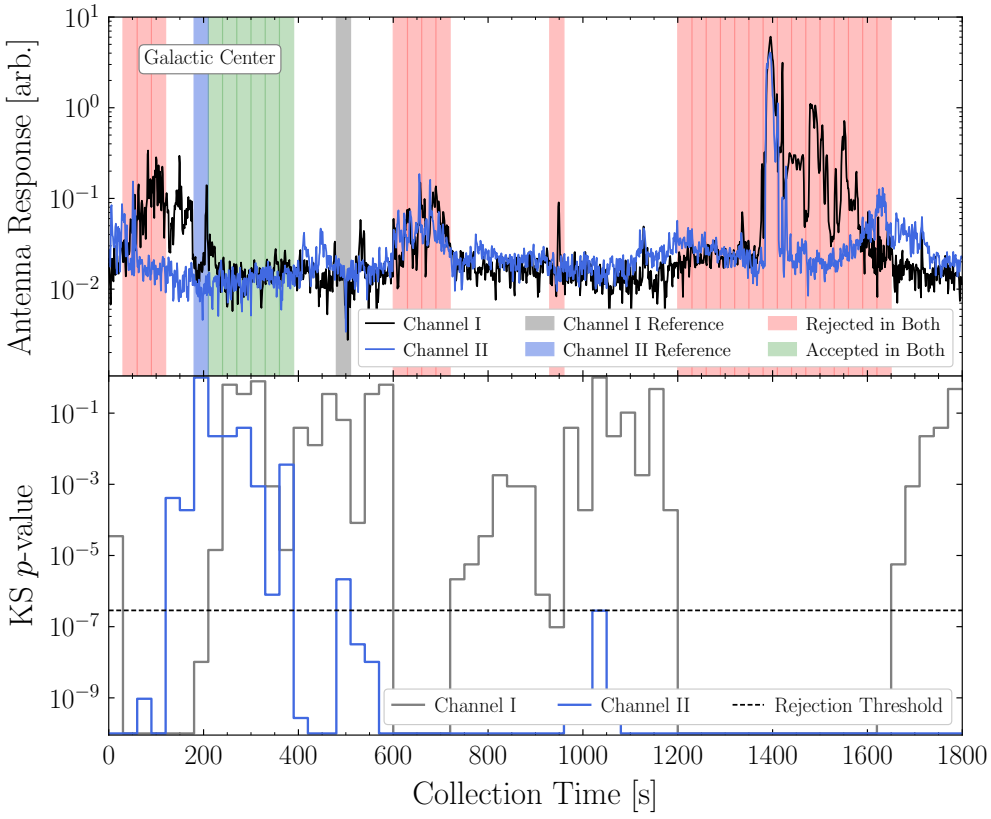



Figure F.3: The interval-by-interval acceptance for two adjacent frequency channels for data taken from the GC observation by the GBT. Channels I and II are located at 1.61908569 GHz and 1.61899414 GHz, respectively. Data for each channel are shown in black and blue, respectively, with correspondingly colored highlighted regions identifying the reference interval for each channel. The antenna response is shown in arbitrary units. Time intervals accepted in both channels are highlighted in green, with those rejected in both channels highlighted in red. Intervals which are accepted in only one channel are not highlighted.

interval is even-sized ensures that an equal amount of noise diode-on and noise diode-off data is contained in each interval and therefore in the accepted ensemble.

An example of the data filtering as applied to two adjacent frequencies in ON position data collected at the GC can be seen in Fig. F.3, where the advantage of independently filtering the data at each channel can be seen by the nontrivial differences in the data and the acceptance results in each channel. The polarization-averaged acceptances for the ON-position data for each observation, sorted by observing session, can be seen in Fig. F.4. For the INSs, the acceptance fraction is quite high, while the acceptance fraction tends to be lower from lower frequency resolution sources like M31 and the GC. This may be due to the reduction in statistical uncertainties when going to larger bandwidths and the increased importance of systematic uncertainties.

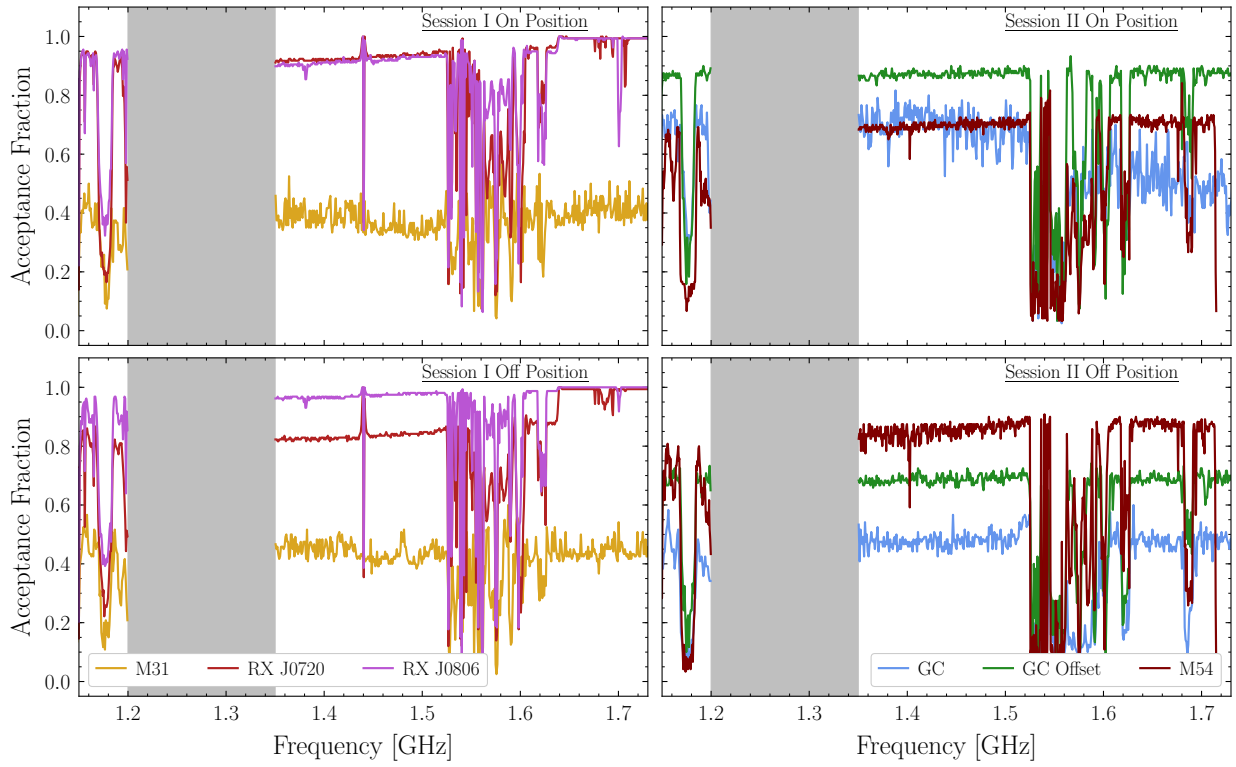



Figure F.4: The channel-by-channel acceptance fraction for each ON-position measurements of the observation target in each observing session. Acceptances are averaged over the two polarizations and downsampled by a factor of 50 for presentation purposes. In the top row, we display the acceptances in the ON observing position; in the bottom row, the acceptances in the OFF observing position.

With a successful filtering procedure applied to the data, the calculation of the system temperature is modified so that only data which is accepted in the filtering is included in the calculation of the system temperature, and subsequently the antenna temperature. Once we have determined antenna temperatures in each polarization, we determine a single antenna temperature by averaging over the polarizations, which typically differ at the percent level or below, as

$$T_{a,j} = \frac{T_{a,j}^{XX} + T_{a,j}^{YY}}{2} \, . \tag{F.3}$$

Since the antenna temperature $T_{a,j}$ is proportional to the flux density, we down-bin by averaging in antenna temperature.

### F.2.2 Calibrating to a Flux Density

We calibrate to a flux density through an observation of a flux calibrator. For Session I Observations (M31, RX J0720.4−3125, and RX J0806.4−4123), we observed the quasar 3C48. For Session II Observations (GC, GC Offset, M54), we observed the quasar 3C286. Several observations of each calibrator were made so that each target observation can be calibrated to a calibration observation with identical spectrometer configuration and frequency resolution.

Because the observations of the flux calibrators were relatively brief, the data are not amenable to a time-series filtering, so we directly determine an antenna temperature using (F.1), (F.2). For a particular calibrator, we have a calibrated flux density scale from [438, 439]. We then determine a frequency-dependent calibration scale for our data, $c_j$, by comparing the observed antenna temperature $T_{a,j}$ from the calibrator to the flux density expected from the source:

$$c_j = \frac{e_j}{\mathrm{med}(T_{a,j}, 31)} \tag{F.4}$$

Here $e_j$ is the expected flux density at the $j^{\mathrm{th}}$ frequency channel computed from [438, 439]. We smooth the antenna temperature at the analysis frequency resolution from the calibrator by a median filter with kernel size of 31 channels, which is wider than our analysis window, in order avoid calibrating against small-scale features that appear in the data, since the calibration data was taken at a lower time resolution. This frequency-dependent calibration factor then allows us to calibrate our antenna temperature measurement from our target observations.

## F.3 Effelsberg Data Processing

We note that the Effelsberg P217mm receiver provides circular polarizations with a receiver system equivalent flux density of $\sim 17$ Jy prior to averaging the two polarizations. The S110mm receiver similarly provides circular polarizations with a receiver system equivalent flux density of $\sim 11$ Jy. Effelsberg data at L- and S-Band were calibrated to a flux density scale via observations of the reference source $\mathrm{NGC}\,7027$ and using the publicly available software package `PSRCHIVE`[2] [440]. Unfortunately, an insufficient signal-to-noise ratio for the noise diode signal, in the fine frequency channels used in this work, resulted in strong frequency-dependent artifacts in the resulting calibrated spectra. Because these effects impinged on efforts to identify lines of interest, we assumed an ideal flat frequency response for the telescope receiver (except at the band edges) and defined calibration factors from the integrated spectrum for both ON and OFF target positions; in this case the GC.

However, we find that the maximum flux density in our data under this calibration procedure is in agreement with the expected flux density of the Galactic plane synchotron radiation flux density reported in [441]. We therefore calibrate both our L-Band and S-Band data by instead the calibration factor

$$c_j = \frac{e_j}{\mathrm{med}(d_j, 201)}\,, \tag{F.5}$$

where $e_j$ is the expected flux density at the $j^{\mathrm{th}}$ frequency channel from [441] and $\mathrm{med}(d_j, 201)$ is a median smoothing with kernel size 201 of the antenna data. This calibration assumes there to be no instrumental backgrounds on the antenna, which may not be true, and it also assumes that the expected flux density spectrum from [441], which was constructed by fitting a functional form to data from different telescopes across a range of frequencies, may be applied to the Effelsberg angular beam size, which is also likely not completely true. A comparison of the calibrated Effelsberg data to the calibrated GBT data for the GC is shown in Fig. F.5. The absolute magnitude of the Effelsberg flux density is typically $20\%$ larger in the L-Band than in the corresponding GBT data. The discrepancy between the GBT and the Effelsberg spectra is likely due to the Effelsberg calibration procedure. By comparison, in [442], a 21 cm observation of the GC using the Effelsberg telescope determined a flux density of 450 Jy, which is more compatible with the GBT GC spectra found in this work. Because the axion constraints scale like the square-root of the flux density, however, these differences have only a roughly 10% effect on the constraint strength and so we do not pursue them further. Moreover, we note that systematic errors in radio data are typically on the order of $\mathcal{O}(10\%)$, independent of calibration scheme.

[2] http://psrchive.sourceforge.net/

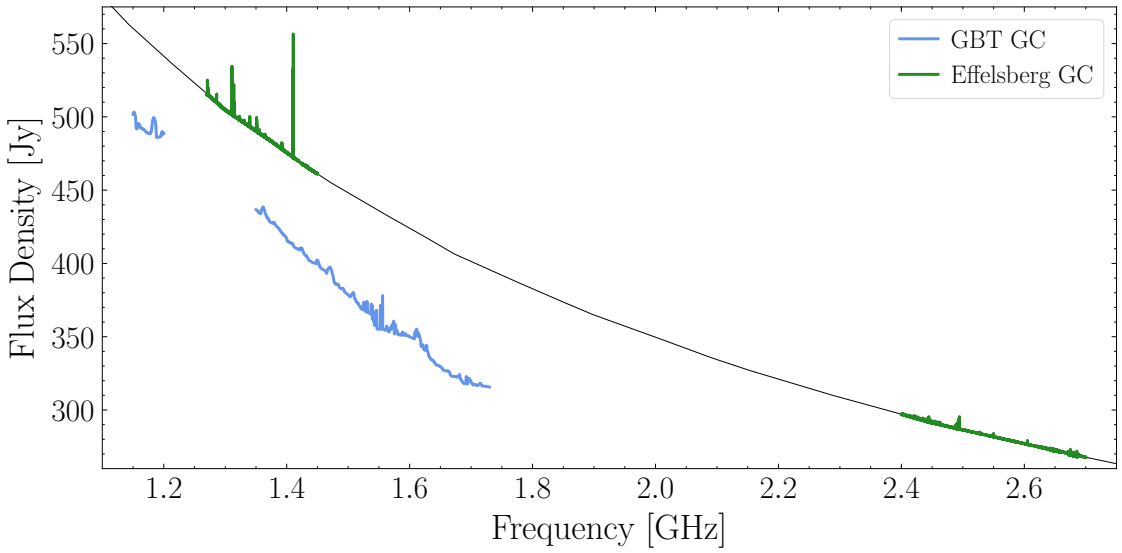



Figure F.5: A comparison of the calibrated flux density for the GBT observation of the GC (*blue*) to the Effelsberg observations of the GC in the L-Band and S-Band (*green*). Note that the Effelsberg data is calibrated to follow the black curve, averaged over large frequency scales. The calibrated L-band Effelsberg data is around 20% different than the calibrated GBT data, suggesting that errors from the calibration procedure impacting sensitivity to $g_{a\gamma\gamma}$ are only on the order of 10% and subdominant compared to other sources of uncertainty.

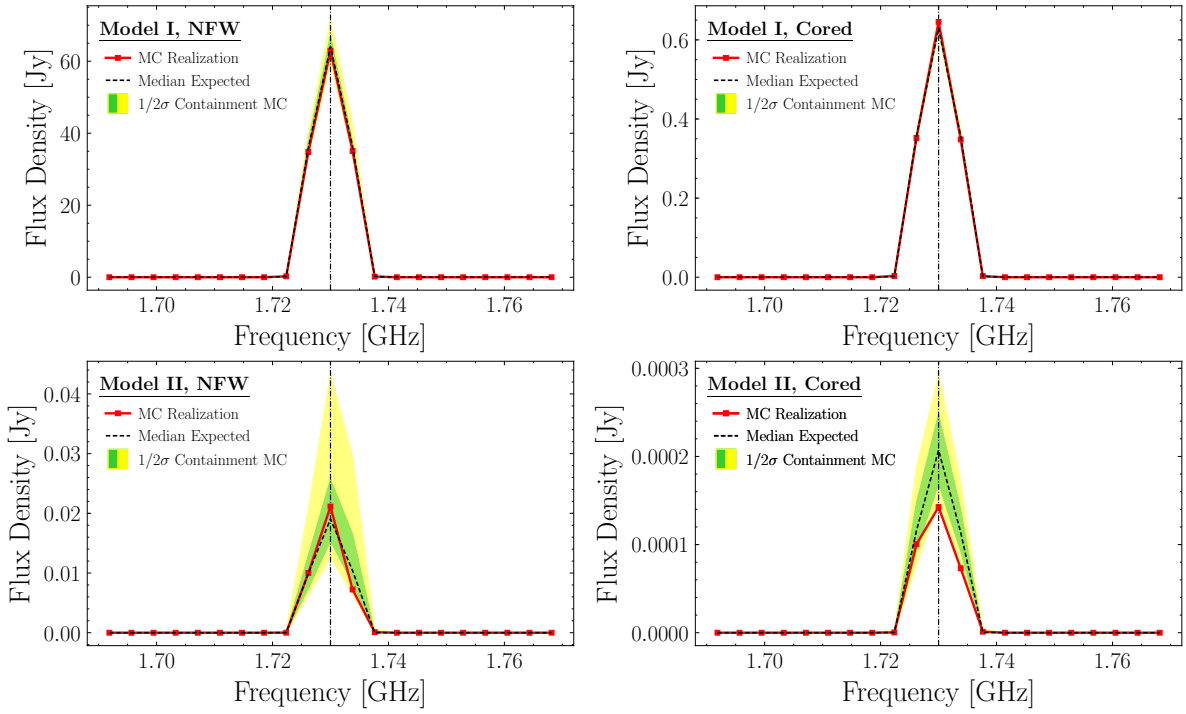



Figure F.6: (*Top Left*) A noise-free example flux density spectra for an axion of mass $m_a = 3.46\pi$ GHz with a coupling strength of $g_{a\gamma\gamma} = 10^{-11}\,\mathrm{GeV}^{-1}$ generated for the GBT observation of the GC at $\delta f_{\mathrm{fid}} = 1.831$ MHz. We assume Model I for the NS population (the model with more NSs participating in the conversion process) and take the DM to follow an NFW density profile. (*Top Right*) As in the top left, but using a cored DM density profile with a core radius of 600 pc. (*Bottom Left*) As in the top left, but assuming the conservative Model II for the NS population. (*Bottom Right*) As in the top right, but assuming the conservative Model II for the NS population.

## F.4 Axion Signal Modeling

Accurately modeling the axion signal is an important step in our limit-setting procedure. In order to model the radio signal from axion-photon conversion from a population of NSs, we choose a DM density profile and use the NS population models from [104] to determine the distributions of the NS locations, magnetic fields, and spin periods. We then draw from these distributions to construct a MC sample. NSs within the beam width will contribute to a potentially observable signal. At each axion mass, we compute a flux density associated with conversion at each NS within the beam that appears in the data, at a frequency corresponding to the axion mass shifted by a randomly-drawn Doppler factor corresponding to the peculiar velocity of that NS along the line of sight.

Examples of axion flux density spectra generated for analysis of the GBT GC observations (for population searches) can be seen in Fig. F.6. Note that these observations are down-binned such that the signals from all NSs appear mostly in the central bin, though there is some leakage to the

neighboring two frequency bins. Fig. F.6 shows the expected flux over multiple MC realizations for an axion with mass $m_a = 3.46 \times \pi$ GHz and $g_{a\gamma\gamma} = 10^{-11}$ GeV$^{-1}$. The dashed black curves show the median expected power, while the green and yellow bands show the 68% and 95% containment regions, respectively, for the power. The red data points illustrate one representative MC realization. Note that in Fig. F.6 we assume NFW DM or cored profile for the Milky Way, as indicated. We also alternate between NS Model I, which has more relevant, converting NSs, and our fiducial NS Model II, again as indicated. Note that the spread is greater for Model II since in this case there are less NSs participating in the conversion process.

By contrast, the Effelsberg data is of sufficiently high frequency resolution to enable the search for the brightest converting NS. Example flux density spectra are shown in Fig. F.7. The different panels indicate which NS population model is used in the analysis along with the assumed form of the DM density profile (NFW or cored NFW). Note that in these figures we have shifted the MC in frequency so that the brightest converting NS always shows up at the same frequency (here taken to be 2.5 GHz). Importantly, we see that the brightest converting NS typically makes the central frequency bin much brighter than the neighboring frequency bins, which are shown at the frequency resolution used in our analysis. This is especially true when simulating with NS model II, since in this case there are less converting NSs. The difference between the central and neighboring frequency bins also becomes more apparent with the NFW DM profile, relative to the cored DM profile, since the NFW profile causes those NSs close to the GC to have a bright axion-induced radio flux.

The situation is slightly more complicated in the NS Model I, cored DM profile case, since in this case the difference between the signal flux in the central frequency bin and the neighboring bins is not as pronounced, as seen in the MC example in top right panel of Fig. F.7. In practice this means that using the narrow-frequency bin approach leads to some difficulty in detecting an axion signal in the case of NS Model I and a cored DM profile with this analysis approach. For example, we may perform multiple MC simulations of the signal only (as shown in red in the top right panel of Fig. F.7) and then perform our analysis framework on this signal-only data. This is an idealized analysis, since we are neglecting other sources of noise, but it allows us to quantify the degradation to our detection capability because of signal leakage to the sidebands. We find, for example, that in this case (for Model I and cored DM profile) that at 95% confidence the discovery TS should be bigger than $\sim$30. On the other hand, for our fiducial scenario (Model II and an NFW DM profile) at 95% confidence the TS is greater than $\sim 10^3$, which indicates that signal leakage into the sidebands is not a concern in this case. For NS Model I and an NFW DM profile the TS is greater than $\sim$130 at 95% confidence, which is also greater than our pre-determined discovery threshold. Thus, the only concern is that in principle an axion signal produced in the S-band for NS Model I and a cored DM profile would not produce a discovery TS above our threshold of 100,

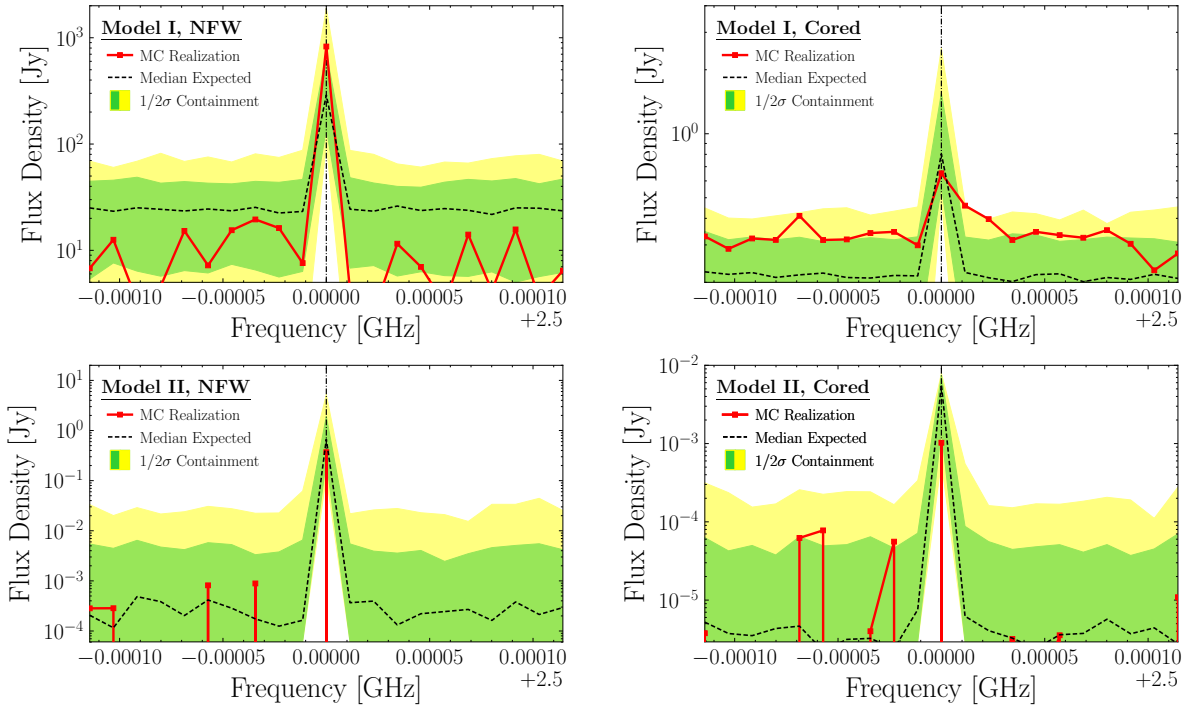



Figure F.7: As in Fig. F.6 but for the Effelsberg observations of the GC in the S-band, for an axion with $m_a = 2\pi \times 2.5$ GHz and $g_{a\gamma\gamma} = 10^{-11}\,\mathrm{GeV}^{-1}$. The panels indicate the assumed DM density profile for the Milky Way (NFW or cored NFW with a core radius of 0.6 kpc) and also the NS population model (Model I or Model II, as described in the text). Note that in this case we search for the brightest converting NS. We have shifted each of the MC realizations around in frequency space such that the brightest converting NS appears at $f = 2.5$ GHz. Note that in the scenario with Model 1 and a cored DM profile, the brightest converting NS is not always that much brighter than the signal flux in the sidebands, from other NSs within the Effelsberg beam, which makes it harder to discover an axion signal in this case.

at 95% confidence. Still, we note that in the S-band there are only two frequency channels with un-vetoed excesses with TS $> 30$ (one with TS $= 37$ and the other TS $= 41$), and in both of these cases the OFF data shows spectral features in the central frequency bin as well, though not at a significant enough level to be vetoed. Thus, we find it unlikely that these signals arise from axions. These excesses, along with others are depicted in Fig. F.12. Note that we account for the signal variance in the population models when setting 95% upper limits on the axion-photon coupling from the radio data (see Sec. F.5.3).

## F.5 Data Analysis

Here we describe the procedures by which we seek to discover an axion signal and by which we impose upper limits on the axion-photon coupling $g_{a\gamma\gamma}$ in the absence of signal detection.

### F.5.1 Profile Likelihood for Antenna Temperature Excesses

After our data processing and calibration procedure, we have obtained an antenna temperature $T_{a,j}$ and calibration factor $c_j$ over a range of frequency channels around a central frequency $f_j$. Axion-sourced signals are expected to appear as excesses at a narrow frequency bandwidth, so we analyze the antenna temperature downsampled to a resolution at which most (if not all) of the axion signal is expected to appear in a single frequency channel. In principle, we could analyze the calibrated flux density $c_j \times T_{a,j}$ for excesses, but we chose to not to do this so as to avoid injecting spectral features from a calibration spectrum into the calibrated flux density prior to analysis, which could result in the detection of spurious excesses. Analyzing the antenna temperature is acceptable because our smoothed calibration scale $c_j$ should vary slowly as a function of frequency on the scale on which the signal is expected to appear. If the calibration scale were to vary significantly from frequency channel to frequency channel, it would indicate regions in which the data is untrustworthy due to uncontrolled instrumental response. Indeed, this is the case for the GBT population analyses, which indicates that these analyses are dominated by instrumental systematics, as discussed previously in this SM.

We search for an excess in a single channel $f_i$ using the Gaussian likelihood

$$\mathcal{L}_i(\vec{T_a}|A, \mathbf{a}) = \prod_{k=i-j}^{i+j} \frac{\exp\left[-\frac{(T_{a,k}-\mu(f_k|\mathbf{a})-A\delta_{ik})^2}{2\sigma_k^2}\right]}{2\pi\sigma_k^2}\,, \tag{F.6}$$

where $A$ is signal parameter describing the size of the excess at the $i^{\text{th}}$ channel and $\mu$ is a quadratic polynomial describing the background determined by the nuisance parameter vector $\mathbf{a}$. The likelihood is calculated including data from the central channel where we seek to identify a putative excess and the $j$ sideband channels to both the left and right. In our fiducial analysis, we take $j = 10$. In each analysis, we exclude the sideband channels immediately adjacent to the central frequency channel in order to not be biased by less than 100% signal containment. The variance of the data $\sigma_k^2$ is determined by $\sigma_k^2 = \sigma^2/\alpha_k$ where $\alpha_k$ is the acceptance fraction of the data at the $k^{\text{th}}$ frequency channel and $\sigma^2$ is a fitted parameter. We fit $\sigma^2$ by fitting the null model ($A = 0$) to the analysis window with the central frequency channel (and hence any putative excess) masked out. We then do not profile over the value of $\sigma^2$ in the analysis. This is done to avoid biasing the calculation of the variance, and therefore the likelihood, in the presence of a large central-channel excess.

Equipped with our full likelihood in terms of both a signal and background model, we determine a profile likelihood purely as a function of the central-channel excess strength $A$ by profiling over

the nuisance parameter vector $\mathbf{a}$:

$$\mathcal{L}_i(\vec{T}_a|A) = \max_{\mathbf{a}}[\mathcal{L}_i(\vec{T}_a|A, \mathbf{a})] \, . \tag{F.7}$$

We determine the significance of the evidence for a flux density excess described by the best-fit excess parameter $\hat{A}$ using the test statistic (TS)

$$\mathrm{TS}_i = 2 \times \left[\log \mathcal{L}_i(\vec{T}_a|\hat{A}) - \log \mathcal{L}_i(\vec{T}_a|0)\right] , \tag{F.8}$$

unless $\hat{A} < 0$, in which case $\mathrm{TS}_i = 0$. By Wilks' theorem, this test statistic will be asymptotically half-$\chi^2$ distributed [78]. In our analysis, our threshold for discovery is set at $\mathrm{TS} = 100$. For details regarding the discovery threshold and TS distribution as informed by Monte Carlo, see Sec. F.8.1.

### F.5.2 Excess Vetoing Procedure

We find multiple excesses in the Antenna temperature data, such as excesses corresponding to 21 cm line emission, that may be vetoed by finding similar excesses in the OFF-position data. We analyze the OFF-position data corresponding to each observation, *i.e.*, the $\tilde{d}_{i,j}^{ZZ}$ that were used in (F.2) to subtract off backgrounds and instrumental baselines. This is important because narrow spectral features present in the OFF-position data indicate locations where backgrounds or antenna responses are systematically mismodeled and narrow spectral features in the calibrated spectrum do not require an axion interpretation. Note that by mismodeled we do not mean that a real astrophysical line is not necessarily present (such as the 21 cm line), but we mean that our simple quadratic background model is not sufficient to describe all of the astrophysical emission or instrumental backgrounds.

We analyze the OFF dataset $\vec{\tilde{d}}$, which is constructed from the polarization-summed OFF data that has been accepted by our time-series filtering. We analyze this data with a modified version of the likelihood in (F.6), now allowing a signal to appear as a flat excess across across one, three, or five frequency channels centered on the central frequency channel. This allows us to identify spurious features slightly wider than the signal we search for in the antenna temperature data that would nonetheless indicate locations of uncontrolled backgrounds and instrumental response. The

likelihood is given by

$$\mathcal{L}_i(\vec{d}|A,\mathbf{a},w) = \prod_{k=i-j}^{i-w-1} \frac{\exp\left[-\frac{(\tilde{d}_k-\mu(f_k|\mathbf{a}))^2}{2\sigma_k^2}\right]}{2\pi\sigma_k^2} \prod_{k=i+w+1}^{i+j} \frac{\exp\left[-\frac{(\tilde{d}_k-\mu(f_k|\mathbf{a}))^2}{2\sigma_k^2}\right]}{2\pi\sigma_k^2} \times \frac{\exp\left[-\frac{(\sum_{k=i-w}^{i+w}\tilde{d}_k-\mu(f_k|\mathbf{a})-A)^2}{2\sum_{k=i-w}^{i+w}\sigma_k^2}\right]}{2\pi\sum_{k=i-w}^{i+w}\sigma_k^2}. \tag{F.9}$$

This likelihood is at full spectral resolution in the sidebands (the top line in Fig. (F.9) accounts for the left and right sidebands) but at downbinned resolution corresponding to the window width in the signal region (the bottom line in Fig. (F.9)). Note that we consider three values for $w$ in (F.9): $w = 0, 1, 2$, corresponding to signal-region widths of 1, 3, and 5 bins. Our variance parameter $\sigma_k^2$ is treated analogously to before, and the TS is taken to be

$$\tilde{\mathrm{TS}}_i(w) = 2 \times [\max_{A,\mathbf{a}}\mathcal{L}_i(\vec{T}_a|A,\mathbf{a},w) - \max_{\mathbf{a}}\mathcal{L}_i(\vec{T}_a|0,\mathbf{a},w)]. \tag{F.10}$$

Note that we calculate the TS for each choice of $w$. Here we do not zero out TSs associated with negative best-fit signal parameters, as our interest is only in identifying locations of spurious features of any sign. We also do not mask out the frequency channels adjacent to the signal region in this analysis as we are not seeking to identify an axion signal in this data. Large TSs in the antenna temperature data are then vetoed if they are at identical or directly adjacent frequency channels to a channel in the OFF-position data with a TS above the $97.5^{\mathrm{th}}$ percentile value of the OFF-position TS ensemble. We perform this test for each of the three choices for $w$. In practice, we find that vetoing on the five-channel wide analysis has negligible effect. An example of vetoed excesses can be see in Fig. F.8.

### F.5.3 Axion Constraints

Setting constraints on the axion-photon coupling $g_{a\gamma\gamma}$ requires a more involved data analysis procedure than our discovery search. This is because variance in the expected axion signal due to scatter in the NS population prevents a direct conversion of a flux density limit, obtained through Wilks' theorem using our profile likelihood, to a constraint on $g_{a\gamma\gamma}$.

Instead, we take a profile-likelihood MC approach that allows us to place limits that account for signal variance. To do this, we analyze the real data under the likelihood of (F.6) to determine the best-fit signal strength parameter $\hat{A}$ and the best-fit null-model background specified by the best-fit model parameters $\hat{\mathbf{a}}$ and $\sigma^2$. We generate 200 MC realizations of the data under the best-fit null model. For each null model spectra, we construct a unique realization of an axion signal generated

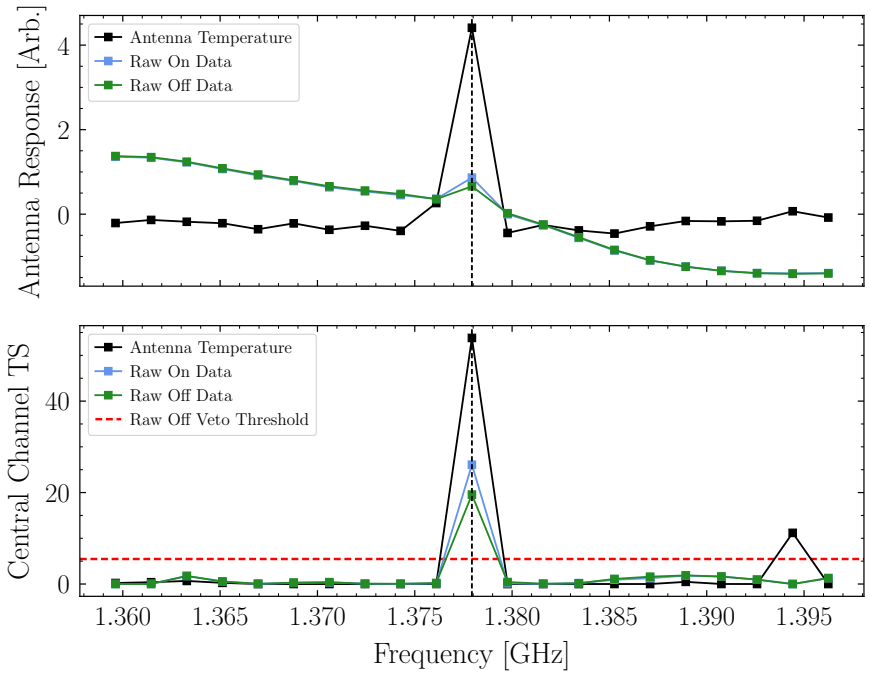



Figure F.8: (*Above*) The ON-position antenna temperature and raw antenna data for ON- and OFF-position measurements for the M31 observation. A narrow feature appears at the indicated central frequency channel in each of the datasets, although with larger relative magnitude in the antenna temperature. The fact that the feature appears in all datasets suggests it is not an axion signal. (*Below*) The test statistics for the central channel excess as a function of the central channel for the analysis of the ON-position antenna temperature and the raw antenna data for ON- and OFF-position measurements for the M31 observation. At the location of the narrow feature, the test statistic is quite large for all analyzed datasets, and the excess in the antenna temperature is vetoed as the test statistic in the OFF-position data exceeds the veto threshold.

under the assumed NS population and DM density profile. The axion signal is injected in the null model spectra with varying strength corresponding to adjusting the value of $g_{a\gamma\gamma}$. We analyze each MC data realization in order to determine the fraction which result in a best-fit signal strength parameter $\hat{A}_{MC}$ that is less than the best-fit signal strength parameter $\hat{A}$ in the real data. We can then obtain a frequentist 95% upper limit on the $g_{a\gamma\gamma}$ by determining the value of $g_{a\gamma\gamma}$ at which only 5% of the $\hat{A}_{MC}$ are less than $\hat{A}$. The advantage of this procedure is a safe accounting for variance in the signal strength and Doppler shifting that lead to less than perfect signal containment in the central frequency channel.

## F.6 Signal Injection Tests

To demonstrate that our framework is capable of setting appropriate upper limits and discovering an axion were such a signal present in the data, we perform signal injection tests where we inject synthetic axion signals into the actual data. Here we demonstrate these tests for signal flux density spectra injected at several frequencies in Effelsberg S-Band data. We inject the signal flux density spectra with varying amplitudes corresponding to varying the strength of $g_{a\gamma\gamma}$. The results of the signal injection tests can be seen in Fig. F.9. In the left panel we show the 95% one-sided upper limits $g_{a\gamma\gamma}^{95\%}$ as a function of the injected signal strength $g_{a\gamma\gamma}^{\rm inj}$ for three different and randomly chosen axion masses. Importantly, we never exclude the injected signal strength, which is indicated by the dashed black line. In the right panel we show the discovery TS in favor of the axion model as a function of the injected signal strength, with our $\mathrm{TS} = 100$ discovery threshold indicated in horizontal dashed black. As expected, the TS increases for increasing signal strengths.

## F.7 Maser Line Detection Test

The observation of known spectral line sources enables additional opportunities to test our analysis pipeline. During the March $10^{\rm th}$ observing session, 10 minutes of observing time was used to collect data in the L-Band at the compact HII region W3OH. Data was collected at 92 kHz frequency resolution for five minutes in the ON position, followed by five minutes of data collection at identical frequency resolution in the OFF position. Data was collected over the frequency range 1.15-1.73 GHz with a notch filter applied in the 1.2 to 1.34 GHz region as in all other GBT observations. Narrow maser emission at rest frequencies approximately 1.612 GHz, 1.665 GHz, 1.667 GHz, and 1.72 GHz associated with transitions in hydroxyl molecules have been observed in W3OH [443, 444], enabling a direct test of the ability of our analysis pipeline to identify bright spectral lines of a similar character to those that might appear due to axion-to-photon conversion in NS magnetospheres.

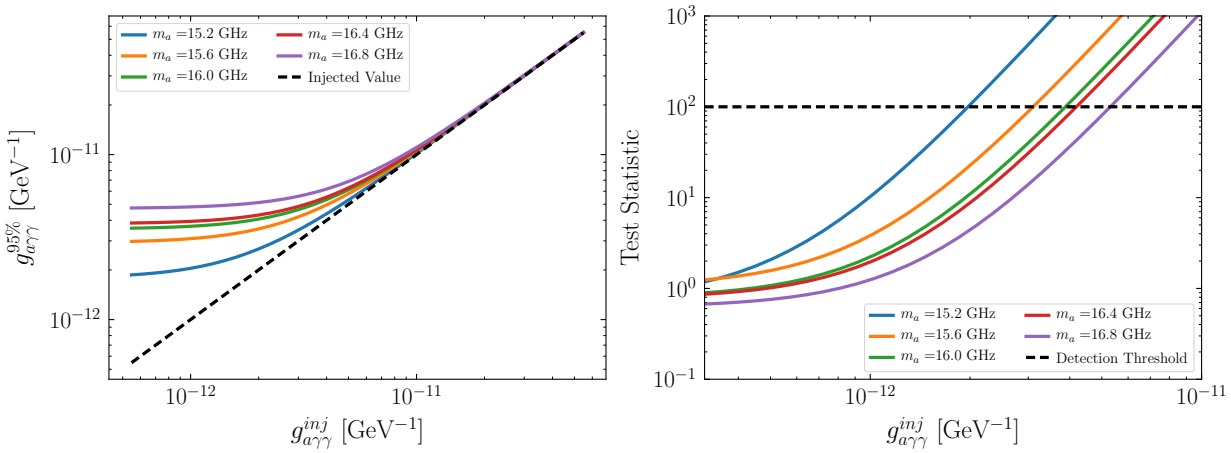



Figure F.9: (*Left*) The one-sided 95% upper limit on the axion-photon coupling as a function of the injected signal strength. The limit lies above the injected signal strength, indicating that we are not excluding an axion signal when present. (*Right*) The test statistic (TS) for discovery as a function of the injected signal strength. For sufficiently large signal strengths the TS exceeds our $\mathrm{TS} = 100$ threshold for an axion signal to be discovered.

We analyze the collected data over the full frequency range at the original 92 kHz frequency resolution with the procedure described in Sec. F.5 in order to search for any spectral lines which may appear. All OH maser lines are detected at their expected frequency locations consistent with the measured Doppler shifting and line-widths of 1665-MHz OH emission within our observing beam [445]. Moreover, none of the detected OH maser lines are vetoed by a coincident excess in the OFF data, validating our exclusion criteria. Results are shown in Fig. F.10. Additional high significance excesses $\mathrm{TS} \geq 100$ are observed at the frequencies 1.35065308, 1.374823, and 1.64664307 GHz. The excess at 1.64664307 GHz is coincident with an excess observed in the L-Band during GBT RFI scans,[3] while the origins and appropriate interpretations of the other two excesses are less certain. Despite the presence of two unidentified lines, these tests effectively demonstrate the ability of our framework to recover spectral lines present in the data. Analogous data were collected for 1665-MHz OH maser 351.7754−0.536 [446] during the March 29$^{\mathrm{th}}$ observing session, with similar results regarding maser line detection.

## F.8 Extended Results

Here we provide extended results for the analyses presented in the main Letter and for the additional GBT observations described at the beginning of this appendix.

[3]RFI Scan data available from the NRAO website

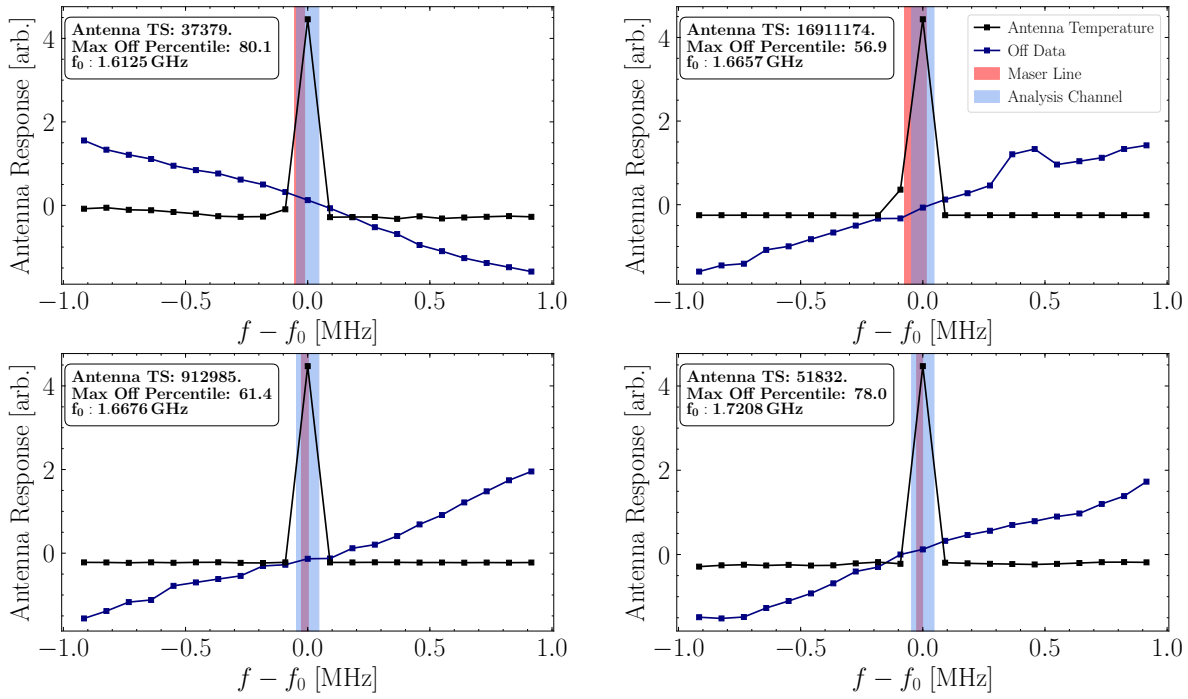



Figure F.10: Maser lines as detected in the GBT data. For each maser line, we show the antenna temperature (black) and the raw OFF data (blue), with each independently rescaled so as to fit within the same figure. The expected frequency location and width of the maser line, which is set by the line-of-sight velocity of W3OH, is indicated by the light red band. The width of the central frequency channel in which the maser line is detected is indicated by the light blue band. We additionally provide the TS associated with the maser line detection in the antenna temperature and the maximum percentile of the variable-width OFF position TS for each line. None of the detections are vetoed as none the maximum OFF position TS percentiles exceed the $97.5^{\text{th}}$ percentile value that triggers vetoing.

### F.8.1 Survival Functions and Excess Candidates

For each observation, we compute the survival function associated with the observed TS distribution. In all observations, no significant TSs exceeding our discovery threshold are observed, and the survival functions appear to match the expectation under the null hypothesis, as shown in Fig. F.11. Note that the survival function figure shows the faction of TSs at or above the indicated TS on the $x$-axis.

From Wilks' theorem we would expect the distribution of TSs to be asymptotically $\chi^2$-distributed under the null hypothesis. This asymptotic expectation is indicated in Fig. F.11. However, our expected distribution of TSs does not follow this asymptotic expectation, even under the null hypothesis, because our analysis window contains only a finite number of frequency bins in the sidebands. When we perform MC simulations under the null hypothesis (first, we fit the null model to the data to determine the model parameters, and then we simulate data from those best-fit parameters and re-analyze the simulated data) we find that passing that simulated data through our analysis pipeline results in a survival function, under the null hypothesis, that is not $\chi^2$-distributed. In Fig. F.11 we show the simulated survival function under the null hypothesis ("MC Expected"), which is constructed by averaging the simulated survival functions across all observations (though we find that all observations produce consistent expectations for the survival function). We note that if we modify our analysis framework to include more frequency channels in the sidebands, the survival function better approaches the $\chi^2$ distribution. On the other hand, the fact that under the null hypothesis our survival function is not precisely $\chi^2$-distributed does not mean that our analysis framework is in any way not valid. It simply means that when assigning TS values significances (*e.g.*, $p$-values), we should use the simulated survival function and not the asymptotic expectation from Wilks' theorem. For example, $\mathrm{TS} = 100$ corresponds to $10\sigma$ significance under Wilks' theorem. On the other hand, using the simulated TS distribution under the null hypothesis we find that in fact for our analysis framework $\mathrm{TS} = 100$ corresponds to approximately $5\sigma$ (local) significance ($p$-value of approximately $6 \times 10^{-7}$).

In the left panel of Fig. F.11 we have already applied the vetoes from analyses of the OFF data. That is, frequencies that are vetoed from the OFF data analyses are not shown in Fig. F.11. In this case, the observed TS distributions are found to closely match the MC expecations under the null hypothesis. In the right panel of Fig. F.11 we show what would happen if we did not apply the OFF data vetoes. In this case, there are many significant detections. Some of these detections correspond to real astrophysical lines, such as the 21 cm line, that also appear in the OFF data, while other lines may be due to RFI or instrumental backgrounds.

Note that we veto excesses at 27 of 56362 analyzed channels in the RX J0720.4−3125 data, 37 of 56362 analyzed channels in the RX J0806.4−4123 data, 24 of 26214 analyzed channels in the Effelsberg S-Band data, and 22 of 24576 analyzed channels in the Effelsberg L-Band data.

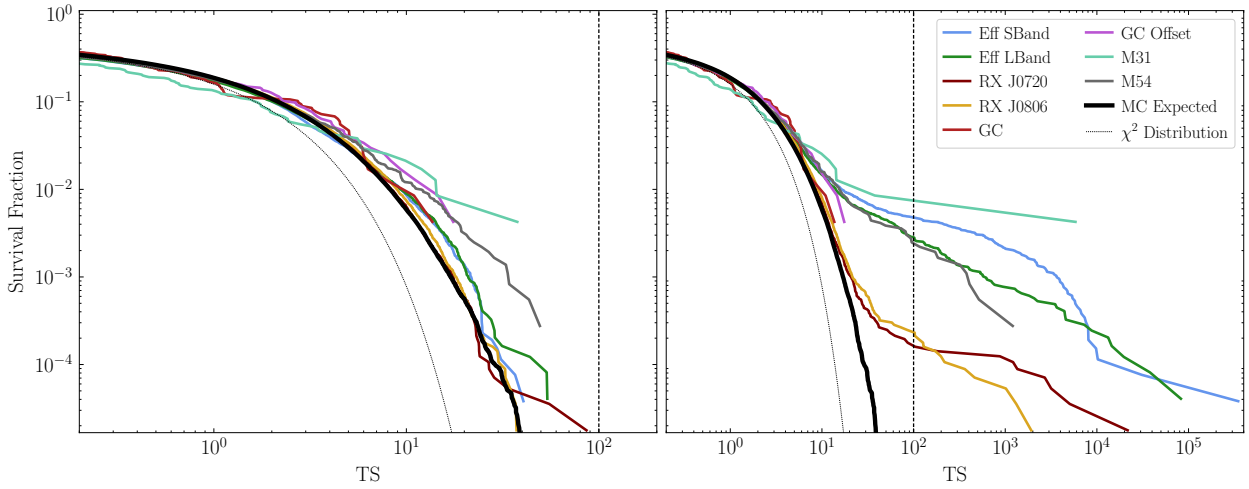



Figure F.11: (*Left*) The discovery TS survival function for all of the observations considered in this Letter. Note that the survival function is defined as the fraction of TSs observed at or above the indicated value. This figure excludes frequencies that are vetoed from the OFF position observation analyses. The "MC Expected" curve shows the expectation under the null hypothesis, as determined by MC simulations. We note that all observations are from GBT except those labeled "Eff", which are from the Effelsberg telescope. (*Right*) As in the left panel, but including frequencies that would be vetoed by the OFF data. Without the OFF vetoes there would be a significant number of frequencies with TSs exceeding the TS detection threshold, which emphasizes the importance of the OFF position vetoing procedure.

We do observe some moderate to high significance excess in our data in various observations which are depicted in F.12 with corresponding TS values and frequency locations. In the Effelsberg observation of the GC in the S-Band we observe two unvetoed excesses, both with $\mathrm{TS} \approx 40$. While these are not exceptionally large TS values, as studied in our MC signal construction, even strong axion conversion signals can result in TSs as small as 30 in the scenario with a cored DM density profile assuming the NS population is described by Model I. Similarly, we have a moderate significance excess ($\mathrm{TS} \approx 52$) in Effelsberg observation of the GC in the L-Band. A relatively higher significance excess ($\mathrm{TS} \approx 90$) is observed in the GBT observation of RX J0720.4−3125. We show the data corresponding to the excesses in Fig. F.12. We remain skeptical that these excesses requires an axion interpretation, as similar or unusual features appear in the OFF-data but at lower significances (below our veto threshold). Additionally, the location of the RX J0720.4−3125 excess at approximately 1.59 GHz is known to be subject to strong RFI. Follow-up observations with longer exposures and at complementary positions on the sky would be necessary to confirm or exclude the persistence of such excesses.

Note that we also do not see any significant excesses ($\mathrm{TS} > 100$) when shifting the frequency bins by half a bin size. As described in the main Letter, we analyze the data shifted by half a frequency bin to account for the possibility that the axion mass falls between our frequency bins and the flux is thus split between neighboring bins. In the Supplementary Data [303] we present the flux and $g_{a\gamma\gamma}$ limits at each frequency point, shifted and unshifted, along with the corresponding TSs.

Additionally, our choice of a quadratic background model is somewhat arbitrary, but has limited impact on our results. What is important is to have a background model with enough parameter freedom to describe the data under the null hypothesis but not so much freedom that the background model can be degenerate with the signal model. The strength of axion limits set under our likelihood procedure is not appreciably affected by the background model, for small variations to the model. In Fig. F.13, we show that in the Effelsberg S-Band data, the detection significances are not appreciably different across flat, linear, and quadratic background models. The flat background model appears to be too simplistic to accurately model the form of the data, but the quadratic and linear background models give quantitatively similar results.

In Fig. F.14, we show that in the INS GBT data the large scale features of the survival function are not appreciably changed by the application of our time-series data filtering procedure. However, the number of moderate-to-high significance excesses in RX J0720.4−3125 and RX J0806.4−4123 observations does decrease without the application of data filtering, which is consistent with the expectation that the excesses are largely sourced by transient RFI.

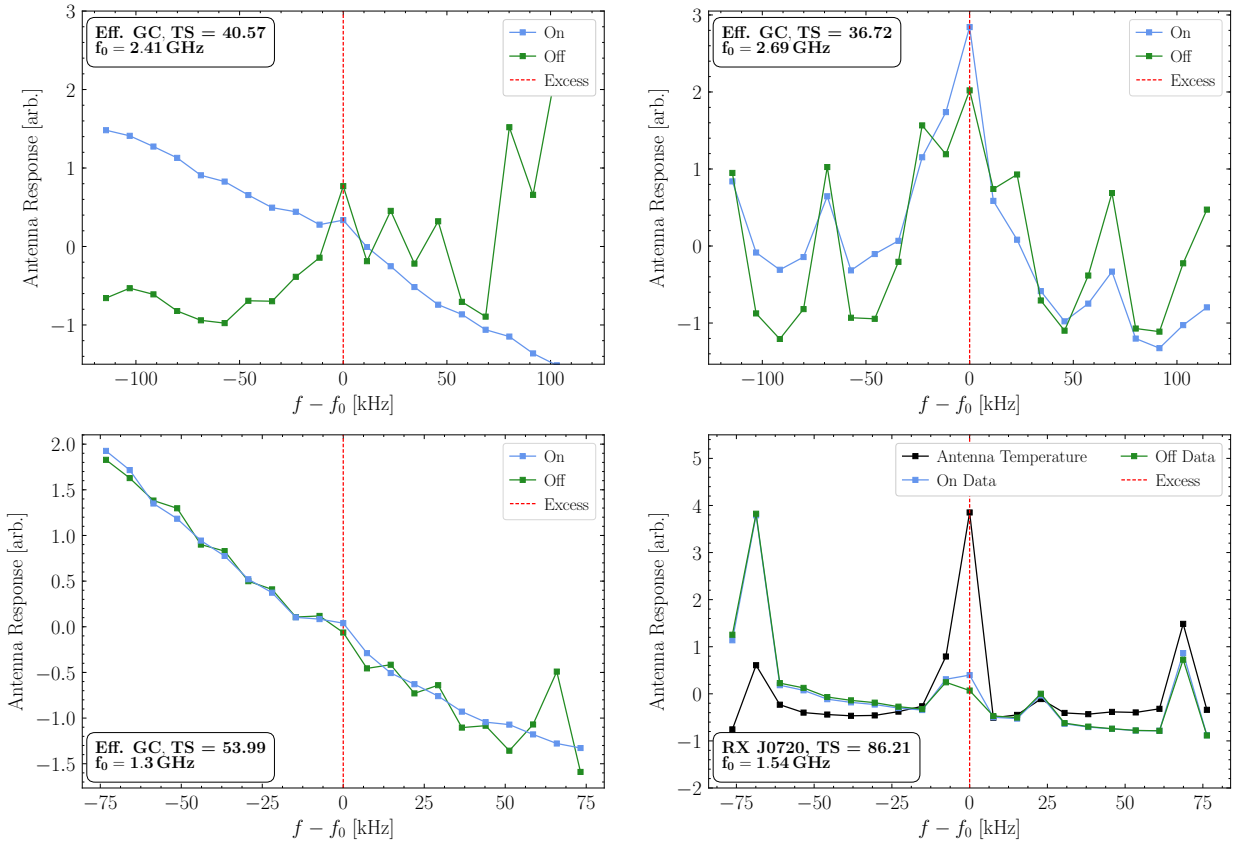



Figure F.12: (*Top Left*) The Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 2.51 GHz in the S-band observation of the GC. Frequencies are plotted relative to the frequency corresponding to the excess channel frequency. This excess has TS $\approx$ 41. While this excess is not vetoed by the OFF data analysis, the OFF data does should a feature at the central frequency. (*Top Right*) Similarly, the Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 2.69 GHz in the S-band observation of the GC. This excess is also not vetoed, but like the previous excess there does appear to be a corresponding feature in the OFF data. (*Bottom Left*) The Effelsberg data shown in the analysis window around the excess candidate located at a central frequency of approximately 1.34 GHz in the L-band observation of the GC. It also appears that there is a similar, though not so significant, feature in the OFF data. (*Bottom Right*) The GBT data shown in the analysis window around the excess candidate located at a central frequency of approximately 1.59 GHz in the observation of RX J0720.4−3125. As before, frequencies are plotted relative to the frequency corresponding to the excess channel frequency. The excess only appears at high significance in the antenna temperature; similar coincident features are observed in both ON and OFF data, coincident features appear in the raw ON and OFF data, although not at high enough significance in the OFF data to result in a veto of the excess. As before, this excess does not exceed our detection threshold, although it does come closer, with TS $\approx$ 90.

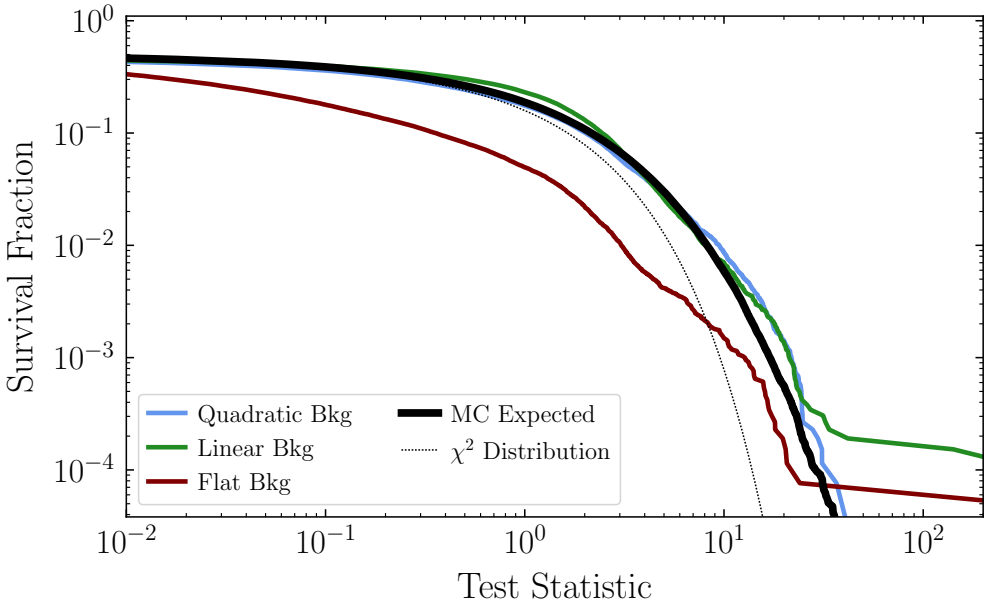



Figure F.13: A comparison of survival functions for various polynomial background models for the analysis of Effelsberg S-Band data. The flat background model is unable to accurately model the null hypothesis and a significant improvement in the quality of the fits is seen by going to the linear background model. On the other hand, there is little improvement to the quality of the fits when going from the linear to quadratic background models, except at very high TS values. Note that we use the quadratic background model in our fiducial analyses. Cubic background models produce similar results but are most computationally intensive to implement.

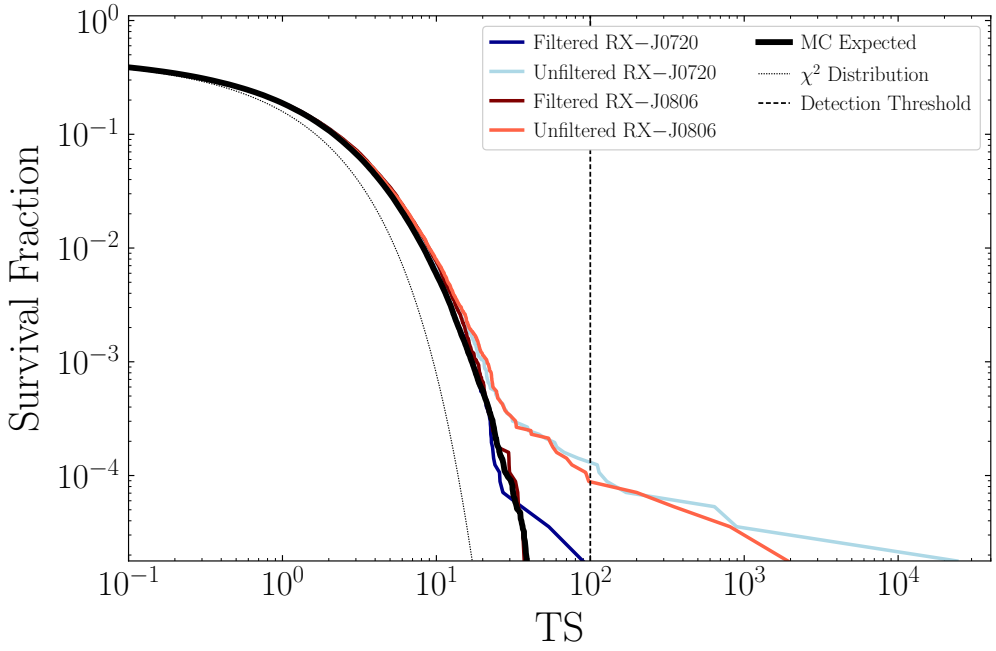



Figure F.14: The discovery TS survival function for the INSs analyzed with and without time-series data filtering applied. Applying the time-series filtering eliminates a number of high significance excesses that appear due to transient noise that appears in the data.

### F.8.2 Variations to the NS and DM Density Modeling

In this section, we show how variations to the NS population models and assumed DM density profiles affect the strength of the limits we are able to set. Under NS population Model I from [104], and assuming all DM density profiles are perfectly NFW, we obtain our strongest limits, as shown in the top left panel of Fig. F.15. Note that for M54 the NFW profile is for the host Sagittarius Dwarf Spheroidal Galaxy (see [104] for details). Limits are made successively weaker by the assumption of our fiducial NS population model (Model II from [104]), as shown in Fig. F.15. See that figure for all four combinations of NS models and DM density profile choices. Note that the Milky Way and M31 cores are taken to be 0.6 kpc, with the DM density profile following an NFW profile outside of this radius and flat within the core radius. For Sagittarius, relevant for M54, under the cored DM scenario we model the DM density distribution as an isothermal sphere with scale radius of 0.2 kpc (see [104]).

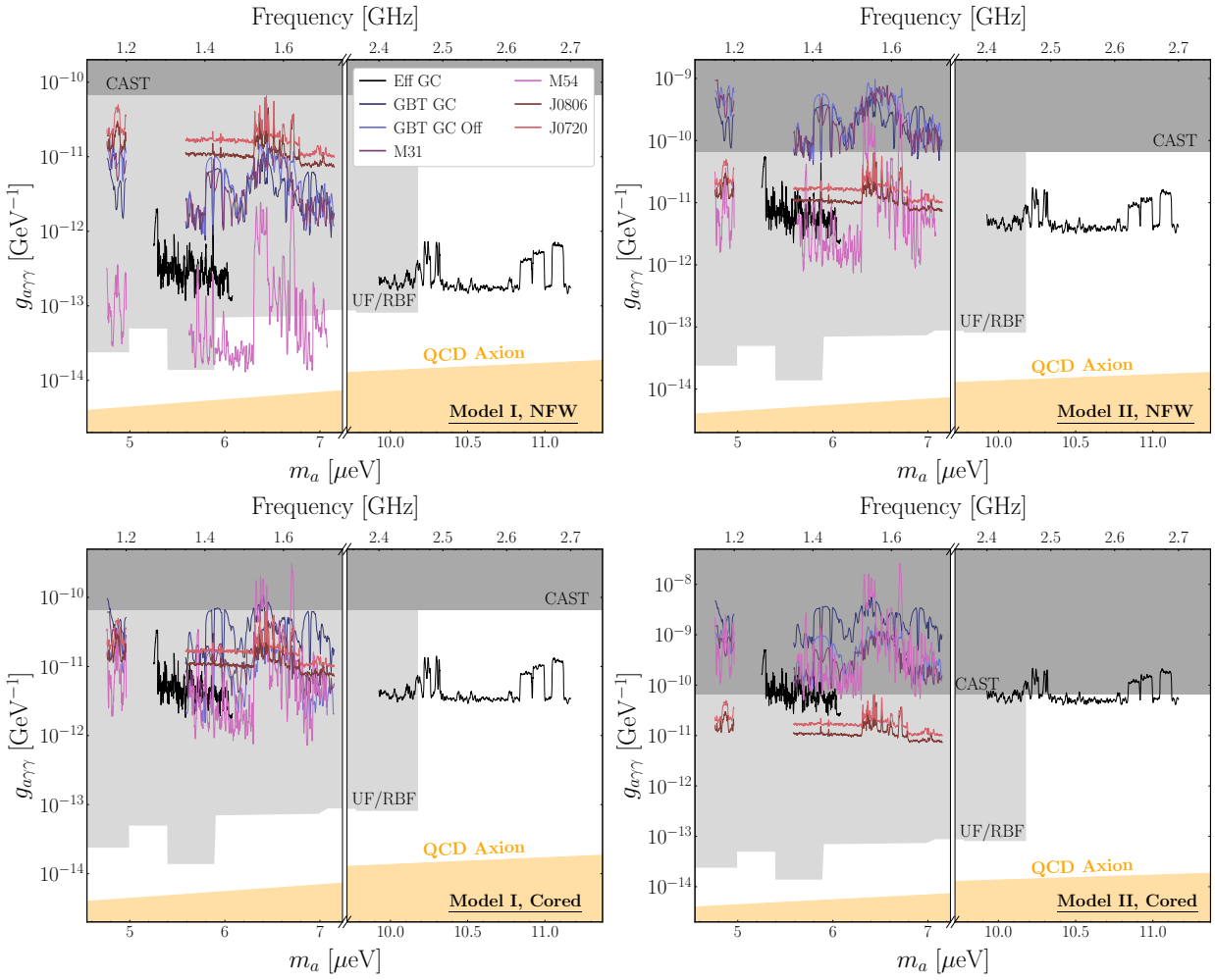



Figure F.15: Limits on the axion-photon coupling for different combinations of assumptions about the DM density profiles in the observed galaxies and the properties of the NSs within those galaxies (see [104] and text for details).

## F.9 Alternative Flux Density Limits

As a further test of the validity of our flux density limits, we recompute them using an alternative approach. The alternative 95% C.L. flux density upper limits ("percentile limits method") are constructed as follows. First, we smooth the calibrated spectrum with a median filter in order to remove any low-significance noise and to emphasize any potential spectral lines present in the data. The median filter assumes frequency windows containing 20 channels (for simplicity we used the same number of channels for both GBT and Effelsberg data). The outcome of this step is a data-driven estimate of the background in each respective window. Second, we single out spectral line signals $T(f_j)$ by subtracting the smoothed background from the raw data. Third, we sort the $T(f_j)$ lines in ascending order using an "order filter." Fourth, we compute the $68^{\text{th}}$ percentile band ($1\sigma$ error) at each frequency window. Finally, we obtain approximate 95% C.L. flux density upper limits by calculating the quantity $(T(f_j) + 2\sigma)$. Note that this method assumes that the uncertainties are well-behaved, *i.e.*, going from $1\sigma$ (68% C.L. ) to $2\sigma$ (95% C.L.) is fairly linear. Furthermore, in order to avoid excessively strong constraints, we floor any signals that were smaller than the negative $68^{\text{th}}$ percentile error band ("power limited constraints" [172]), just as we do for our profile-likelihood limit setting procedure.

The results of our comparison for GBT observations of M54 is shown in Fig. F.16 and for Effelsberg GC observations in Fig. F.17. As can be seen, the two different methods display very good agreement. This illustrates that the profile likelihood and the percentile upper limits methods are essentially equivalent and that the upper limits obtained in this work do not depend sensitively on the limit-setting procedure. We note that the advantage of the percentile method is its efficiency: it avoids computationally expensive log-likelihood maximization computations.

## F.10 Bandwidth estimate from refraction in moving medium

In Ref. [287] it was pointed out that the reflected photons may acquire Doppler boosts for a rotating and misaligned magnetosphere. In effect, one may think of the conversion surface as a mirror that reflects the incoming axions into outgoing photons. In order for the mirror to transfer momentum to the photons, the mirror must have a component of its velocity in the normal direction. For an aligned rotator, where the magnetic axis is aligned with the axis of rotation, the velocity vector at any point on the conversion surface is orthogonal to the normal vector from the surface and thus there is no induced Doppler broadening from reflection. However, as pointed out in [287], for misaligned rotators the conversion surface locally has a velocity that has a component in the direction of the normal vector, thus inducing a frequency shift.

Here, we consider a similar frequency shift induced by *refraction* for axions that are converted

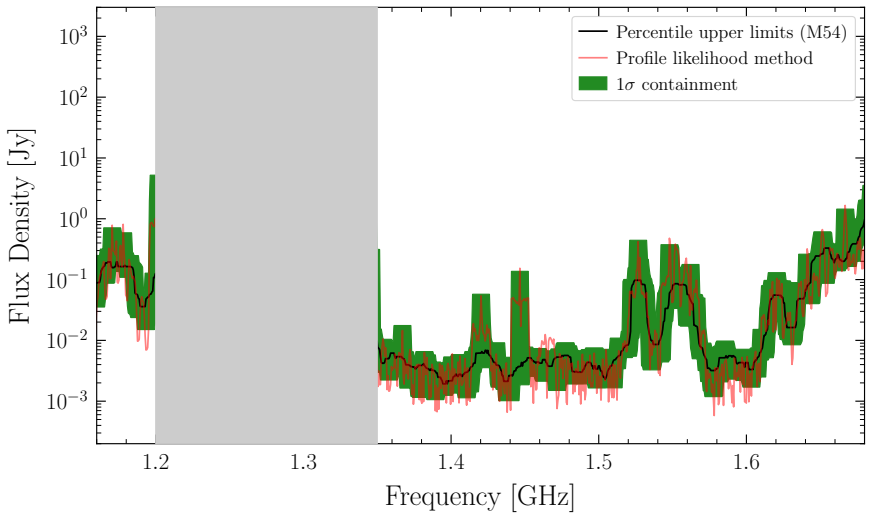



Figure F.16: Comparison between the profile likelihood and percentile upper limits methods for M54 observations with GBT. The black line (green area) shows the 95% C.L upper limits ($1\sigma$ containment band) obtained with the percentile method. The red line shows the upper limits obtained with the profile likelihood method and calibration used as default in our main pipeline; these flux limit curves are used in the main text.

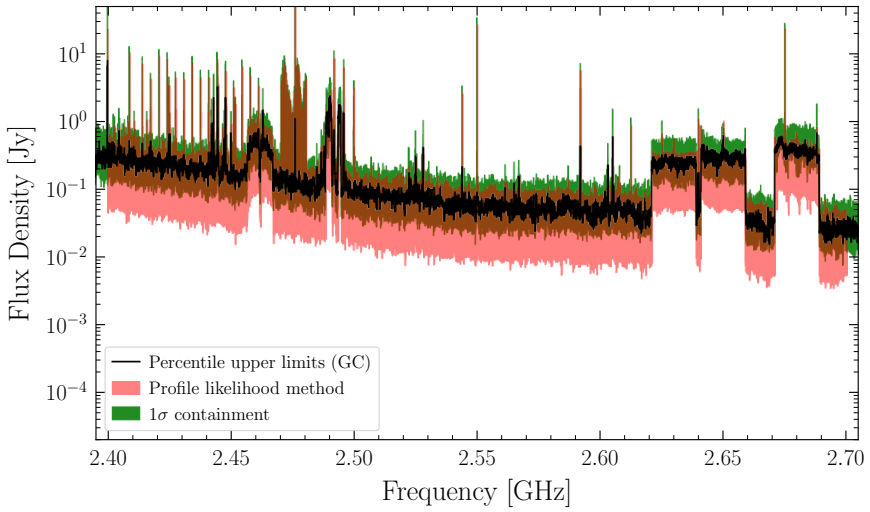



Figure F.17: Same as Fig. F.16, for Effelsberg GC observations.

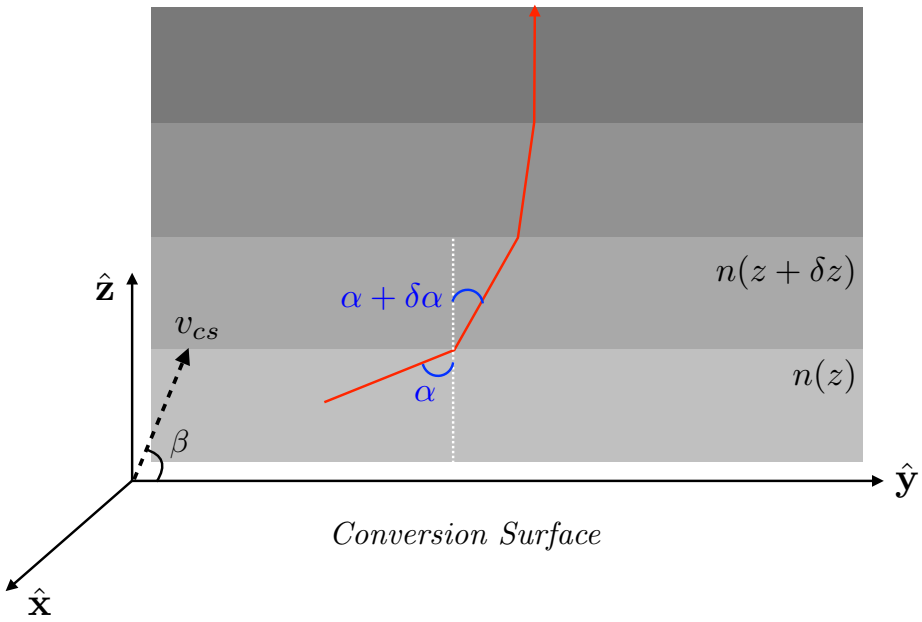



Figure F.18: An illustration of how outgoing electromagnetic waves are refracted towards the normal vector to the conversion surface, labeled here by $\hat{\mathbf{z}}$.

to photons in the *outgoing* direction. A key point to note is that the plasma frequency profile of the magnetosphere may be interpreted as a spatially-dependent index of refraction. For example, for the polarization component parallel to the magnetic field direction we may write the index of refraction as

$$n \approx \sqrt{\frac{1-\left(\frac{r_c}{r}\right)^3}{1-\left(\frac{r_c}{r}\right)^3 \cos^2\widetilde{\theta}}}, \tag{F.11}$$

for $r > r_c$, where $r_c$ is the conversion radius, $r$ is the radial direction, and $\widetilde{\theta}$ is the angle between the magnetic field and the propagation direction. The index of refraction is anisotropic, with a dependence on $\widetilde{\theta}$, because the plasma is strongly magnetized. Note that the group velocity is always smaller than the speed of light, consistent with special relativity.

Locally around the conversion surface, we are interested in describing the scenario illustrated in Fig. F.18. We choose coordinates such that the tangent plane to the conversion surface is spanned by the unit vectors $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, with the normal given by $\hat{\mathbf{z}}$, with $\hat{\mathbf{y}}$ chosen such that the magnetic field lies in the $\hat{\mathbf{y}}$-$\hat{\mathbf{z}}$ plane. In the large-field limit the two linearly independent polarization states do not mix, so to locally describe the trajectory of outgoing electromagnetic waves, we only need to consider the waves propagating in the $\hat{\mathbf{x}}$-$\hat{\mathbf{z}}$ plane. Without loss of generality, we imagine that the local conversion surface is traveling at a speed $v_{cs}$ in the $\hat{\mathbf{x}}$-$\hat{\mathbf{z}}$ plane at an angle $\beta$ from the conversion surface, as indicated in Fig. F.18. We will work to leading order in the speed $v_{cs}$, in

natural units. We are interested in two properties of the outgoing wave: (i) the outgoing angle with respect to the normal $\hat{\mathbf{z}}$, given an initial angle $\alpha_i$ near the conversion surface, and (ii) the frequency shift $\Delta\omega \equiv \omega(r = \infty) - \omega(r = r_c)$ between the wave asymptotically far from the conversion surface and the wave at the conversion surface.

The outgoing wave quickly turns towards the $\hat{\mathbf{z}}$ direction, as illustrated in Fig. F.18, because radiation refracts towards the direction of increasing index of refraction and $n$ increases with the distance from the conversion surface. More concretely, by considering a differential form of Snell's law one may show that (assuming $\widetilde{\theta} = \pi/2$ for simplicity and since having non-trivial $\widetilde{\theta}$ does not qualitatively change these results)

$$\frac{dz}{dy} = \sqrt{\frac{1 - z^{-3}}{1 - z_i^{-3}} \frac{1}{\sin(\alpha_i)^2} - 1}\,, \tag{F.12}$$

where $z(y)$ is the trajectory of the wave that starts a distance $z_i$ from the conversion surface at an initial angle $\alpha_i$. In the limit $z_i \to 0$, all trajectories asymptotically approach the $\hat{\mathbf{z}}$ direction, regardless of $\alpha_i$. In fact, $z_i \sim r_c v^2$ [103], where $v \sim 0.1$ is the axion velocity at the conversion surface in the frame of the NS, which is close enough to zero in practice that the asymptotics hold. This is also true regardless of the magnetic field direction; the full differential equation for the trajectory is a complicated nonlinear equation because $\widetilde{\theta}$ depends on the trajectory, but such dependence is washed out because of the sharp change of index of refraction very near the conversion surface. This itself is interesting because it says that while the initial $\alpha_i$ are isotropically distributed, since the DM phase space is isotropic, the outgoing radiation is collimated in the direction normal to the local conversion surface.

Next, we consider the frequency shift $\Delta\omega$ induced by the finite velocity $v_{cs}$ of the medium, to lowest order in $v_{cs}$. Roughly speaking, such a frequency shift results from the electromagnetic wave being created in an index of refraction which is already moving, and being measured in a stationary frame at infinity where the index of refraction is unity. To derive a differential equation for the evolution of the frequency it is useful to consider a differential step as shown in Fig. F.18 whereby we transition from a layer at distance $z$ to one at distance $z + \delta z$, with initial angle $\alpha$ and refracted angle $\alpha + \delta\alpha$. We perform the following set of steps. Let the frequency of the initial state be $\omega(z)$. First, we perform a Galilean boost by $v_{cs} \sin\beta$ in the $\hat{\mathbf{z}}$ direction so that the conversion surface is stationary in that direction. Under this boost the material becomes birefringent, with an angle-dependent index of refraction [447]. To leading order in $v_{cs}$ the index of refraction that the incoming wave sees in the boosted frame is $\tilde{n}(z) = n(z) + (n^2(z) - 1)v_{cs}\cos\alpha(z)\sin\beta$. The frequency in the boosted frame (to this order in $v_{cs}$) is $\tilde{\omega}(z) = \omega(1 - n(z)v_{cs}\cos\alpha(z)\sin\beta)$. We may then use Snell's law to refract the wave over the interface, where the material has index of refraction $n(z + \delta z)$. This changes the angle $\alpha$ but does not change the frequency. Then, we

boost again by $v_{cs} \sin\beta$ but now in the negative $\hat{\mathbf{z}}$ direction. Taking the limit $\delta z \to 0$ we find the differential equation:

$$\frac{d \log \omega(z)}{dz} = v_{cs} \sin\beta \left[ n'(z) \cos\alpha(z) - n(z) \alpha'(z) \sin\alpha(z) \right] . \tag{F.13}$$

In practice, since $\alpha(z)$ quickly approaches 0 (see Fig. F.18) the second term tends to be subdominant to the first, which remains non-zero in the limit $\alpha \to 0$. In this approximation, taking $\alpha = 0$, the right-hand side of (F.13) is a total derivative, and thus $\omega(z)$ only depends on the difference in index of refraction between the conversion surface (approximately zero) and infinity (approximately unity): $\delta\omega/\omega = v_{cs} \sin\beta$, to leading order in $v_{cs}$. Again, in the limit $z_i \to 0$ this result is independent of the initial angle $\alpha_i$ and independent of the anisotropy of the index of refraction from the magnetic field direction.

There are few interesting implications of this result. First, when averaging the axion signal over the phase of the NS rotation there will be a frequency broadening induced by the spread in $v_{cs} \sin\beta$ across the conversion surface, appropriately averaged. For the INSs considered in the main Letter and assuming misalignment angles $\sim 45^\circ$, we find that the frequency broadening is less than $\delta f/f \sim 5 \times 10^{-6}$ at 68% containment for both NSs (not much larger than the intrinsic bandwidth $\delta f/f \sim v_0^2 \sim 10^{-6}$), justifying the bandwidths used in our fiducal analyses. Second, this result reasserts the possibility of strong time dependence of the signal over the NS period, since the outgoing radiation is beamed normal to the conversion surface by refraction. While Ref. [288] claimed that the outgoing signal would not be strongly time-dependent because of the fact that the DM velocity distribution is isotropic, we have shown here that this result is modified due to the refraction of the outgoing photons. Moreover, since the frequency shift appears to leading order to be independent of the initial angle relative to the conversion surface normal vector, it is possible that when phase-resolved, the radio signal again becomes order $v_0^2$ wide in terms of $\delta\omega/\omega$, with a central frequency that shifts by an amount $\delta\omega/\omega \sim v_{cs}$ over the period. We leave both the theoretical analysis and an investigation of this effect in the data to future work.

# APPENDIX G

# X-Ray Searches for Axions from Super Star Clusters

## G.1 Methods: Data Reduction, Analysis, Simulations, and Calculations

In this section we first provide additional details needed to reproduce our NuSTAR data reduction, before giving extended discussions of our MESA simulations, axion luminosity calculations, and conversion probability calculations.

### G.1.1 Data Reduction and analysis

To perform the NuSTAR data reduction, we use the NuSTARDAS software included with HEASoft 6.24 [358]. We first reprocess the data with the NuSTARDAS task `nupipeline`, which outputs calibrated and screened events files. We use the strict filtering for the South Atlantic Anomaly. We then create counts maps for both focal plane modules (FPMs) of the full NuSTAR FOV with `nuproducts` in energy bins of width $5$ keV from $5-80$ keV.[1] We additionally generate the ancillary response files (ARFs) and the redistribution matrix files (RMFs) for each FPM. We generate the corresponding exposure maps with `nuexpomap`, which produces exposure maps with units [s]. To obtain maps in exposure units [cm$^2$ s keV] that we can use to convert from counts to flux, we multiply in the mean effective areas in each bin with no PSF or vignetting correction.

Once the data is reduced, we apply the analysis procedure described in the main text to measure the spectrum associated with the signal template in each energy bin. However, to compare the signal-template spectrum to the axion model prediction, we need to know how to forward-model the predicted axion-induced flux, which is described in more detail later in the SM, through the

[1] We use 5 keV-wide energy bins as a compromise between having narrow energy bins that allow us to resolve the spectral features in our putative signal (see Fig. 8.2) and having wide-enough bins that allow to accurately determine the background template normalizations in our profile likelihood analysis procedure. However, small-to-moderate changes to the bins sizes (*e.g.*, increasing them by a factor of 2) lead to virtually identical results.

instrument response. In particular, we pass the signal flux prediction through the detector response to obtain the expected signal counts that we can compare to the data:

$$\mu^e_{S,i}(\boldsymbol{\theta}_{\rm S}) = t^e \int dE' {\rm RMF}^e_i(E') {\rm ARF}^e(E') S(E'|\boldsymbol{\theta}_{\rm S})\,. \tag{G.1}$$

Here, $t^e$ is the exposure time corresponding to the exposure $e$ in [s], while the signal is the expected intensity spectrum in [erg/cm$^2$/s/keV]. We have now obtained the expected signal counts $\mu^e_{S,i}(\boldsymbol{\theta}_{\rm S})$ that may be integrated into the likelihood given in (8.1).

### G.1.2 MESA Simulations

MESA is a one-dimensional stellar evolution code which solves the equations of stellar structure to simulate the stellar interior at any point in the evolution. In our fiducial analysis, we construct models at a metallicity Z = 0.035, initial stellar masses from 15 to 200 $M_\odot$, and initial surface rotations from 0 km/s to 500 km/s as indicated in the main text. We use the default inlist for high-mass stars provided with MESA. This inlist sets a number of parameters required for high-mass evolution, namely the use of Type 2 opacities. We additionally use the Dutch wind scheme [448] as in the high rotation module.

On this grid, we simulate each star from the pre-MS phase until the onset of neon burning around $1.2 \times 10^9$ K. At that point, the star only has a few years before undergoing supernova. Given that no supernova has been observed in the SSCs since the observations in 2012-2015, this end-point represents the most evolved possible state of stars in the SSCs at time of observation. The output is a set of radial profiles at many time steps along the stellar evolution. The profiles describe, for example, the temperature, density, and composition of the star. These profiles allow us to compute the axion spectrum at each time step by integrating the axion volume emissivity over the interior.

Here we show detailed results for a representative star of mass 85 $M_\odot$ with initial surface rotation of 300 km/s. This star is a template star for the WC phase (and other WR phases) in the Quintuplet Cluster, which dominates the Quintuplet axion spectrum in the energies of interest. In the left panel of Fig. G.1, we show the Hertzsprung–Russell (HR) diagram for our template star. The star's life begins on the MS, where it initiates core hydrogen burning. Eventually, the core runs out of hydrogen fuel and is forced to ignite helium to prevent core collapse (see Fig. G.2 left). Because helium burns at higher temperatures, the star contracts the core to obtain the thermal energy required to ignite helium (see Fig. G.3). At the same time, the radiation pressure in stellar winds cause heavy mass loss in the outer layers, which peels off the hydrogen envelope (see Fig. G.4). When the surface is 40% hydrogen, the star enters the WNh phase; when it is 5% hydrogen, the star enters the WN phase. Further mass loss begins to peel off even the helium layers, and the star

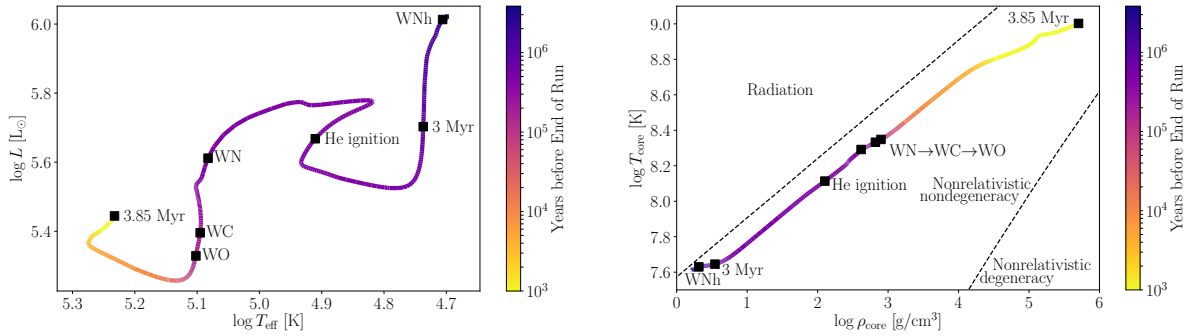



Figure G.1: (Left) The HR diagram for the Quintuplet template star of mass 85 $M_\odot$ and initial surface rotation of 300 km/s. The coloring indicates the year before the run was stopped, approximately a few years from supernova. We mark with black squares, in order of occurrence, when the star enters the WNh phase, when it is 3 Myr old, when its core undergoes helium ignition, when it enters the WN, WC, and WO phases, and finally when the run ends at 3.85 Myr. (Right) A $\log$T-$\log\rho$ diagram for the template star with the same points of interest marked. We also show the relevant degeneracy zones, showing that the star is entirely in the nonrelativistic nondegenerate regime.

enters the WC and WO phases when its surface is 2% carbon and oxygen by abundance [335], respectively (see Fig. G.2 right).

### G.1.3 Axion Production in SSCs

In this section we overview how we use the output of the MESA simulations to compute axion luminosities and spectra.

#### G.1.3.1 The Axion Energy Spectrum

Here we focus on the calculation of the axion energy spectrum [erg/cm$^2$/s/keV]. The axion production rate is [367]

$$\Gamma_p(E) = \frac{g_{a\gamma\gamma}^2 T\kappa^2}{32\pi}\left[\left(1+\frac{\kappa^2}{4E^2}\right)\ln\left(1+\frac{4E^2}{\kappa^2}\right)-1\right], \tag{G.2}$$

where $\kappa^2 = \frac{4\pi\alpha}{T}\sum_i Z_i^2 n_i$ gives the Debye screening scale, which is the finite reach of the Coulomb field in a plasma and cuts off the amplitude. To obtain the axion energy spectrum, this is to be

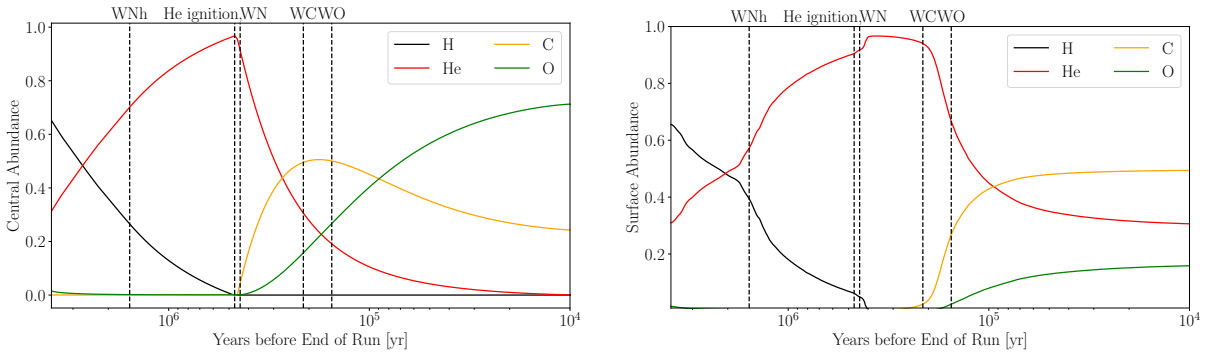



Figure G.2: (Left) The abundances of hydrogen (black), helium (red), carbon (yellow), and oxygen (green) in the center of the star as a function of time, for the simulation described in Fig. G.1. With dashed-black vertical lines, we mark several points of interest: "WNh" indicates the time the star enters the WNh phase, "He ignition" when its core undergoes helium ignition, and "WN","WC", and "WO" indicate the beginning of the WN, WC, and WO phases, respectively. (Right) The same as in the left panel, but for surface abundances.

convolved with the photon density, such that

$$\begin{aligned}\frac{dL_p}{dE}(E) &= \frac{1}{\pi^2}\frac{E^3}{e^{E/T}-1}\Gamma_p(E) \\ &= \frac{g_{a\gamma\gamma}^2}{8\pi^3}\frac{\xi^2 T^3 E}{e^{E/T}-1}\left[\left(E^2+\xi^2T^2\right)\ln\left(1+\frac{E^2}{\xi^2T^2}\right)-E^2\right],\end{aligned} \tag{G.3}$$

where we have defined the dimensionless parameter $\xi^2 = \dfrac{\kappa^2}{4T^2}$. To obtain the axion emissivity for a whole star, we integrate over the profiles produced with MESA, and we show results for this calculation in the next section. Finally, the axion-induced photon spectrum at Earth is given by

$$\frac{dF}{dE}(E) = P_{a\to\gamma}(E)\frac{1}{4\pi d^2}\frac{dL_a}{dE}(E)\,, \tag{G.4}$$

with the conversion probability $P_{a\to\gamma}$ computed later.

### G.1.3.2 Results for Template Star

In this section, we show our expectation for the axion luminosity from our template star.

In the left panel of Fig. G.5, we show the axion emissivity from the radial slices of the MESA profile, using the model at the start of the WC evolutionary stage. As expected, the stellar core is by far the most emissive due to its high temperature and density. We also show the temperature profile in the star. Note that the axion volume emissivity does not have the same profile shape as the temperature because the emissivity also depends on the density and composition which are

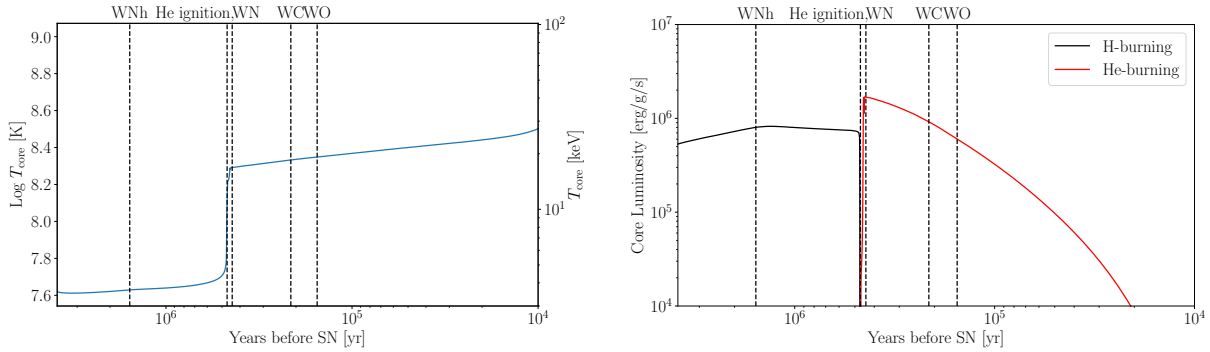



Figure G.3: (Left) The stellar core temperature as a function of time for the simulation described in Fig. G.1. (Right) The hydrogen and helium luminosities in the core through the CNO cycle and the triple-alpha process, respectively. The dashed-black vertical lines retain their meanings from Fig. G.2.

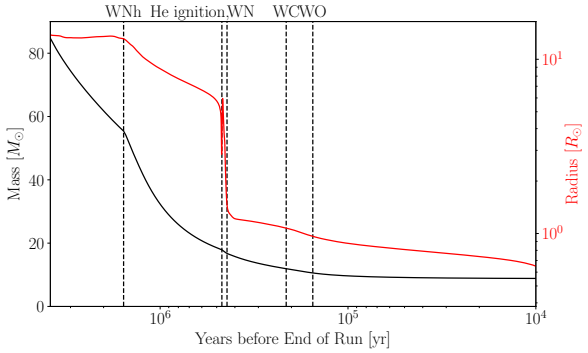



Figure G.4: The stellar mass (black) and radius (red) as a function of time from the simulation described in Fig. G.1. The dashed-black vertical lines retain their meanings from Fig. G.2.

highly nonuniform over the interior.

In the right panel of Fig. G.5, we show how the axion luminosity changes over the stellar lifetime. We see that before helium ignition, the axion luminosity is rather low, and the axion spectrum reaches its maximum around 10 keV, owing to the low core temperature—the star is still hydrogen burning at core temperatures well below 10 keV. During helium ignition, the luminosity increases quickly due to the sudden increase in temperature. During helium burning, the core temperature continues to increase; for this reason, more evolved stars will be more luminous in axions.

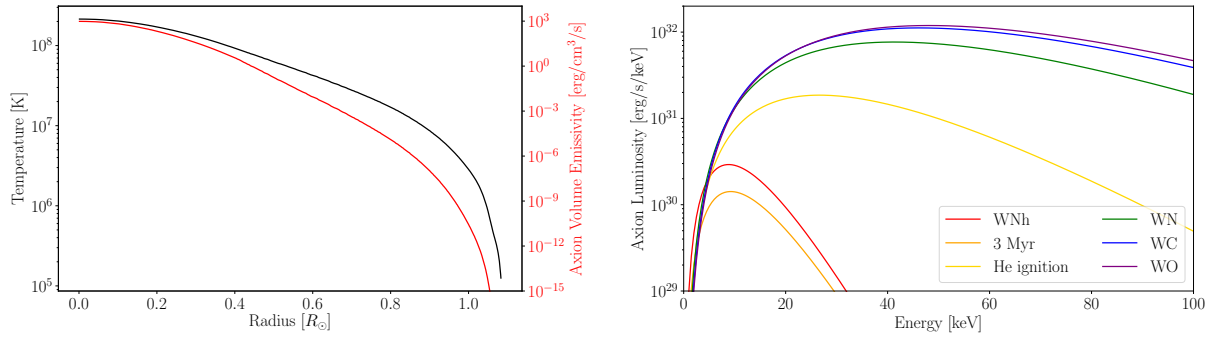



Figure G.5: (Left) Axion volume emissivity over the interior of the star. In this figure we have taken the stellar model to be the one at the start of the WC stage and fixed $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$. For comparison purposes, we also show the temperature profile. (Right) Axion luminosity spectrum for those same stages marked in Fig. G.2.

### G.1.4 Magnetic field model and conversion probability

When the axion-to-photon conversion probability $p_{a\to\gamma}$ is sufficiently less than unity, it may be approximated by [341]:

$$p_{a\to\gamma} = \frac{g_{a\gamma\gamma}^2}{4} \sum_{i=1,2} \left| \int_0^d dr' B_i(r') e^{i\Delta_a r' - i\int_0^{r'} dr'' \Delta_{||}(r'')} \right|^2 , \tag{G.5}$$

where $B_i$, for $i = 1, 2$, denote the two orthogonal projections of the magnetic field onto axes perpendicular to the direction of propagation. The integrals are over the line of sight, with the source located a distance $d$ from Earth, and $r = 0$ denoting the location of the source. We have also defined $\Delta_a \equiv -m_a^2/(2E)$ and $\Delta_{||}(r) \equiv -\omega_{\rm pl}(r)^2/(2E)$, with $E$ the axion energy and $\omega_{\rm pl}(r)$ the location-dependent plasma mass. The plasma mass may be related to the number density of free electrons $n_e$ by $\omega_{\rm pl} \approx 3.7 \times 10^{-12} (n_e/10^{-2}\ {\rm cm}^{-3})^{-1/2}$ eV. To perform the integral we need to know (i) the free electron density along the line of sight to the target, and (ii) the orthogonal projections of the magnetic field along the line-of-sight. In this section we give further details behind the electron-density and magnetic-field profiles used in this Letter.

The Quintuplet and Arches SSCs are both $\sim$30 pc away from the GC and thus are expected to have approximately the same conversion probabilities for conversion on the ambient Galactic magnetic fields. It is possible, however, that local field configurations near the GC could enhanced the conversion probabilities for one or both of these sources. For example, the axions are expected to travel through or close to the GC radio arc, which has a strong magnetic field $\sim$mG over a cross-section $\sim(10\ {\rm pc})^2$ [355]. Magnetic fields within the clusters themselves may also be important.

Our fiducial magnetic field model for Quintuplet and Arches is illustrated in the left panel of

Fig. G.6. In the right panel we show the magnetic field profiles relevant for the Wd1 observations.

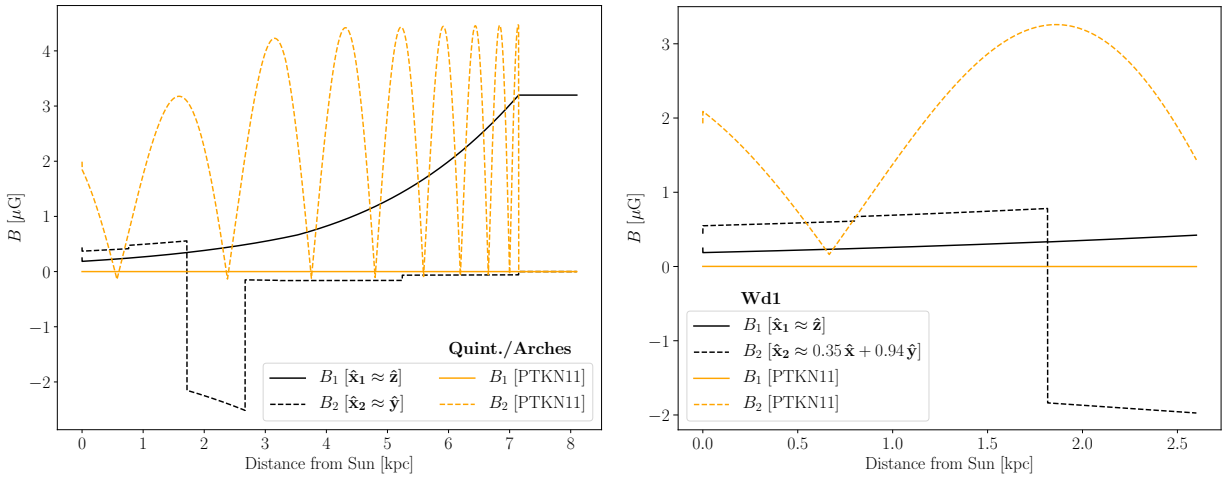



Figure G.6: We denote the projections of the Galactic magnetic field onto the plane normal to the propagation direction by $B_1$, $B_2$. (Left) The transverse magnetic field components in our fiducial model (the JF12 model, black) and alternate model (PTKN11, orange) towards the Quintuplet and Arches clusters. Note that in our fiducial $B$-field model we extend the JF12 model to distances less than 1 kpc from the GC using the field values at 1 kpc. The true magnetic field values in the inner kpc almost certainly surpass those from this conservative model (see text for details). (Right) The two field components towards the Wd1 cluster, which is taken to be at a distance of 2.6 kpc from the Sun. The conversion probabilities towards Wd1 are much larger in the alternate model (PTKN11) than in our fiducial model (JF12), though we stress that random fields are not included and could play an important role in the conversion probabilities towards Wd1.

The components of the $B$-field along the two transverse directions are denoted by $B_1$ and $B_2$. For the Quintuplet and Arches analyses, the propagation direction is very nearly aligned with $-\hat{\mathbf{x}}$ (in Galactic coordinates), so we may take $B_1$ to point in the $\hat{\mathbf{z}}$ direction, towards the north Galactic pole, and $B_2$ to point in the direction $\hat{\mathbf{y}}$ (the approximate direction of the local rotation). Note that the targets are slightly offset from the origin of the Galactic coordinate system, so the actual basis vectors have small components in the other directions. As Wd1 is essentially within the plane of the disk, one of the transverse components points approximately in the $\hat{\mathbf{z}}$ direction ($B_1$).

The dominant magnetic field towards the GC within our fiducial $B$-field model is the vertical direction ($B_1$), which is due to the out-of-plane $X$-shaped halo component in the JF12 model [343, 344]. However, in the JF12 model that component is cut off within 1 kpc of the GC, due to the fact that in becomes difficult to model the $B$-field near the GC. The $B$-field is expected to continue rising near the GC – for example, in [353] it was claimed that the $B$-field should be at least 50 $\mu$G (and likely 100 $\mu$G) within the inner 400 pc. However, to be conservative in our fiducial $B$-field model we simply extend the $B$-field to the GC by assuming it takes the value at 1 kpc (about 3

$\mu$G) at all distances less than 1 kpc from the GC. We stress that this field value is likely orders of magnitude less than the actual field strength, but this assumption serves to make our results more robust. The extended field model is illustrated in Fig. G.6.

To understand the level of systematic uncertainty arising from the $B$-field models we also show in Fig. G.6 the magnetic field profiles for the alternative ordered $B$-field model PTKN11 [347]. This model has no out-of-plane component, but the regular $B$-field within the disk is stronger than in the JF12 model. In the case of Quintuplet and Arches we find, as discussed below, that the PTKN11 model leads to similar but slightly enhanced conversion probabilities relative to the JF12 model. On the other hand, the conversion probabilities in the PTKN11 model towards Wd1 are significantly larger than in the JF12 model.

There is a clear discrepancy in Fig. G.6 between the magnetic field values observed at the solar location, in both the JF12 model and the PTKN11 model, and the local magnetic field strength, which is $\sim$3 $\mu$G [349]. The reason is that the magnetic field profiles shown in Fig. G.6 are only the regular components; additional random field components are expected. For example, in the JF12 model the average root-mean-square random field value at the solar location is $\sim$6.6 $\mu$G [343, 344]. The random field components could play an important role in the axion-to-photon conversion probabilities, especially for the nearby source Wd1, but to accurately account for the random field components one needs to know the domains over which the random fields are coherent. It is expected that these domains are $\sim$100 pc [344], in which case the random fields may dominate the conversion probabilities, but since the result depends sensitively on the domain sizes, which remain uncertain, we conservatively neglect the random-field components from the analyses in this work (though this would be an interesting subject for future work).

To compute the conversion probabilities we also need the free-electron densities. We use the YMW16 model [345] as our fiducial model, but we also compare our results to those obtained with the older `ne2001` model [346] to assess the possible effects of mismodeling the free-electron density. In the left panel of Fig. G.7 we compare the free electron densities between the two models as a function of distance away from the Sun towards the GC, while in the right panel we show the free electron densities towards Wd1. The differences between these models result in modest differences between the computed conversion probabilities, as discussed below.

Combining the magnetic field models in Fig. G.6 and the free-electron models in Fig. G.7 we may compute the axion-photon conversion probabilities, for a given axion energy $E$. These conversion probabilities are presented in the left panels of Fig. G.8 (assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ and $m_a \ll 10^{-11}$ eV). In the top left panel we show the results for Quintuplet and Arches, while the bottom left panel gives the conversion probabilities for Wd1, computed under both free-electron models and various magnetic field configurations.

In the top left panel our fiducial conversion probability model is shown in solid black. Changing

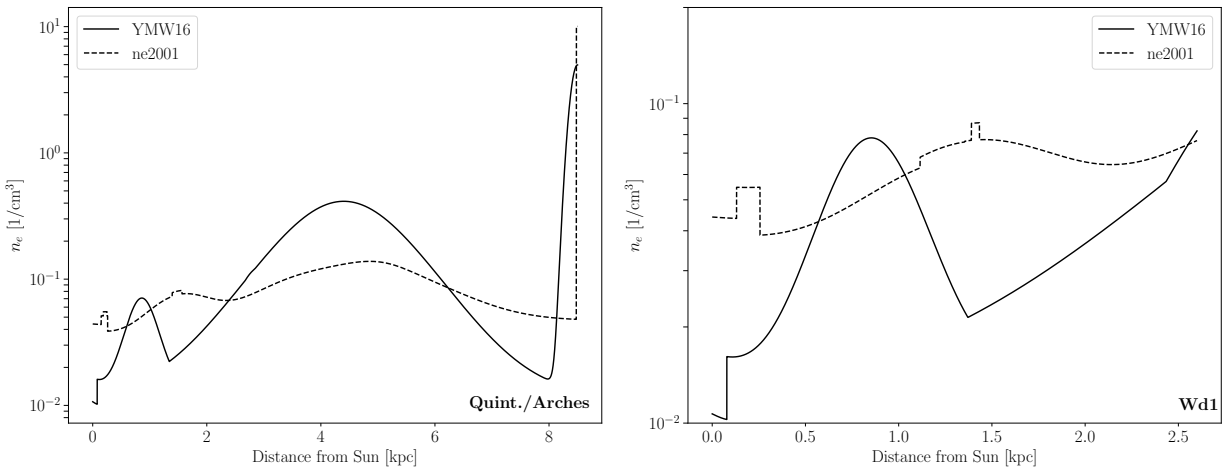



Figure G.7: (Left) The free electron density $n_e$ towards the GC in our fiducial model (YMW16) and the alternate model (`ne2001`). (Right) As in the left panel but towards the Wd1 cluster. The free-electron density gives the photon an effective mass and thus affects the axion-photon conversion probability.

to the `ne2001` model would in fact slightly enhance the conversion probabilities at most energies, as shown in the dotted black, though the change is modest. Completely removing the $B$-field within 1 kpc of the GC leads only to a small reduction to the conversion probabilities, as indicated in red. Changing magnetic field models to that of [347] (PTKN11), while also removing the $B$-field within the inner kpc, leads to slightly enhanced conversion probabilities, as shown in orange (for both the YMW16 and `ne2001` $n_e$ models). Note that the conversion probabilities exhibit clear constructive and destructive interference behavior in this case at low energies, related to the periodic nature of the disk-field component, though including the random field component it is expected that this behavior would be largely smoothed out.

As discussed previously the magnetic field is expected to be significantly larger closer in towards the GC than in our fiducial $B$-field model. As an illustration in blue we show the conversion probabilities computed, from the two different free-electron models, when we only include a $B$-field component of magnitude $50$ $\mu$G pointing in the $\hat{\mathbf{z}}$ direction within the inner 400 kpc (explicitly, in this case we do not include any other $B$-field model outside of the inner 400 kpc). The conversion probabilities are enhanced in this case by about an order of magnitude across most energies relative to in our fiducial model. The inner Galaxy also likely contains localized regions of even strong field strengths, such as non-thermal filaments with $\sim$mG ordered fields. As an illustration of the possible effects of such fields on the conversion probabilities, in Fig. G.8 we show in grey the result we obtain for the conversion probability when we assume that the axions traverse the GC radio arc, which we model as a 10 kpc wide region with a vertical field strength of 3 mG

and a free-electron density $n_e = 10$ cm$^{-3}$ [355, 449]. Due to modeling uncertainties in the non-thermal filaments and the ambient halo field in the inner hundreds of pc, we do not include such magnetic-field components in our fiducial conversion probability model. However, we stress that in the future, with a better understanding of the Galactic field structure in the inner Galaxy, our results could be reinterpreted to give stronger constraints.

The Wd1 conversion probabilities change by over an order of magnitude going between the JF12 and PTKN11 models, as seen in the bottom left panel of Fig. G.8, though it is possible that this difference would be smaller when random fields are properly included on top of the JF12 model (though again, we chose not to do this because of sensitivity to the random-field domain sizes).

The effects of the different conversion probabilities on the $g_{a\gamma\gamma}$ limits may be seen in the top right panel for Quintuplet (Arches gives similar results, since the conversion probabilities are the same) and Wd1 in the bottom right panel of Fig. G.8. Note that the observed fluxes scale linearly with $p_{a\to\gamma}$ but scale like $g^4_{a\gamma\gamma}$, so differences between conversion probability models result in modest differences to the $g_{a\gamma\gamma}$ limits. Still, it is interesting to note that the Wd1 limits with the PTKN11 model are stronger than the fiducial Quintuplet limits, which emphasizes the importance of better understanding the $B$-field profile towards Wd1. For Quintuplet (and also Arches) we see that depending on the field structure in the inner $\sim$kpc, the limits may be slightly stronger and extend to slightly larger masses (because of field structure on smaller spatial scales) than in our fiducial $B$-field mode.

## G.2 Extended Data Analysis Results

In this section we present additional results from the data analyses summarized in the main Letter.

### G.2.1 Quintuplet

In this subsection we give extended results for the Quintuplet data analysis. Our main focus is to establish the robustness of the flux spectra from the NuSTAR data analysis (shown in Fig. 8.2) that go into producing the limits on $g_{a\gamma\gamma}$ shown in Fig. 8.3.

#### G.2.1.1 Data and templates

First we take a closer look at the stacked data and models that go into the Quintuplet data analysis. The stacked counts data in the vicinity of Quintuplet are shown in the left panel of Fig. G.9. We show the counts summed from 10 - 80 keV. Note that the circle in that figure indicates $2'$, which the

radius of our fiducial analysis ROI.[2] As in Fig. 8.1 we also indicate the locations of the individuals stars in Quintuplet that may contribute axion-induced $X$-ray flux. The middle panel shows the expected background flux from our background template. The template is generally uniform over the ROI, with small variations. On the other hand, the right panel shows the axion-induced signal counts template, normalized for $g_{a\gamma\gamma} = 7 \times 10^{-12}$ GeV$^{-1}$, which is localized about the center of the SSC. Note that the signal template is generated by accounting for the PSF of NuSTAR in addition to the locations and predicted fluxes of the individual stars.

#### G.2.1.2 Axion Luminosity

We now show the axion luminosity and spectra that go into the right panel of Fig. G.9. For each star in the cluster, we assign it a set of possible MESA models based on its spectral classification as described in the main text. In the upper left panel of Fig. G.10, we show the mean expected axion luminosity, as a function of energy, of the Quintuplet cluster, assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$. The luminosity peaks around 40 keV, but the effective area of NuSTAR, also shown, rapidly drops above 10 keV. Due to the much higher effective area at low energies, most of the sensitivity is at lower energies. There is also considerable flux above 80 keV, although NuSTAR does not have sensitivity at these energies. In the upper right panel, we show the median contribution of each spectral classification in Quintuplet to this luminosity, summed over all stars with the given classification. For all energies of interest, the WC stars dominate the cluster luminosity. This is because WR stars have the hottest cores and there are 13 WC stars in Quintuplet (there is 1 WN star). In the bottom panel, we show the 10 - 80 keV luminosity distribution for each spectral classification, along with the $1\sigma$ containment bands and the mean expectation. The distribution depends principally on whether or not core helium is ignited while the star is assigned a given classification. The O, BSG, and WNh stars all can be either hydrogen or helium burning, in which case they have 10 - 80 keV luminosities of $\sim 10^{31}$ or $\sim 10^{33}$ erg/s, respectively—recall that the jump in temperature during helium ignition is a factor $\sim 3$. The LBV phase is always core helium burning, and the star may go supernova in this phase. The same is true of the WR phases WN and WC, although the stars undergoing supernova in this phase are typically more massive.

The luminosities in Fig. G.10 are computed for our fiducial choices of $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s. To better understand the importance of these choices we show in Tab. G.1 how the luminosities depend on the initial metallicity $Z$ and mean rotation speed $\mu_{\rm rot}$. Note that each entry in that table shows the luminosity summed over the stellar sub-types (with the number of stars indicated), and except in the two last columns the luminosities are summed over all stars. The uncertainties in the entries in Tab. G.1 come from performing 500 draws from the representative

[2]Note that ROIs for all of our analyses are centered upon the center of axion fluxes in RA and DEC, though the distinction between the center of fluxes and the SSC center is minimal for all of our targets.

| | O | BSG | LBV | WNh | WC + WN | tot | tot (10-80 keV) |
|---|---|---|---|---|---|---|---|
| $N_{\rm star}$ | 37 | 7 | 2 | 5 | 14 | 65 | 65 |
| $z = 0.018$ $\mu_{\rm rot} = 100$ km/s | $3.0^{+1.7}_{-1.5} \times 10^{33}$ | $1.3^{+0.9}_{-0.9} \times 10^{33}$ | $1.9^{+2.1}_{-1.7} \times 10^{34}$ | $5.9^{+5.8}_{-5.8} \times 10^{34}$ | $2.8^{+2.6}_{-0.8} \times 10^{35}$ | $3.8^{+2.6}_{-1.0} \times 10^{35}$ | $2.8^{+1.6}_{-0.6} \times 10^{35}$ |
| $z = 0.035$ $\mu_{\rm rot} = 100$ km/s | $1.9^{+2.9}_{-0.9} \times 10^{34}$ | $3.5^{+1.2}_{-1.2} \times 10^{33}$ | $1.4^{+1.1}_{-0.7} \times 10^{34}$ | $7.4^{+30}_{-7.3} \times 10^{33}$ | $1.7^{+0.9}_{-0.4} \times 10^{35}$ | $2.3^{+0.9}_{-0.5} \times 10^{35}$ | $1.7^{+0.5}_{-0.3} \times 10^{35}$ |
| $z = 0.035$ $\mu_{\rm rot} = 150$ km/s | $3.4^{+2.4}_{-2.3} \times 10^{34}$ | $3.6^{+1.2}_{-1.3} \times 10^{33}$ | $1.4^{+1.2}_{-0.8} \times 10^{34}$ | $4.3^{+22}_{-4.2} \times 10^{33}$ | $1.5^{+0.7}_{-0.3} \times 10^{35}$ | $2.1^{+0.7}_{-0.4} \times 10^{35}$ | $1.7^{+0.4}_{-0.3} \times 10^{35}$ |

Table G.1: The number of stars $N_{\rm star}$ for each stellar class in the Quintuplet cluster, along with the predicted axion luminosities (all in erg/s). Note that Quintuplet is $\sim$30 pc away from the GC. Except in the last column, the axion luminosities are summed over all energies. All entries assume $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ and are summed over all stars for the given stellar class.

models and account for the variance expected from star-to-star within a given classification. As discussed in the main text, the 10 - 80 keV luminosity could be $\sim$70% larger than in our fiducial model, depending on the initial $Z$ and $\mu_{\rm rot}$.

#### G.2.1.3 Injecting an axion signal

As a first test of the robustness of the Quintuplet analysis we inject a synthetic axion signal into the real stacked data and then pass the hybrid real plus synthetic data through our analysis pipeline. Our goal from this test is to ensure that if a real axion signal were in the data with sufficiently high coupling to photons then we would be able to detect it. The results from this test are shown in Fig. G.11.

The left panel of Fig. G.11 shows the best-fit $g^{\rm rec.}_{a\gamma\gamma}$ as a function of the simulated $g^{\rm inj.}_{a\gamma\gamma}$ used to produce the axion-induced counts that are added to the real NuSTAR stacked data. Importantly, as we increase the injected signal strength the recovered signal parameter converges towards the injected value, which is indicated by the dashed curve. Note that the band shows the 68% containment region for the recovered signal parameter from the analysis. As the injected signal strength increases, so to does the significance of the axion detection. This is illustrated in the middle panel, which shows the discovery TS as a function of the injected signal strength. Recall that the significance is approximately $\sqrt{\rm TS}$. Perhaps most importantly, we also verify that the 95% upper limit does not exclude the injected signal strength. In the right panel of Fig. G.11 we show the 95% upper limit found from the analyses of the hybrid data sets at different $g^{\rm inj}_{a\gamma\gamma}$. Recall that all couplings above the $g^{\rm rec}_{a\gamma\gamma}$ curve are excluded, implying that indeed we do not exclude the injected signal strength. Moreover, the 95% upper limit is consistent with the expectation for the limit under the signal hypothesis, as indicated by the shaded regions at $1\sigma$ (green) and $2\sigma$ (yellow) containment. Note that we do not show the lower $2\sigma$ containment region, since we power-constrain the limits. These regions were computed following the Asimov procedure [195].

#### G.2.1.4 Changing region size

As a systematic test of the data analysis we consider the sensitivity of the inferred spectrum associated with the axion model template to the ROI size. In our fiducial analysis, with spectrum shown in Fig. 8.2, we use an ROI size of $r_{\rm max} = 2'$. Here we consider changing the ROI size to $r_{\rm max} = 1.5'$ and $2.5'$. The resulting spectra are shown in Fig. G.12. The spectrum does not appear to vary significantly when extracted using these alternate ROIs, indicating that significant sources of systematic uncertainty related to background mismodeling are likely not at play.

### G.2.2 Westerlund 1

In this subsection we provide additional details and cross-checks of the Wd1 analysis.

#### G.2.2.1 Data and templates

In Fig. G.13 we show, in analogy with Fig. G.9, the data, background, and signal maps summed from 15 - 80 keV. We note that the background templates are summed using their best-fit normalizations from the fits to the null hypothesis of background-only emission. The signal template is noticeably extended in this case beyond a point-source template and is shown for $g_{a\gamma\gamma} = 8 \times 10^{-12}$ GeV$^{-1}$ and $m_a \ll 10^{-11}$ eV. The location of the magnetar CXOU J164710.2–45521 is indicated by the red star.

#### G.2.2.2 Axion Luminosity

We now show the axion luminosity and spectra that go into the right panel of Fig. G.13. In the upper left panel of Fig. G.14, we show the mean expected axion luminosity, as a function of energy, of the Wd1 cluster, assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$. In the upper right panel, we show the contribution of each spectral classification in Wd1 to this luminosity, summed over all stars with the given classification. For all energies of interest, the WN stars dominate the cluster luminosity, although the WC stars are important as well. As in Quintuplet, this is due to the fact that WR stars have the hottest cores, but in this case there are more WN stars than WC stars. In the bottom panel, we show the 10 - 80 keV luminosity distribution for each spectral classification, along with the $1\sigma$ bands and the mean expectation. Again, the more evolved stars produce more axion flux, because their core temperatures increase with time. As in the case of Quintuplet, the O and BSG stars may be pre- or post-helium ignition. The luminous blue variable (LBV), yellow hypergiant (YHG), and cool red supergiant (RSG) stars are all post-helium ignition, although have generically cooler cores than the WR stars. The WNh stars are entirely helium burning.

| | O | (B/R)SG/YHG | LBV | WNh | WC/WN | tot | tot (10-80 keV) |
|---|---|---|---|---|---|---|---|
| $N_{\rm star}$ | 72 | 56 | 1 | 2 | 22 | 153 | 153 |
| $z = 0.018$ $\mu_{\rm rot} = 100$ km/s | $1.6^{+0.9}_{-0.6} \times 10^{35}$ | $2.4^{+1.3}_{-0.8} \times 10^{35}$ | $1.4^{+2.8}_{-1.3} \times 10^{34}$ | $2.2^{+4.8}_{-1.5} \times 10^{35}$ | $4.3^{+1.7}_{-1.4} \times 10^{36}$ | $5.2^{+1.7}_{-1.4} \times 10^{36}$ | $1.3^{+0.2}_{-0.2} \times 10^{36}$ |
| $z = 0.035$ $\mu_{\rm rot} = 100$ km/s | $2.6^{+1.5}_{-1.1} \times 10^{35}$ | $3.9^{+2.6}_{-1.5} \times 10^{35}$ | $7.1^{+10}_{-6.5} \times 10^{33}$ | $8.7^{+37}_{-4.5} \times 10^{34}$ | $2.0^{+1.0}_{-0.7} \times 10^{36}$ | $3.1^{+1.1}_{-0.9} \times 10^{36}$ | $9.9^{+1.4}_{-1.4} \times 10^{35}$ |
| $z = 0.035$ $\mu_{\rm rot} = 150$ km/s | $2.3^{+1.3}_{-1.0} \times 10^{35}$ | $3.5^{+2.6}_{-1.5} \times 10^{35}$ | $7.1^{+9.0}_{-6.5} \times 10^{33}$ | $6.2^{+31}_{-2.8} \times 10^{34}$ | $1.8^{+1.0}_{-0.7} \times 10^{36}$ | $2.6^{+1.0}_{-0.8} \times 10^{36}$ | $9.0^{+1.0}_{-1.0} \times 10^{35}$ |

Table G.2: As in Tab. G.1 but for Wd1.

In Tab. G.2 we provide detailed luminosities for each of the stellar sub-types for different choices of initial $Z$ and $\mu_{\rm rot}$ for Wd1, as we did in Tab. G.1. Note that we assume $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s for our fiducial analysis, even though it is likely that the initial $Z$ is closer to solar (in which case the luminosities would be enhanced, as seen in Tab. G.2).

#### G.2.2.3 Systematics on the extracted spectrum

In analogy to the Quintuplet analysis we may profile over emission associated with the background template to measure the spectrum from 15 - 80 keV associated with the axion-induced signal template shown in Fig. G.13. That spectrum is reproduced in Fig. G.15. For our default analysis we use the ROI with all pixels contained with $r_{\rm max} = 2.0'$ of the cluster center, except for those in the magnetar mask, as indicated in Fig. G.13. However, as a systematic test we also compute the spectrum associated with the axion-induced template for $r_{\rm max} = 2.5'$ and $1.5'$, as shown in Fig. G.15. We measure a consistent spectrum across ROIs at these energies.

### G.2.3 Arches

In this subsection we present results from the analysis of archival NuSTAR data for an axion-induced signal from the Arches cluster. The Arches cluster is at a similar location, $\sim$30 pc from the GC, as the Quintuplet cluster. Arches hosts even younger and more extreme (*e.g.*, hotter and more massive) stars than the nearby Quintuplet cluster. Indeed, it is estimated that all $\sim$105 spectroscopically classified stars within Arches may become core-collapse supernovae within the next $\sim$10 Myr [450]. *A priori*, the Arches and Quintuplet clusters should have similar sensitivities to axions, though as we discuss below the axion prediction from Arches is less robust to uncertainties in the initial metallicity than the Quintuplet prediction.

#### G.2.3.1 Axion Luminosity

We now describe the axion luminosity and spectra for Arches. In the upper left panel of Fig. G.16, we show the mean expected axion luminosity, as a function of energy, of the Arches cluster, as-

suming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$. The luminosity peaks at very low energies, although we could not analyze these energies due to contamination from the molecular cloud. As shown by the upper right panel, the Arches luminosity is dominated by the O stars, since the WNh stars are always hydrogen burning with our assumed metallicity of $Z = 0.035$ and there are many more O stars than WNh stars. In the bottom panel, we show the 10 - 80 keV luminosity distribution for the O and WNh stars, along with the $1\sigma$ bands and the mean expectation.

However, unlike for the Quintuplet and Wd1 clusters we find that the Arches luminosity is a strong function of the initial metallicity $Z$, as illustrated in Tab. G.3. As seen in that table, changing the metallicity from $Z = 0.035$ to $Z = 0.018$ increases the flux by over an order of magnitude. This is because at the higher metallicity values the WNh stars are typically not in the He burning phase, while decreasing the initial metallicity slightly causes the WNh stars to enter the He burning phase. Note that at solar initial metallicity ($Z = 0.02$, and also taking $\mu_{\rm rot} = 100$ km/s) we find that the 10-80 keV flux is $8.7^{+9.4}_{-5.6} \times 10^{34}$ erg/s, comparable to but slightly larger than that found for $Z = 0.018$. Thus, it is possible that the sensitivity of the Arches observations is comparable to that from Quintuplet, but given the larger uncertainties related to the stellar modeling of the Arches stars the limit is, at present, less robust. We stress that the qualitative difference between Arches and Quintuplet that is responsible for this difference is that Quintuplet has a large cohort of WC and WN stars, which are robustly He burning, while Arches does not have any stars in these stellar classes.

| | O | (B/R)SG/YHG | LBV | WNh | WC/WN | tot | tot (10-80 keV) |
|---|---|---|---|---|---|---|---|
| $N_{\rm star}$ | 96 | 0 | 0 | 13 | 0 | 109 | 109 |
| $z = 0.018$ $\mu_{\rm rot} = 100$ km/s | $2.3^{+0.2}_{-0.1} \times 10^{33}$ | 0 | 0 | $8.7^{+6.5}_{-5.2} \times 10^{34}$ | 0 | $8.9^{+6.5}_{-5.2} \times 10^{34}$ | $6.6^{+5.6}_{-3.6} \times 10^{34}$ |
| $z = 0.035$ $\mu_{\rm rot} = 100$ km/s | $3.9^{+1.8}_{-1.9} \times 10^{35}$ | 0 | 0 | $3.9^{+217}_{-0.6} \times 10^{32}$ | 0 | $7.2^{+16}_{-4.9} \times 10^{33}$ | $5.7^{+23}_{-2.8} \times 10^{33}$ |
| $z = 0.035$ $\mu_{\rm rot} = 150$ km/s | $3.5^{+2.1}_{-1.6} \times 10^{33}$ | 0 | 0 | $3.6^{+125}_{-0.3} \times 10^{32}$ | 0 | $4.7^{+12}_{-2.2} \times 10^{33}$ | $3.7^{+13}_{-2.4} \times 10^{33}$ |

Table G.3: As in Tab. G.1 but for Arches.

#### G.2.3.2 Data analysis, results, and systematic tests

We reduce and analyze 370 ks of archival NuSTAR data from Arches. The Arches observations (IDs `40010005001`, `40101001004`, `40101001002`, `40202001002`, `40010003001`) were performed as part of the same GC survey as the Quintuplet observations as well as for dedicated studies of the Arches cluster below 20 keV. Note that we discard data from the Focal Plane Module B instrument for observations `40101001004`, `40101001002`, `40202001002`, and `40010003001` because of ghost-ray contamination. We perform astrometric calibration using the low-energy data on the Arches cluster itself, which is a bright point source

above 3 keV.

In the Arches analysis it is known that there is a nearby molecular cloud that emits in hard $X$-rays [451]. We follow [451] and model emission associated with this extended cloud as a 2D Gaussian centered at R.A.=$17^h45^m50.62^s$, Dec.=$-28^\circ49'47.17''$ with a FWHM of $72.4''$. The hard $X$-ray spectrum associated with the molecular cloud has been observed to extend to approximately 40 keV [451]; indeed, we see that including the molecular cloud template, with a free normalization parameter, at energies below 40 keV affects the spectrum that we extract for the axion template, but it does not significantly affect the spectrum extraction above 40 keV. The non-thermal flux associated with the molecular cloud is expected to be well described by a power-law with spectral index $\Gamma \approx 1.6$ and may arise from the collision of cosmic-ray ions generated within the star cluster with gas in the nearby molecular cloud [452]. With this spectral index the molecular cloud should be a sub-dominant source of flux above $\sim$20 keV, and we thus exclude the 10-20 keV energy range from the Arches analysis, though *e.g.* including the 15-20 keV bin results in nearly identical results (as does excluding the 20 - 40 keV energy range).

The molecular cloud template is illustrated in the bottom left panel of Fig. G.17. In that figure we also show the data, background templates, signal template, and background-subtracted counts, as in Fig. G.9 for the Quintuplet analysis. Note that we profile over emission associated both the background template and with the halo template when constraining the flux in each energy bin associated with the signal template.

As a systematic test of our signal extraction procedure we show in Fig. G.18 (left panel) the spectrum extracted for axion emission from the Arches cluster both with and without the halo template. The two spectra diverge below $\sim$20 keV but give consistent results above this energy. Similarly, we find that the spectrum is relatively insensitive to the ROI size for energies above $\sim$20 keV, as shown in the right panel of Fig. G.18, which is analogous to the Quintuplet Fig. G.12.

In Fig. G.19 we show the 95% upper limit we obtain on $g_{a\gamma\gamma}$ from the Arches analysis, using the conservative modeling with $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s. We find no evidence for an axion-induced signal from this search. Note that, as in indicated in Fig. G.18, we do not include data below 20 keV in this analysis.

## G.3 Initial metallicity determination for Quintuplet and Arches

In our fiducial analysis we assumed the cluster metallicity was $Z = 0.035$, which we take as the highest allowed metallicity in the Quintuplet cluster. In this subsection we show how we arrived at this value. The cluster metallicity is an important parameter in that it affects the mass loss rates

in the stellar winds, the lifetime of individual evolutionary stages, and the surface abundances. Here we use measurements of the nitrogren abundances of WNh stars in the Arches cluster to estimate the uncertainty on the cluster metallicities. The nitrogen abundance during the WNh phase reaches a maximum that depends only on the original CNO content, and as such is a direct tracer of stellar metallicity (and increases with increasing metallicity). Ref. [329] measured the nitrogen abundance in the WNh stars in the Arches cluster at present to be $0.0157 \pm 0.0045$. We run MESA simulations of the Arches WNh stars on a grid of metallicities from $Z = 0.01$ to $Z = 0.04$ and find this measurement implies that the Arches initial metallicity is between $Z = 0.018$ and $Z = 0.035$. The results are shown in Fig. G.20, where we see that the nitrogen abundance during the WNh phase intersects with the measurement only for the initial metallicities in that range. Although there are no measurements of the Quintuplet WNh nitrogren abundance, note that a similar abundance was found in the nearby GC SSC of $0.0143 \pm 0.0042$ [453]. Given the similarity of these two measurements, we assume the same metallicity range for Quintuplet as computed for Arches.

## G.4 Variation of upper limits with initial conditions

In this section we show the variation in the upper limits as we vary over our initial conditions $Z \in (0.018, 0.035)$ and $\mu_{\rm rot} \in (50, 150)$ km/s. These initial conditions represent the dominant uncertainties in our stellar modeling. Recall that in our fiducial analysis we assume the initial metallicity and rotation giving the most conservative upper limits: $Z = 0.035$ and $\mu_{\rm rot} = 150$ km/s. Fig. G.21 shows, for both Quintuplet and Wd1, how our 95% upper limit varies as we scan over $Z$ and $\mu_{\rm rot}$. In particular, the shaded blue regions show the minimum and maximum limit obtained when varying $Z$ and $\mu_{\rm rot}$. Note that our fiducial limits, solid black, are the most conservative across most axion masses, though the effect of the $Z$ and $\mu_{\rm rot}$ is relatively minimal, especially for Wd1.

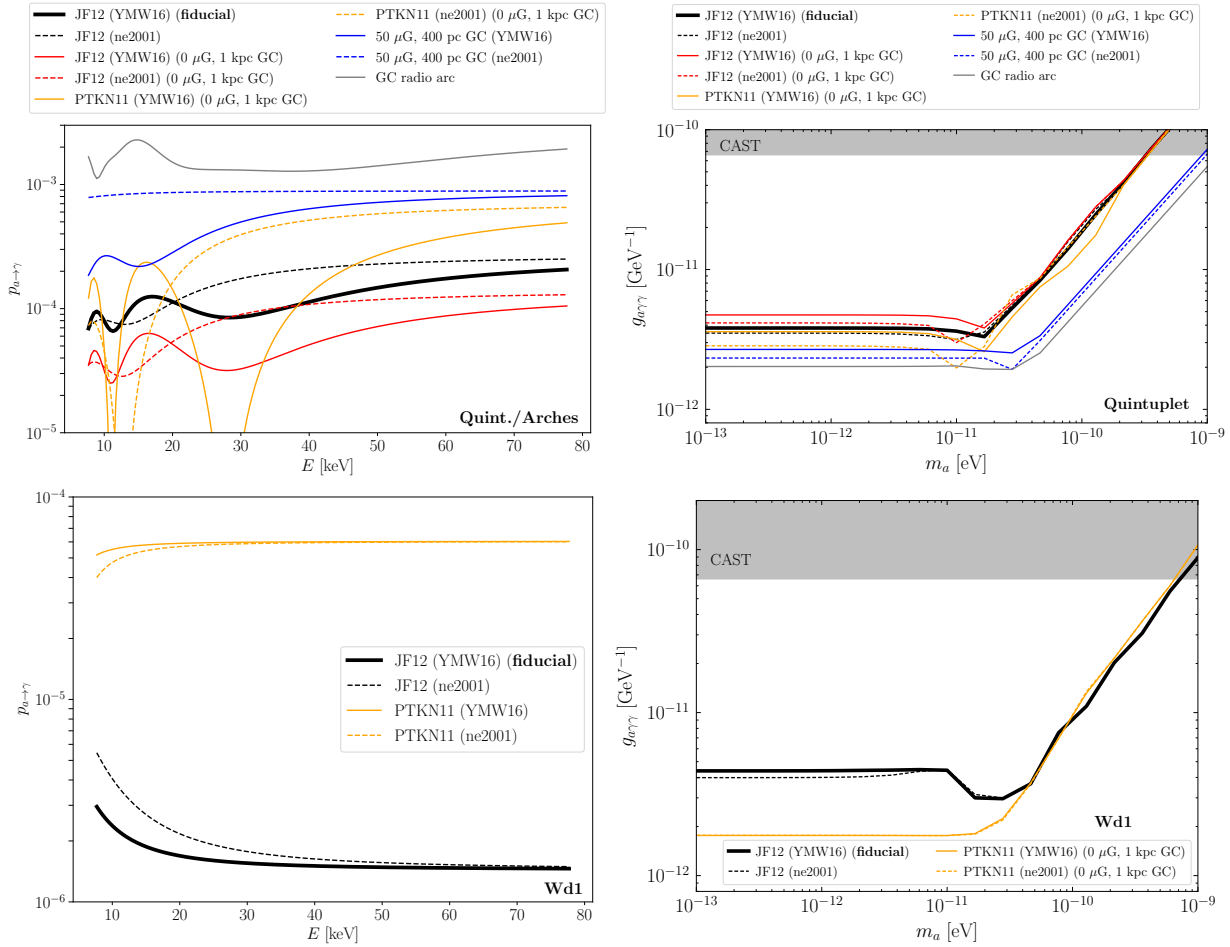



Figure G.8: (Left Column) The axion-photon conversion probabilities $p_{a\to\gamma}$, assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$, computed as a function of the axion energy $E$ (and assuming $m_a \ll 10^{-10}$ eV) using the formula given in (G.5). (Top Left) The conversion probabilities for axions produced in the Quintuplet or Arches clusters for different modeling assumptions for the Galactic magnetic field and free-electron density. Our fiducial result is shown in solid black. Note that the plasma mass, induced by the free-electron density, becomes more important at lower axion energies and induces the lower-energy features. The dashed black curve shows the effect of changing from the YMW16 free-electron model to the `ne2001` model. Removing the $B$-field within the inner kpc leads to the results in red, while only modeling a 50 $\mu$G field in the inner 400 pc leads to the results in blue. Changing to the PTKN11 model (and masking the inner kpc) gives the results in orange. We estimate that if the axions traverse the GC radio arc, located near the Quintuplet and Arches clusters, the conversion probabilities could be enhanced to the values shown in grey. (Bottom Left) As in the top left panel but for axions emitted from the Wd1 cluster. (Right Column) The effects of the different conversion probability models on the 95% upper limits on $g_{a\gamma\gamma}$ for Quintuplet (top right) and Wd1 (bottom right). Note that Arches is similar to Quintuplet, since they are both assumed to have the same conversion-probability models.

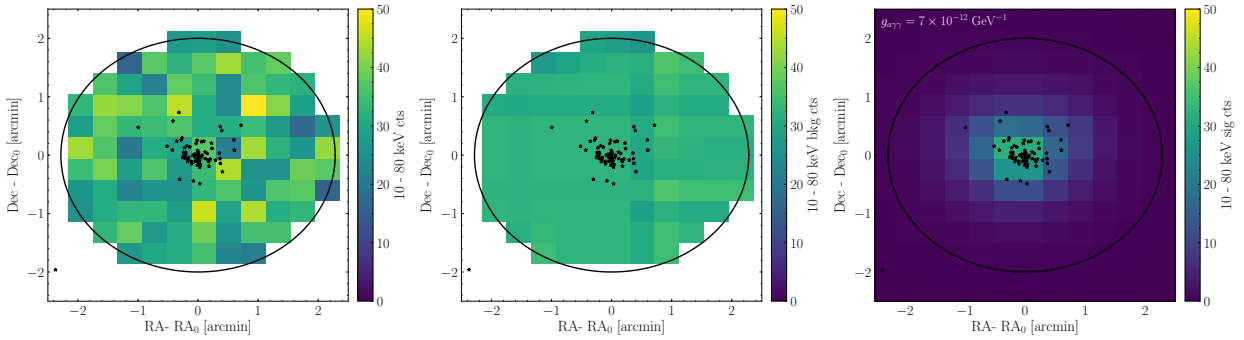



Figure G.9: (Left) As in Fig. 8.1, but for the total observed counts between 10 - 80 keV instead of the background-subtracted counts. (Center) The best-fit background model, summed from 10 - 80 keV, for the Quintuplet data set shown in the left panel. (Right) The predicted axion-induced signal template from Quintuplet, in counts, normalized for an axion with $g_{a\gamma\gamma} = 7 \times 10^{-12}$ GeV$^{-1}$ and $m_a \ll 10^{-11}$ eV.

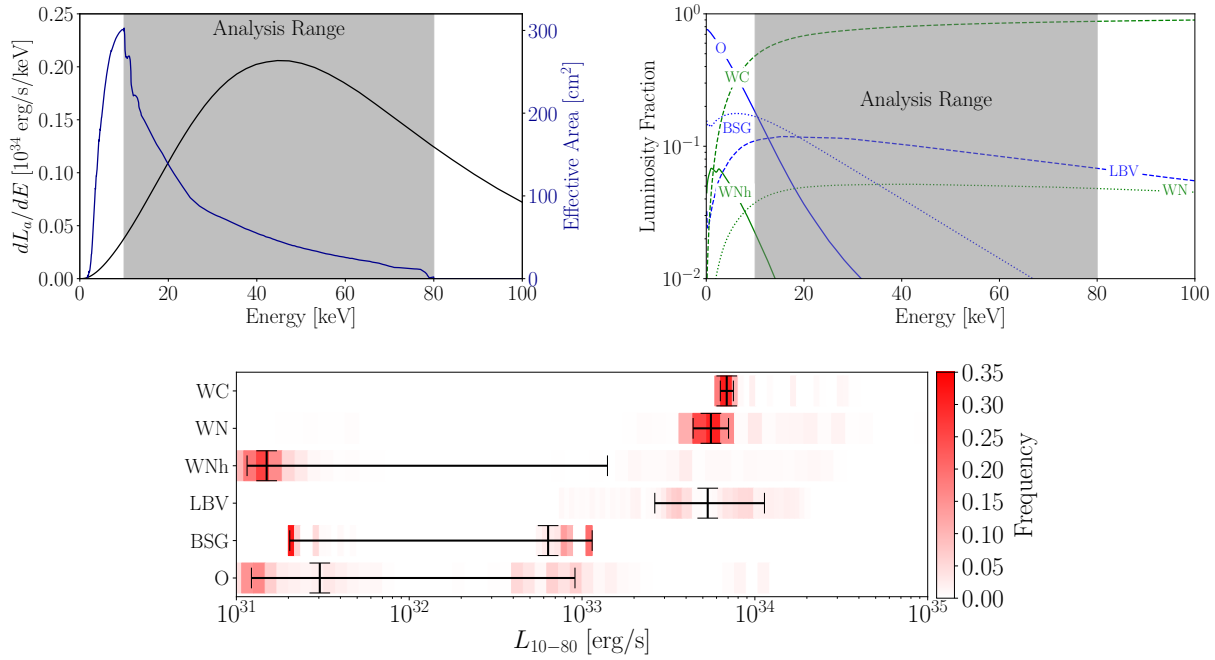



Figure G.10: (Upper Left) The Quintuplet axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 10 - 80 keV, is shaded in red. (Upper Right) The individual contributions of each stellar classification to the Quintuplet axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification (per star) in Quintuplet. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band.

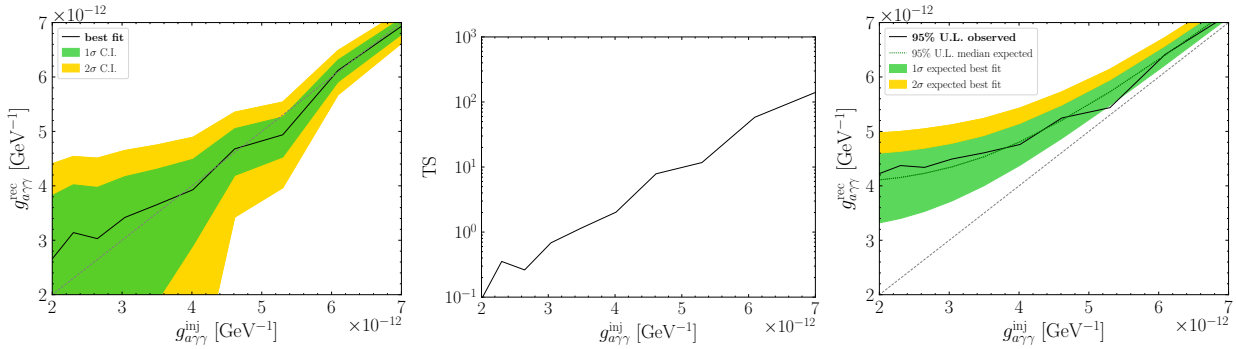



Figure G.11: (Left) We inject a synthetic axion signal into the Quintuplet NuSTAR data with axion coupling $g^{\rm inj}_{a\gamma\gamma}$. We then pass the hybrid synthetic plus real data through our analysis pipeline and show the best-fit coupling $g^{\rm rec}_{a\gamma\gamma}$, along with the recovered $1\sigma$ and $2\sigma$ uncertainties. (Middle) The discovery TS for the axion signal for the test illustrated in the left panel. The square root of the TS is approximately the discovery significance. (Right) The 95% upper limit recovered for the injected signal test. Importantly, the 95% upper limit is above the injected signal value, for all injected signal strengths, and the upper limit is consistent with the 68% and 95% expectations for the upper limit under the null hypothesis, which are indicated in green and gold, respectively.

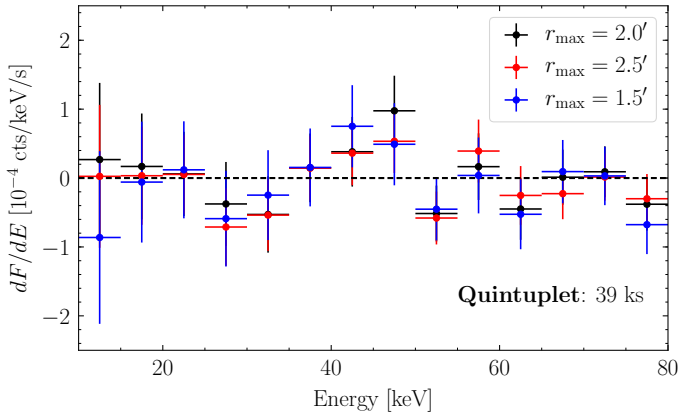



Figure G.12: As in Fig. 8.2, except for different ROI sizes, as indicated.

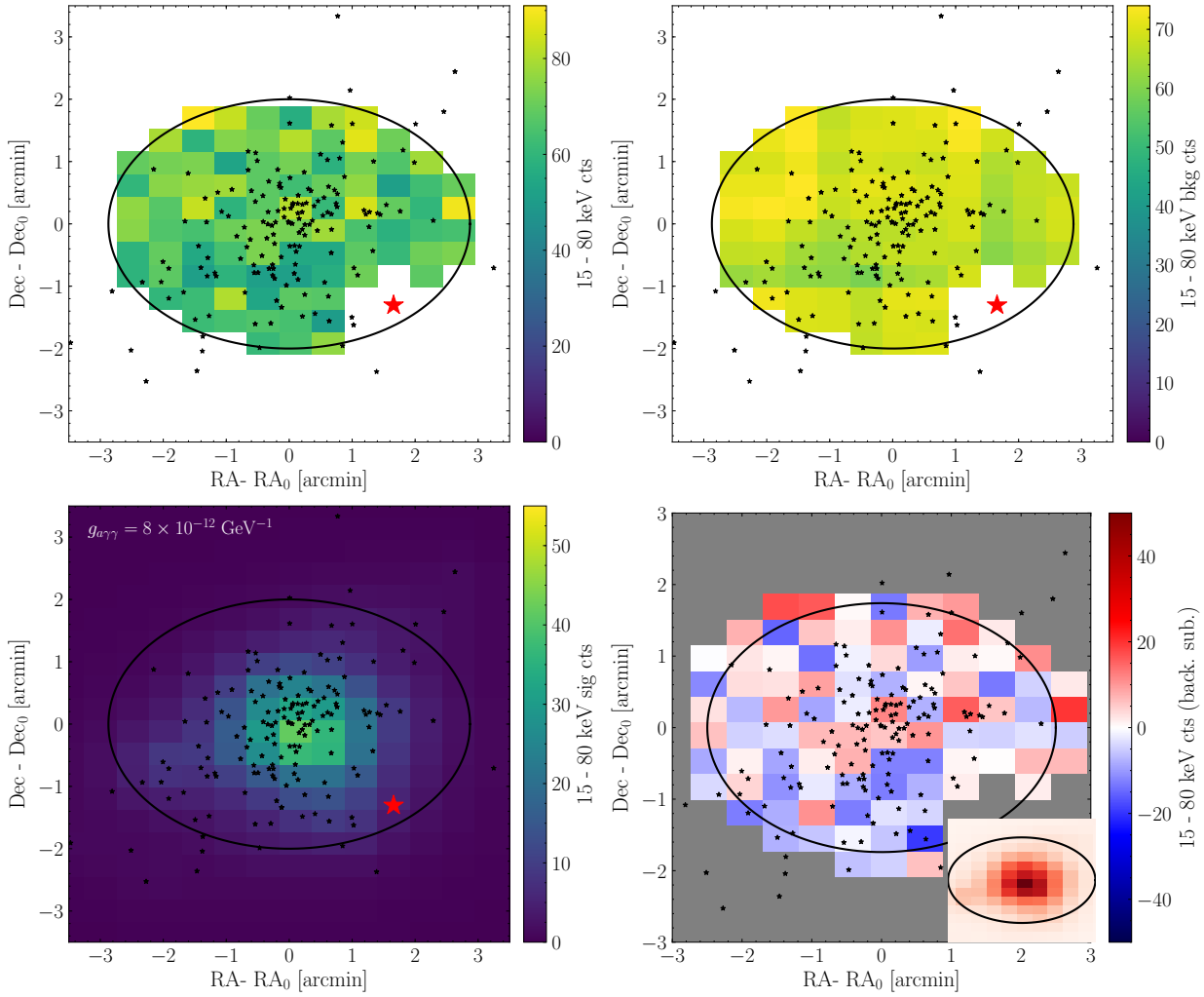



Figure G.13: As in Fig. G.9, but for the Wd1 cluster NuSTAR analysis. The red star indicates the location of the magnetar CXOU J164710.2–45521, which is masked at 0.5'. Also shown is the background-subtracted count data, as in Fig. 8.1.

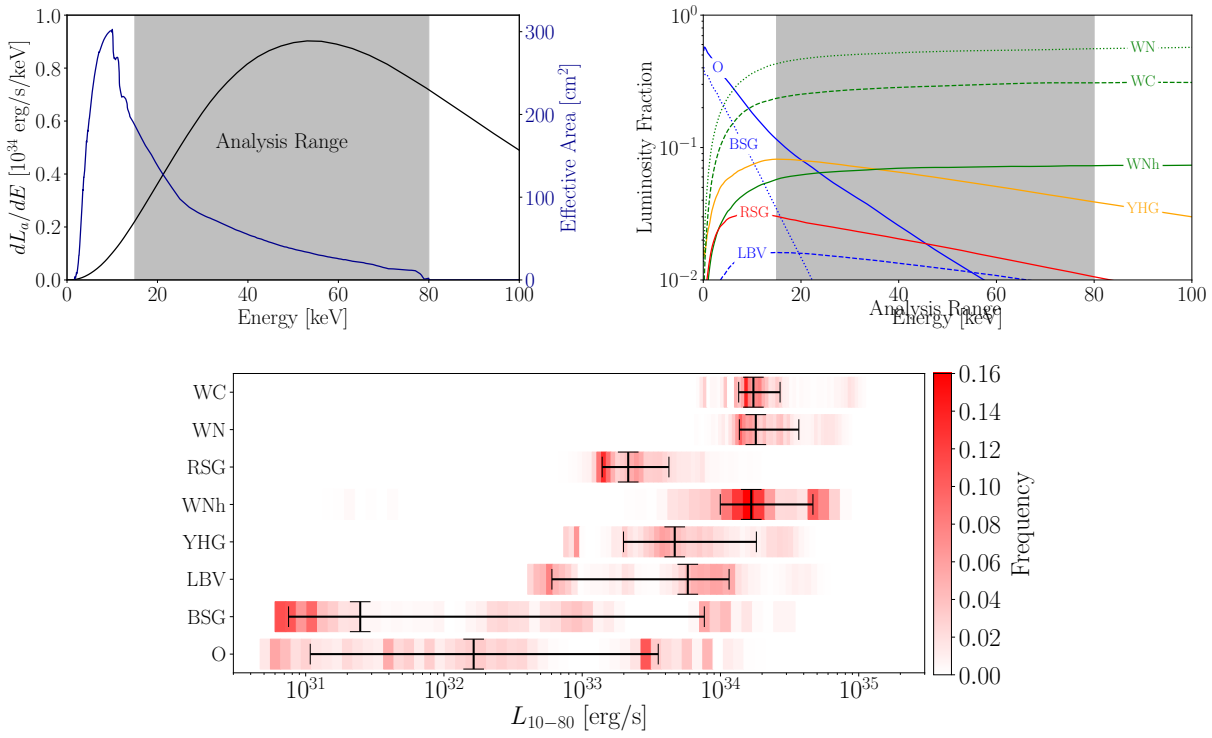



Figure G.14: (Upper Left) The Wd1 axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 15 - 80 keV, is shaded in gray. (Upper Right) The individual contributions of each stellar classification to the Wd1 axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification in Wd1. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band.

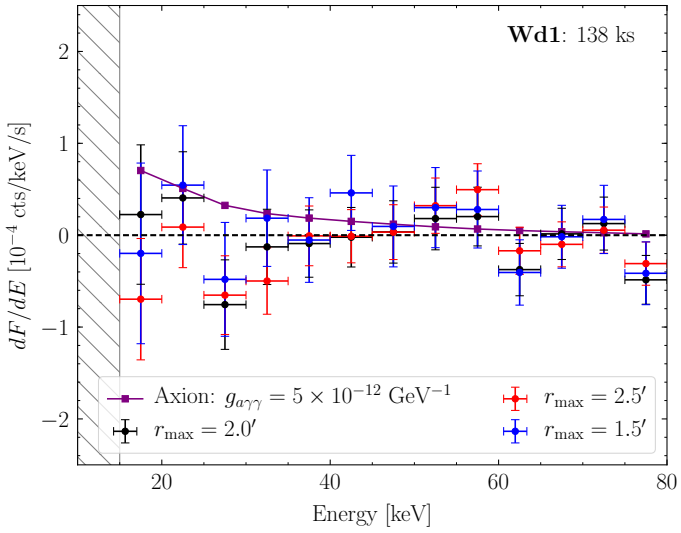



Figure G.15: As in Fig. G.12 but for the Wd1 analysis. Note that we only include energies above 15 keV in our analysis because of ghost-ray contamination.

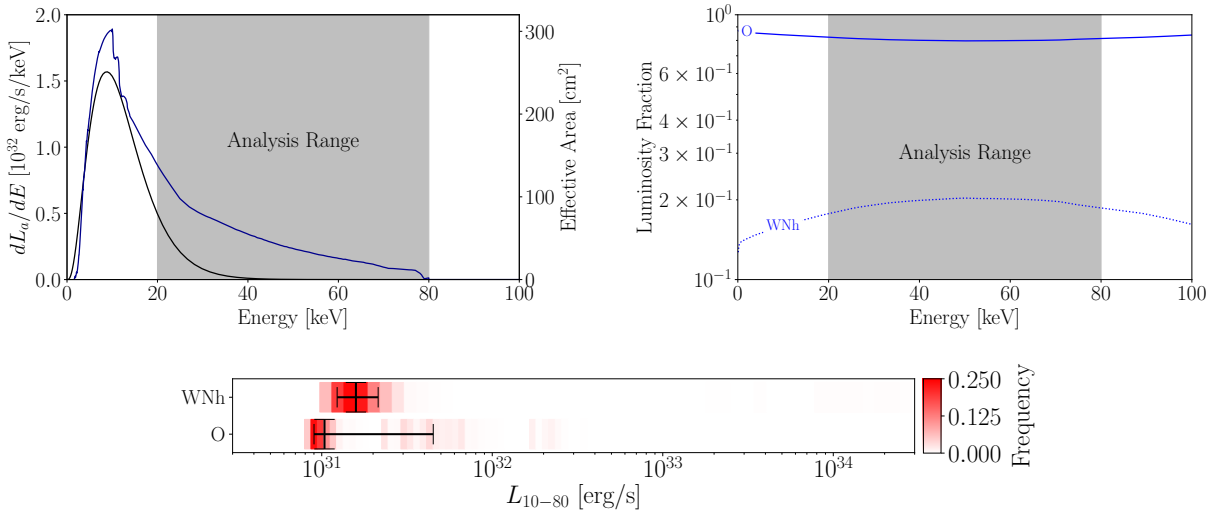



Figure G.16: (Upper Left) The Arches axion spectrum assuming $g_{a\gamma\gamma} = 10^{-12}$ GeV$^{-1}$ (black) plotted against the NuSTAR effective area (blue). The analysis range, from 20 - 80 keV, is shaded in gray. (Upper Right) The individual contributions of each stellar classification to the Arches axion spectrum. The analysis range is again shaded. (Bottom) The 10-80 keV luminosity distribution assigned to each stellar classification in Arches. In red we show the frequency with which each luminosity occurs, while the black error bars show the mean and $1\sigma$ band.

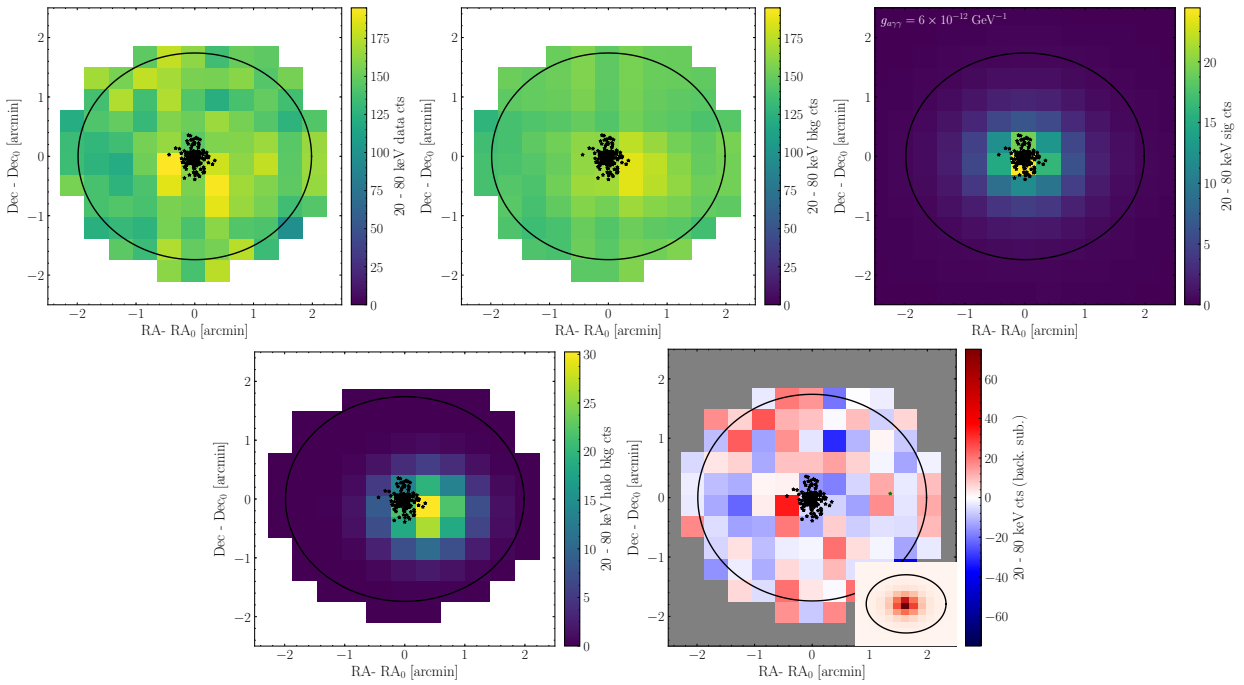



Figure G.17: (Top Panel) As in Fig. G.9, but for the Arches cluster. (Bottom left) We show the best-fit emission associated with the halo template that describes emission from the nearby molecular cloud. (Bottom right) As in in Fig. 8.1, but for Arches.

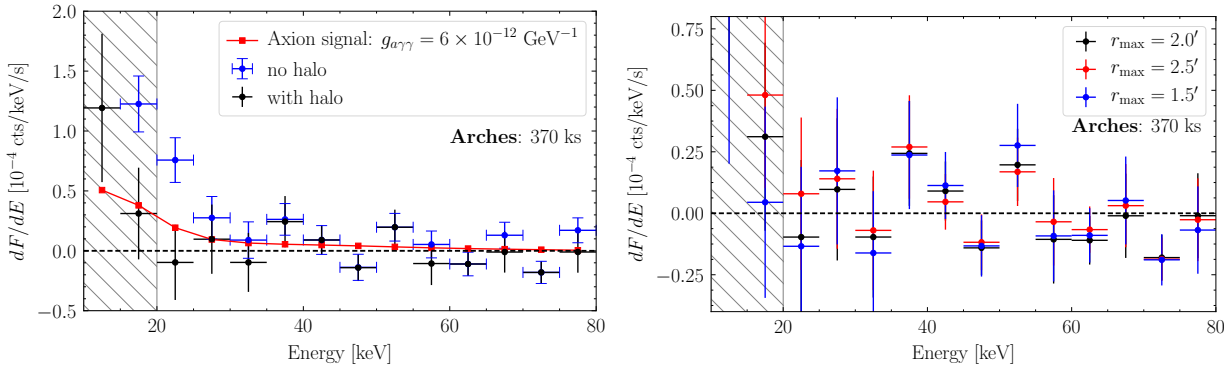



Figure G.18: (Left) The Arches spectrum measured with and without the halo template. Note that we use the spectrum with the halo template in our fiducial analysis, though the difference between the two results is relatively minor above ∼20 keV. (Right) As in Fig. G.12 but for the Quintuplet analysis. Note that these spectra are computed while profiling over halo emission. Above ∼20 keV the different ROIs produce consistent results.

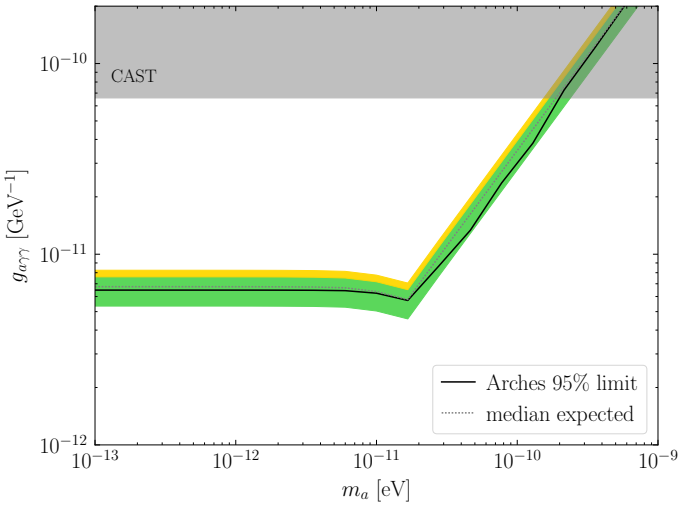



Figure G.19: As in Fig. 8.3 but from the analysis towards the Arches SSC. No evidence for axions is found from this search.

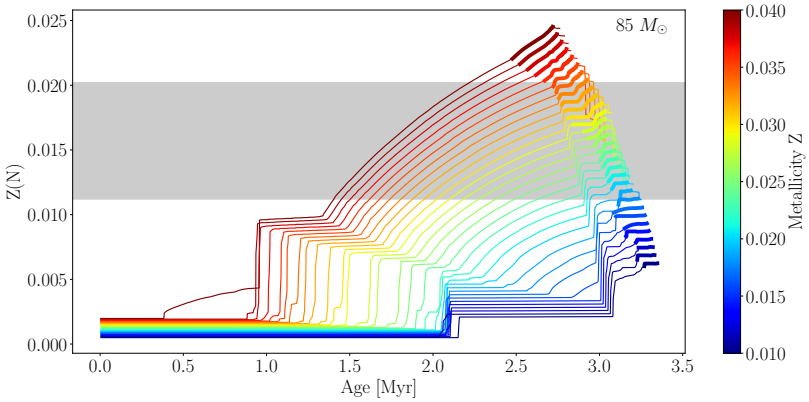



Figure G.20: (Left) The evolution of the nitrogen abundance Z(N) over time from MESA simulations of a non-rotating 85 $M_\odot$ star with initial metallicity $Z = 0.01$ to $Z = 0.04$. The bolded sections of the lines correspond to the WNh phase. The gray shaded region indicates the measurements of nitrogren abundances of the Arches WNh stars from [329].

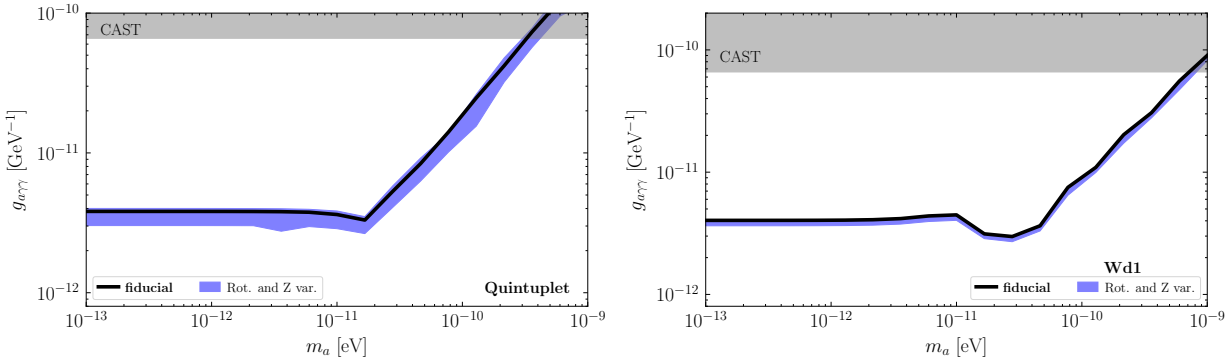



Figure G.21: (Left) The variation to the 95% upper limit found by varying the initial metallicity and rotation in the range $Z \in (0.018, 0.035)$ and $\mu_{\rm rot} \in (50, 150)$ km/s for the Quintuplet analysis. The blue region indicates the maximum and minimum limit found, while the black curve shows our fiducial limit. (Right) As in the left panel but for Wd1.

# APPENDIX H

# X-Ray Searches for Decaying Dark Matter in the Milky Way

## H.1 Data Reduction and Analysis

In this section, we detail our process for data reduction and analysis.

### H.1.1 Data Reduction

We selected all *XMM-Newton* observations performed until September 5, 2018. For each of these observations, we retrieved the raw data products from the XMM-Newton Science Archive. For data reduction, we used the *XMM-Newton* Extended Source Analysis Software (ESAS) package, which is a part of the Science Analysis System [454] (SAS) version 17.0, and used for modeling sources covering the full *XMM-Newton* field-of-view and diffuse backgrounds.

The data reduction process is described in detail in Ref. [378]; here, we summarize the important steps and point out any differences. To reduce a given observation, we obtain the list of science exposures (i.e. pointings taken in a mode usable for scientific purposes) from the summary files. For each exposure (independent of camera), we generate an event list and filter this list to only include events which were recorded during a period of low-background, which cuts contamination from soft-proton flares. We then mask point sources in the field of view which contribute in any energy range (c.f. Ref. [378] where we only masked point sources in the 3-4 keV range). We also mask data from CCDs operating in anomalous states. From the filtered and masked data products we create the photon-count data, the ancillary response file (ARF), and the redistribution matrix file (RMF).

The reduced data contains 11,805 observations, with 21,388 and 8,190 individual MOS and PN exposures, totaling 438 Ms and 109 Ms of data. Given our focus is on searching for DM emission in otherwise dark regions of the sky, we place a cut on these data sets to isolate the astrophysically quietest amongst them. In particular, we construct the integrated flux from $2 - 10$ keV in all

exposures, and determine the median value for MOS and PN separately as 0.09 photons/keV/s and 0.39 photons/keV/s. All observations with integrated fluxes higher than these median values are excluded. This cut will remove observations with above average astrophysical emission, but also those where there is large instrumental or quiescent particle background (QPB) emission (c.f. Ref. [378] where a separate cut on the QPB emission was performed). For regions of the sky that are not focused on a bright Galactic or extra-galactic source, the QPB counts should dominate over the extra-galactic X-ray background [455]. However, the QPB is time-dependent and will vary over exposures because of *e.g.* flaring activity (which, as described in the SM, we filter for). Further, we emphasize that even in the most optimistic scenario, a DM UXL will only provide an exceptionally small contribution to the total integrated flux, and thus this cut will not bias against a potential signal. In addition to removing these bright exposures, we place two additional cuts. Firstly, all exposures with less than 500s of data are removed, as the flux in such short exposures can be poorly characterized. Finally, we exclude all observations within $2^\circ$ of the plane of the Milky Way, which excises only a small amount of the expected DM signal, but a much larger fraction of the expected astrophysical emission associated with emission from our own galaxy. The cuts leave 215 Ms (57 Ms) of the total 438 Ms (109 Ms) of full-sky ($|b| \geq 2^\circ$) exposure time for MOS (PN).

Exposure passing all three cuts are then divided into 30 rings, each of width $6^\circ$ from the GC as described in the main body. The rings, numbered 1-30 starting from the GC, are used to form our signal ROI (rings 1-8) and background ROI (rings 20-30).

### H.1.2 Public Data Products

The processed data used to perform the analysis in this work is made fully publicly available at github.com/bsafdi/XMM_BSO_DATA. There we provide all the data required to reproduce our results. In particular, we provide the data after the cuts described above in each ring for the MOS and PN cameras separately. The instrument response files, appropriately weighted across the exposures in each ring, are provided.

### H.1.3 Analysis

In this section we provide additional details behind the analysis framework used to interpret the data products described above in the context of the decaying DM model. First, we describe how we analyze the flux data in the individual annuli, and then we detail how those results are joined together to constrain the DM lifetime. Lastly, we describe how we test for and incorporate systematic uncertainties.

#### H.1.3.1 Construction of the profile likelihood

Let us first focus on the analysis of the (either MOS or PN) data in an individual ring $k \in [1, 8]$. The data set $d^k$ in this ring consists of background subtracted count rates $d_i^k$ in each energy channel $i$. The count rates have units of cts/s/keV, as illustrated in *e.g.* Fig. H.1, with Poisson counting uncertainties $\sigma_i^k$ that arise from combining the statistical uncertainties in the signal and background data sets in the large-count limit, where the uncertainties become normally distributed. Our goal is then to compute the log-likelihood $\log p(d_k|\theta)$ as a function of the model parameters $\theta = \{A_{\rm sig}, \theta_{\rm nuis}\}$, which consist of our parameter of interest, $A_{\rm sig}$, and our nuisance parameters $\theta_{\rm nuis}$. The nuisance parameters include background line amplitudes, $A_j$, with $j$ indexing the different lines at energies $E_j$, and also hyperparameters for the GP model. In our fiducial analysis the only GP model hyperparameter is the amplitude of the double-exponential kernel $A_{\rm GP}$. Note that, as described shortly, in determining the instrumental line list we also assign nuisance parameters to the locations of the lines. Our goal is to construct the profile likelihood $\log p(d_k|A_{\rm sig}) = \max_{\boldsymbol{\theta}_{\rm nuis}} \log p(d_k|\boldsymbol{\theta})$.

Before describing the log-likelihood function in detail, we note that because we are in the large-count limit, so that the statistical fluctuations are normally distributed, we may interchangeably use the concept of modeling the data as the sum of model components and subtracting model components from the data and considering the residuals. In the small-count limit, where the Poisson fluctuations are not nearly Gaussian, this approach would not be appropriate.

The log-likelihood function that we use is a modification of the zero-mean GP marginal likelihood. The modification that we implement incorporates the background lines and the signal line of interest. For a given set of model parameters $\{A_{\rm sig}, A_j\}$ we construct the modified data vector[1]

$$y_i^k(\boldsymbol{\theta}) \equiv d_i^k - A_{\rm sig}\mu_{{\rm sig},i}^k - \sum_j A_j \mu_{j,i}^k - \left\langle d_i^k - A_{\rm sig}\mu_{{\rm sig},i}^k - \sum_j A_j \mu_{j,i}^k \right\rangle_i , \tag{H.1}$$

where $\mu_{\rm sig}$ is the spectral template of the signal line of interest, with fixed normalization, as obtained by appropriately summing the forward modeling matrices of the individual exposures that compose the observations within the ring of interest, $k$. Similarly, $\mu_{j,i}^k$ denotes the fixed-normalization spectral template of the $j^{\rm th}$ background line, at energy $E_j$, in ring $k$ (recall that $i$ labels the detector energy channel). The quantity $\langle \cdots \rangle_i$ in (H.1) denotes the average over energy bins $i$, which implies that by construction the $y_i^k(\boldsymbol{\theta})$ have zero mean when averaged over the full energy range of the analysis. We postulate that the $y_i^k$ are described by GP models, so that we may

[1]Note that the line energies $E_j$ are fixed in all analyses except those of the background ROI data for constructing our lists of instrumental lines; in those analyses only, the $E_j$ are also model parameters.

use the GP marginal likelihood to compute the hyperparameter $A_{\rm GP}$:

$$\log p(d_k|\boldsymbol{\theta}) = -\frac{1}{2}\mathbf{y}^{k^T}\left[\mathbf{K} + (\sigma^k)^2\mathbf{I}\right]^{-1}\mathbf{y}^k - \frac{1}{2}\log|\mathbf{K} + (\sigma^k)^2\mathbf{I}| - \frac{n}{2}\log(2\pi)\,. \tag{H.2}$$

Above, $n$ is the number of energy channels, and all matrix operations are taken in the space of energy channels, with $(\sigma^k)^2\mathbf{I}$ denoting the diagonal matrix with entries $(\sigma_i^k)^2$. The matrix $\mathbf{K}$ denotes the GP kernel. We implement the non-stationary kernel

$$K(E, E') = A_{\rm GP}\exp\left[-\frac{(E-E')^2}{2EE'\sigma_E^2}\right], \tag{H.3}$$

which has the hyperparameters $A_{\rm GP}$ and $\sigma_E$. Note that later in the appendices we show that similar results are obtained using the more standard double exponential kernel, but we chose the form of the kernel in (H.3) for reasons discussed below.

It is worth emphasizing that we have made the choice to describe the residuals of the background-subtracted data, after also subtracting the contributions from the instrumental lines, by a zero-mean GP model. An alternative strategy would be to allow the GP model to have an energy-dependent mean. Equivalently, we could include a parametric model component (such as a power-law or exponential) to model the clear upward trend in the data at low energies observed in *e.g.* Fig. H.1, with the GP model then describing fluctuations about that parametric component. Such an approach would likely result in smaller values of the hyperparameter $A_{\rm GP}$ and, potentially, increased sensitivity. Such a hybrid parametric plus GP modeling approach could be explored in future work.

Our goal is to look for narrow lines on top of a smooth continuum flux. We know that even a narrow line will manifest itself as a broader feature in the detector-level data due to the detector response. So the correlation-length of the GP kernel has a lower-bound set by the detector resolution. Because the energy resolution $\delta E$ of *XMM-Newton* increases linearly with energy (*i.e.*, $\delta E/E$ is roughly constant), a stationary kernel with a fixed correlation length is not adequate and a kernel of the form in (H.3) is more natural. However, we expect the continuum to be much smoother, even before the smearing induced by the detector resolution. A common approach in GP literature when the hyperparameters are not motivated from some other considerations is to fit them to the data. This approach leads to the best-fit values $\sigma_E \approx 0.608$ ($\sigma_E \approx 0.77$) for MOS (PN) in the first annulus, with comparable results in the annuli further from the Galactic Center. However, we chose to fix $\sigma_E = 0.3$ because this is an intermediate value between the lower-bound of a narrow line given by the energy resolution and the best-fit result reflecting the smoothness of the observed continuum. This choice leads to more conservative limits, since for smaller values of $\sigma_E$ the GP model is able to capture smaller-scale fluctuations in the data, absorbing what would otherwise be

attributed to narrow lines.

Lastly, note that while the marginal likelihood in (H.2) is defined within the context of Bayesian statistics, as it is obtained by integrating the likelihood times prior distribution for the formal GP model parameters, we will use the likelihood to perform frequentist parameter inference. This approach is called the "hybrid approach" in [404]. As noted in [404], the asymptotic expectations for the distribution of the TS constructed from the marginal likelihood may differ from the frequentist expectations [172], because of the use of the Bayesian marginal likelihood, and so in principle the $p$-values and upper-limit criteria should be calibrated on Monte Carlo (MC). However, as we show below, we find through MC simulations that in our examples the TS statistical distributions follow the asymptotic frequentist expectations to high accuracy. With that in mind, we briefly review the asymptotic expectations for translating discovery TS values to $p$-values and forming 95% one-sided upper limits.

As discussed in the main Letter, the TS in favor of the signal model is given by

$$t = -2\big[\max_{\boldsymbol{\theta}} \log p(d_k|\boldsymbol{\theta}) - \max_{\theta_{\rm nuis}} \log p(d_k|\{A_{\rm sig}=0, \boldsymbol{\theta}_{\rm nuis}\}\big]\,, \tag{H.4}$$

where the second term is the maximum marginal likelihood for the null model without a signal line. When searching for evidence of DM, the discovery TS is set to zero for unphysical model parameters (in that case, $\sin^2(2\theta) < 0$), but for the purpose of testing for systematic uncertainties it is useful to allow for both positive and negative signal amplitudes. The discovery TS is asymptotically $\chi^2$ distributed with one degree of freedom under the null hypothesis (see, *e.g.*, [172]). In addition to searching for evidence of the signal model over the null hypothesis using $t$, we also set 95% one-sided upper limits using the likelihood ratio. We define the profile likelihood ratio $q(A_{\rm sig})$ by

$$q(A_{\rm sig}) = -2\big[\max_{\boldsymbol{\theta}_{\rm nuis}} \log p(d_k|\{A_{\rm sig}, \boldsymbol{\theta}_{\rm nuis}\}) - \max_{\boldsymbol{\theta}} \log p(d_k|\boldsymbol{\theta})\big]\,, \tag{H.5}$$

where in the first term we maximize the log-likelihood over the nuisance parameters $\boldsymbol{\theta}_{\rm nuis}$ at fixed signal parameter $A_{\rm sig}$. Let $\hat{A}_{\rm sig}$ be the best-fit signal parameter; *i.e.*, $q(\hat{A}_{\rm sig}) = 0$. Then, the 95% one-sided upper limit $A_{\rm sig}^{95\%}$ is given by the value $A_{\rm sig}^{95\%} > \hat{A}_{\rm sig}$ which satisfies $q(A_{\rm sig}^{95\%}) \approx -2.71$ [172].

Note that the profile likelihood in (H.5) is computed as a function of the signal normalization $A_{\rm sig}$ at fixed UXL energy (or, equivalently, fixed DM mass). All of the analyses presented in the main Letter are performed in this way (*i.e.*, we have a grid of different UXL energies to probe and then for each fixed energy we compute the profile likelihood as a function of the signal-strength amplitude). In Sec. H.3.1, however, we consider our ability to localize a putative signal in $m_a$ and

$\sin^2(2\theta)$ using synthetic data. In that analysis, and that analysis only, we simultaneously constrain the mass and the signal strength.

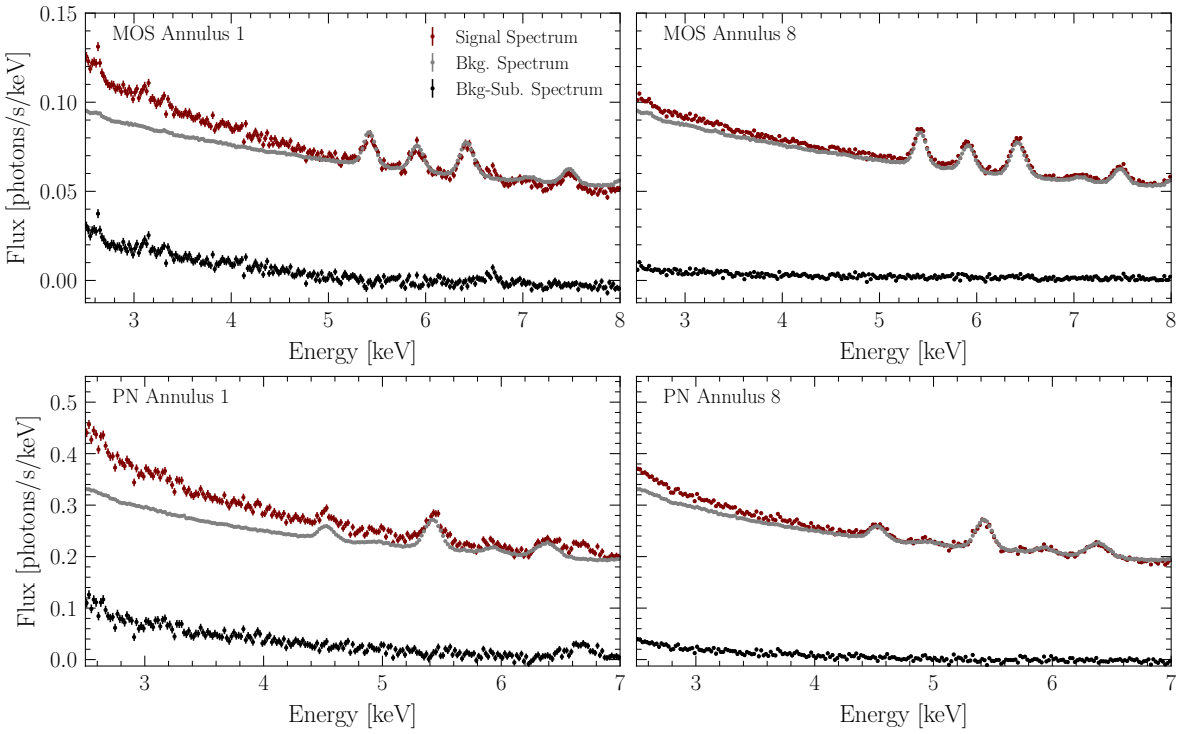



Figure H.1: Examples of the signal region spectra for MOS (top panels) and PN (bottom panels) in Ring 1 (left panels) and Ring 8 (right panels) with and without background subtraction in red and black, respectively. The background-region spectra are shown in grey. Many of the large instrumental features that are removed when looking at the background-subtracted data. Note that for visual clarity these spectra have been down-binned by a factor of 4.

### H.1.3.2 Instrumental and astrophysical background lines

Several instrumental and astrophysical background spectral lines are expected to contribute to the observed flux spectra. Here, we outline the procedure by which we obtain our candidate instrumental and astrophysical background lines, with the complete list of included lines for MOS and PN presented in Tab. H.1 and Tab. H.2, respectively.

We adopt an initial instrumental line list for PN and MOS from Refs. [456, 457]. We then analyze the stacked data, for MOS and PN independently, in the background ROI in order to test for the presence of each candidate line. We use an analysis framework analogous to that we use in the background-subtracted signal ROI data: in particular, our analysis of the background ROI data incorporates GP modeling for the continuum emission, in addition to the set of putative

| Energy [keV] | Origin | Type | Ring 1 | Ring 2 | Ring 3 | Ring 4 | Ring 5 | Ring 6 | Ring 7 | Ring 8 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2.46 | S | Astro. | **12.5** | 0.7 | **7.2** | 0.7 | 2.0 | **5.5** | 0.9 | **5.4** |
| 2.62 | S | Astro. | **36.8** | **7.6** | **4.3** | 1.9 | 0.0 | 2.1 | 0.0 | **6.1** |
| 3.12 | Ar | Astro. | **15.0** | 1.1 | **8.8** | 2.8 | 0.0 | **3.5** | 0.0 | 0.9 |
| 3.90 | Ca | Astro. | 0.3 | 0.0 | 0.0 | **9.8** | 0.0 | 0.2 | **4.1** | 0.0 |
| 5.42 | Cr | Inst. | **8.9** | **7.7** | **22.1** | 0.1 | 0.0 | 1.2 | **7.2** | 0.0 |
| 5.92 | Mn | Inst. | 1.8 | 1.0 | **3.6** | 0.3 | 0.4 | 2.5 | **10.0** | 1.7 |
| 6.42 | Fe | Inst. | 0.1 | **7.1** | **82.0** | 0.3 | 1.5 | 0.0 | 2.7 | **7.2** |
| 6.67 | Fe | Astro. | **55.4** | 1.7 | 0.0 | **4.3** | **5.9** | 2.5 | 0.0 | 0.0 |
| 6.97 | Fe | Astro. | **5.9** | 0.0 | 0.5 | **4.3** | 0.3 | 0.2 | 0.4 | 0.2 |
| 7.08 | Fe | Inst. | 1.5 | 0.0 | **4.1** | 0.2 | 2.1 | 1.3 | 0.0 | 2.2 |
| 7.48 | Ni | Inst. | 2.0 | 0.1 | **11.9** | 0.1 | 0.0 | 0.1 | **3.1** | **4.7** |
| 8.06 | Cu | Inst. | 1.1 | 0.9 | 0.3 | 0.1 | 0.0 | 0.2 | 0.4 | 0.1 |

Table H.1: The list of spectral lines of instrumental and astrophysical origins which are included in our background model for the MOS camera. For the line in each ring, we provide the value of $\Delta\chi^2$ associated with the addition/removal of the line from the best-fit background model which is obtained after our line-dropping procedure. Bolded values indicate the inclusion of a line in a ring's background model.

instrumental lines. We test for known instrumental lines in the vicinity of: 4.51 keV (Ti $K\alpha$), 5.41 keV (Cr K$\alpha$), 5.90 keV (Mn K$\alpha$), 5.95 keV (Cr K$\beta$), 6.40 keV (Fe K$\alpha$), 6.49 keV (Mn K$\beta$), 7.06 keV (Fe K$\beta$), 7.47 keV (Ni K$\alpha$), 8.04 keV (Cu K$\alpha$). During this process, we allow the central location of the background lines to float by up to $25$ eV. Lines which are detected with $t > 16$ ($4\sigma$ local significance) in the background data analysis are accepted at their best-fit energy as a new component of our residual background model. In MOS, we accept instrumental lines at energies: 5.42 keV, 5.915 keV, 6.425 keV, 7.07 keV, 7.485 keV, and 8.06 keV. In PN, we accept instrumental lines at 4.52 keV , 5.42 keV, 5.95 keV, and 6.39 keV.

After constructing our list of instrumental background lines we include them in our analyses of the signal-ROI background-subtracted data sets. In particular, each line is given an intensity nuisance parameter in each ring. Given our procedure of subtracting the background flux from the signal region, variability in the instrumental lines between observations can result in the best fit instrumental line intensity in our data set having a positive or negative normalization. Accordingly, we allow the normalization of the instrumental lines to be either positive or negative.

We also develop an initial list of candidate astrophysical background lines following [379], by selecting those with emissivities greater than $5 \times 10^{-19}$ photons/cm$^3$/s at a temperature of 1 keV, which is the approximate temperature of the hot component of the Galactic Center, using the

| Energy [keV] | Origin | Type | Ring 1 | Ring 2 | Ring 3 | Ring 4 | Ring 5 | Ring 6 | Ring 7 | Ring 8 |
|---|---|---|---|---|---|---|---|---|---|---|
| 2.46 | S | Astro. | **15.3** | 0.7 | **12.5** | **3.9** | 0.0 | **3.5** | 1.1 | **6.8** |
| 2.62 | S | Astro. | **19.0** | **4.5** | **9.1** | **5.1** | 0.0 | 0.0 | 0.1 | **4.4** |
| 3.12 | Ar | Astro. | **6.4** | 2.8 | **6.5** | **13.7** | 0.0 | 1.1 | 0.0 | 0.0 |
| 3.90 | Ca | Astro. | **3.9** | 0.0 | 0.5 | **3.8** | 0.0 | **6.0** | 0.2 | 0.5 |
| 4.52 | Ti | Inst. | 0.6 | 0.7 | 0.8 | 0.6 | 0.1 | 0.3 | 0.5 | 0.0 |
| 5.42 | Cr | Inst. | 0.7 | 0.0 | 0.4 | **7.3** | 0.7 | 1.9 | 2.4 | **6.6** |
| 5.93 | Cr | Inst. | 0.8 | 1.0 | 0.5 | **4.7** | 0.1 | 1.7 | **3.6** | 0.7 |
| 6.39 | Fe | Inst. | 0.0 | 0.1 | **3.2** | 1.2 | 0.0 | 1.6 | **9.3** | 0.0 |
| 6.67 | Fe | Astro. | **79.2** | **5.5** | **8.9** | 2.8 | 0.8 | 2.3 | 0.3 | 2.3 |
| 6.97 | Fe | Astro. | 0.0 | 0.0 | 0.4 | 0.0 | 0.3 | 0.0 | 0.7 | 0.0 |

Table H.2: The list of spectral lines of instrumental and astrophysical origins which are included in our background model for the PN camera. For the line in each ring, we provide the value of $\Delta\chi^2$ associated with the addition/removal of the line from the best-fit background model which is obtained after our line-dropping procedure. Bolded values indicate the inclusion of a line in a ring's background model.

AtomDB database [458]. We include additional iron lines that are known to produce emission in the inner Galaxy [459]. Taking this preliminary list, we then inspect the innermost ring and determine all lines which appear with TS $t > 3$ in either PN or MOS. If such a line meets this criteria in either PN or MOS then we add it to our list of putative astrophysical lines for both instruments. As with their instrumental counterparts, the astrophysical lines are treated with independent nuisance parameters describing their intensity in each annuli. However, for astrophysical background lines, we restrict intensities to values greater than or equal to zero.

The procedure described above leads to a list of astrophysical and instrumental lines, which are shown in Tabs. H.1 and H.2. However, this does not mean that we included all of the those lines in every ring when performing our UXL search. In each annulus we analyze the background-subtracted data to determine which of the lines in Tabs. H.1 and H.2 are detected with moderate significance (we use the criteria $t > 3$) in the background-subtracted data set. Note that in Tabs. H.1 and H.2 we indicate whether the line is included in each annulus. To determine the significance of a given line we proceed iteratively, starting with the full list of lines and then calculating the change in the maximum likelihood when a given line is removed from the model.

In Fig. H.2 we illustrate example fits for our fiducial analyses to the data without the inclusion of an UXL. These fits are to the same background-subtracted data as illustrated in Fig. H.1, as labeled. In black we show the combined best-fit model, which is the sum of the GP model contribution (dark red) and the contributions from the individual astrophysical and instrumental lines (colored

curves). Note that the number of background lines differs between each of the annuli because the important background lines are determined independently for each annulus.

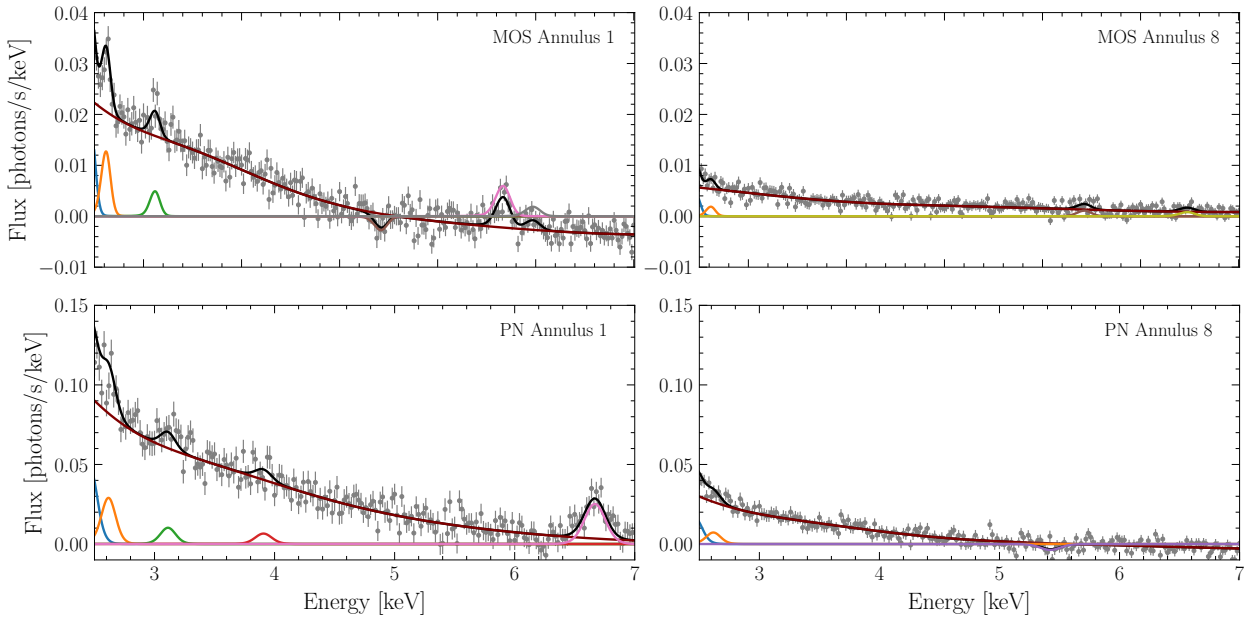



Figure H.2: The same background-subtracted data sets illustrated in Fig. H.1 (also down-binned), but now shown along with their best-fits under the null hypothesis. The best-fit model prediction is shown in black, which may be decomposed into the contribution from the GP model (dark red) and the contributions from the individual background lines (colored curves). Note that the background lines to include in the analysis are determined independently in each annulus, as described in the text.

#### H.1.3.3 The joint likelihood and background mismodeling

After constructing the profile log likelihoods $q_k(A_{\rm sig})$ in each energy annulus ($k = 1, 2, \cdots, 8$) we then convert from $A_{\rm sig}$, which has units of cts/cm$^2$/s/sr, to $\sin^2(2\theta)$ using the relation

$$\Phi \approx 0.26\,\text{photons/cm}^2\text{/s/sr} \times \left(\frac{m_\chi}{7.0\,\text{keV}}\right)^4 \left(\frac{D}{10^{29}\,\text{keV/cm}^2}\right) \left(\frac{\sin^2(2\theta)}{10^{-10}}\right) . \tag{H.6}$$

To do so we use the background-subtracted $D$-factors, as discussed in the main Letter. Then, at each test mass point for the DM model we construct the joint profile likelihood

$$q_{\rm joint}\big(\sin^2(2\theta)\big) = \sum_{k=1}^{8} q_k\big(\sin^2(2\theta)\big) , \tag{H.7}$$

for both MOS and PN independently. Later, we will also combine the MOS and PN profile likelihoods to construct our final joint profile likelihood that we use to search for evidence of decay DM. First, however, we analyze the joint MOS and PN profile likelihoods independently for evidence of background mismodeling.

We test and account for possible background mismodeling by extending the background model to include a component that is totally degenerate with the signal. This is a conservative approach that would remove all sensitivity to a UXL if the amplitude for this additional signal-like component were left free. Therefore we penalize the amplitude of such a signal like feature in the background model with a zero-mean Gaussian likelihood with variance hyperparameter $\sigma^2_{\rm spur}$ . The approach we follow was developed and implemented in [279–281] within the context of searches for narrow spectral features in $\gamma$-ray astronomy and in the context of the Higgs boson search by the ATLAS experiment, where it is called the "suprious signal" [402]. We extend the likelihood to include two "spurious signal" nuisance parameters, one for the MOS data and one for the PN data. The MOS and PN likelihoods are then combined to produce the joint likelihood that we use for probing the DM model.

After extending the background model to include a signal-like component constrained by $\sigma^2_{\rm spur}$, the resulting profile likelihood (for either the MOS or PN data) is given by

$$\tilde{q}_{\rm joint}\big(\sin^2(2\theta)\big) = \max_{A_{\rm spur}} \left[ q_{\rm joint}\big(\sin^2(2\theta) + A_{\rm spur}\big) - \frac{(A_{\rm spur})^2}{\sigma^2_{\rm spur}} \right] , \tag{H.8}$$

where $q_{\rm joint}$ is defined in (H.7). Note that the profile likelihood now depends on the hyperparameter $\sigma^2_{\rm spur}$, which determines the strength of the spurious-signal nuisance parameter For example, in the limit $\sigma^2_{\rm spur} \to 0$ the nuisance parameter becomes fixed at zero ($A_{\rm spur} \to 0$) and the modified profile likelihood $\tilde{q}$ approaches the un-modified likelihood $q$. However, in the opposite limit $\sigma^2_{\rm spur} \to \infty$ we completely lose constraining power and $\tilde{q}_{\rm joint}\big(\sin^2(2\theta)\big) \to 0$ for all $\sin^2(2\theta)$.

In practice, we determine the value of the hyperparameter at each test mass point independently. The philosophy is that if there is evidence that the background model is not properly describing the data in the immediate energy side-bands around a mass point of interest, then we should account for the possibility, through $A_{\rm spur}$, of similar background mismodeling at our mass point of interest. Specifically, we implement the following approach. At a given mass point $m^m_\chi$, where $m$ is the index that labels the mass point, we consider the subset of test mass points in a 2 keV window around $m^m_\chi$, masking: (i) a 0.4 keV window in mass around $m^m_\chi$ and (ii) masking 0.1 keV windows around the locations are all background lines that were included in the analyses of the annuli. Each test mass point within this side-band window has a best-fit $\sin^2(2\theta)$ from the likelihood analysis without the inclusion of the spurious-signal nuisance parameter. The ensemble of best-fit points in the side-band window is denoted by $\{\sin^2(2\theta)\}_m$. We compute the variance over this ensemble of

best-fit points, $\mathrm{Var}\left[\{\sin^2(2\theta)\}_m\right]_{\mathrm{observed}}$. The observed variance is then compared to the expected variance $\mathrm{Var}\left[\{\sin^2(2\theta)\}_m\right]_{\mathrm{expected}}$, and specifically we set

$$\sigma^2_{\mathrm{spur,m}} = \max\left[0, \mathrm{Var}\left[\{\sin^2(2\theta)\}_m\right]_{\mathrm{observed}} - \mathrm{Var}\left[\{\sin^2(2\theta)\}_m\right]_{\mathrm{expected}}\right] , \tag{H.9}$$

where $\sigma^2_{\mathrm{spur,m}}$ denotes the hyperparameter at the mass point $m^m_\chi$. The expected side-band best-fit variance $\mathrm{Var}\left[\{\sin^2(2\theta)\}_m\right]_{\mathrm{expected}}$ is computed from 500 MC simulations of the null hypothesis. The null hypothesis model is that given by the fit of the background model to the data without any extra UXL signal components.

We expect $\sigma^2_{\mathrm{spur,m}}$ to be non-zero if there is background mismodeling in the energy side-band, which increases the variance of observed best-fit points relative to the expectation under the null hypothesis. However, sometimes $\sigma^2_{\mathrm{spur,m}}$ will be non-zero simply because of statistical fluctuations in the observed side-bands, in which case the nuisance parameter will weaken the limits more than intended. However, this occasional weakening of the limits is worth the advantage of having an analysis framework that is more robust to mismodeling. Indeed, we know that there is an opportunity for some degree of background mismodeling because we have chosen to only include background lines that pass some significance threshold, and thus the aggregate effect of the sub-threshold lines could lead to mismodeling that could be partially mitigated by $A_{\mathrm{spur}}$.

In Fig. H.3 we illustrate the values of $\sigma^2_{\mathrm{spur,m}}$ (labeled MOS Syst. and PN Syst.) that we find from the data analyses of the MOS and PN data. We compare the hyperparameter to the statistical uncertainty on $\sin^2(2\theta)$, labeled MOS Stat. and PN Stat. Note that the statistical uncertainties are computed from the Hessian of the log-likelihood, for that data set, about the best-fit coupling at a fixed UXL energy, without the inclusion of $A_{\mathrm{spur}}$. For both MOS and PN we see that the background mismodeling uncertainties, as captured by $A_{\mathrm{spur}}$, may dominate the statistical uncertainties at some low energies, though the nuisance parameter appears more important for PN than for MOS.

It is interesting to consider the ensemble of discovery TSs in favor of the DM model across all tested mass points. We denote this distribution of TSs without the spurious-signal by $\mathbf{T}$, while with the inclusion of the spurious-signal nuisance parameter we call this distribution $\mathbf{T}_{\mathrm{sys}}$. We expect $\mathbf{T}_{\mathrm{sys}}$ to have fewer high-TS points than $\mathbf{T}$. In the left and center panels of Fig. H.4 we illustrate the distributions of TSs for both $\mathbf{T}$ (labeled Data) and $\mathbf{T}_{\mathrm{sys}}$ (labeled Data w/ Nuisance Parameter) for MOS and PN, respectively. More specifically, in that figure we illustrate the survival fractions for the distributions, which show the faction of TSs in $\mathbf{T}$ or $\mathbf{T}_{\mathrm{sys}}$ with a value above the TS indicated on the $x$-axis. Asymptotically we expect, up to the caveat that we used the Bayesian marginal likelihood of the GP to define our TSs, that the TSs should be $\chi^2$ distributed [172]. The survival function of the $\chi^2$ distribution is shown in Fig. H.4. We verify with a large number of MC

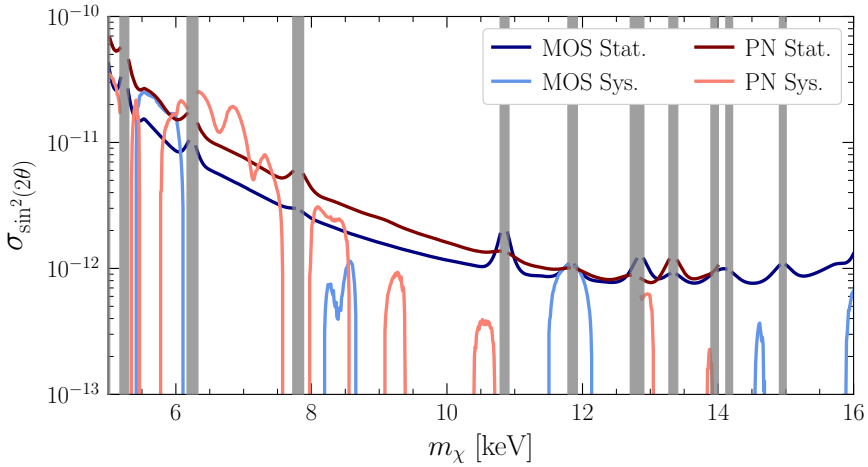



Figure H.3: The spurious-signal hyperparameter $\sigma^2_{\text{spur,m}}$ (labeled MOS Sys. and PN Sys.), as computed in (H.9), as a function of the DM mass. For both MOS and PN the nuisance parameter $A_{\text{spur}}$ is predominantly active at low energies, and it plays a more significant role in PN than in MOS. We compare the hyperparameter to the statistical uncertainties (labeled MOS Stat. and PN Stat.), which are computed from the Hessian of the log-likelihood (without the spurious-signal) about the best-fit mixing angle at a fixed energy. We note that several of the sharp variations of the expected sensitivity shown in Fig. 9.3 arise as a result of the variations of the spurious signal hyperparameter shown here.

simulations that the that the null-distribution of TSs is indeed $\chi^2$-distributed for both MOS and PN datasets. The results of these tests are labeled "Monte Carlo" in Fig. H.4 and overlap with the $\chi^2$ distribution, providing evidence that we are in the asymptotic regime [172].

Because there are a finite number of samples in $\mathbf{T}$ and locations spaced within the detector energy resolution are correlated, the survival function for the observed data is not expected to follow the $\chi^2$-distribution exactly. The green and gold bands in Fig. H.4 show the 68% and 95% containment regions for the survival fraction computed over 500 MC realizations of $\mathbf{T}$. We expect that the data should fall within these bands if no signal is present, which is analogous to the green and gold bands for the significance in Fig. H.5. In the left and center panels of Fig. H.4 we may see that the distributions of $\mathbf{T}$ for MOS and PN are broadly consistent with the MC expectations. The distributions of $\mathbf{T}_{\text{sys}}$, as expected, fall off slightly faster at large values of the TS. The right panel of Fig. H.4 shows the survival fraction for the combined analysis where we combine the MOS and PN profile likelihoods, with and without the spurious-signal. The most significant test point has a significance slightly above $2\sigma$ local significance, which is less than $1\sigma$ in global significance. Thus, we conclude that there is no evidence for decaying DM in our analysis above our $5\sigma$ global predetermined detection threshold.

The effect of the spurious-signal nuisance parameter on the individual MOS and PN limits is

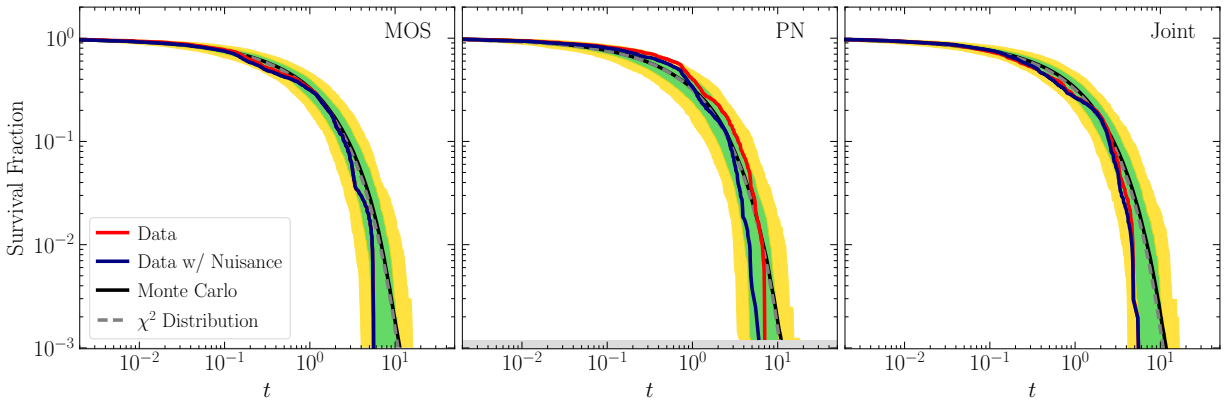



Figure H.4: (Left) The survival function of the test statistic for discovery in the analysis of the MOS data. Under the null hypothesis, and for a large number of samples, the survival fractions are expected to follow the $\chi^2$ distribution, as verified by MC (as labeled). At a finite number of samples the expected chi-square distributions are found from MC to be expected to be contained within the green and gold shaded regions at 68% and 95% confidence, respectively. The negligible effect of the systematic nuisance parameter can be seen by comparing the survival function without the nuisance parameter (red, labelled "Data") and with the nuisance parameter (blue, labeled "Data w/ Nuisance Parameter"). (Center) As in the left panel, but for the PN analysis. (Right) The survival function for the joint analysis of MOS and PN data. In blue, the survival function for the joined PN and MOS analysis without systematic nuisance parameters; in red, the survival function for the joint analysis when the PN and MOS results are corrected by their independently-tuned systematic nuisance parameters prior to joining.

illustrated in Fig. H.5. The inclusion of the nuisance parameter slightly decreases the discovery TSs at low masses and also causes a slight weakening of the limits. Note that the expectations under the null hypothesis are indicated for the spurious-signal-corrected analysis in that figure.

## H.2 Extended Results

In this section we present extended results for the analyses that go into producing Fig. 9.3. First, we provide a measure of the goodness-of-fit of our null model to the data, quantified through the $\chi^2$ per degree of freedom (dof), in each annulus for the PN and MOS data sets in Tab. H.3. Note that we also quote the $p$-value associated with the $\chi^2$ per dof, with smaller numbers indicating a worse null-model fit. We observe acceptable $p$-values ($p \gtrsim 0.1$) in all rings except for Ring 3 of the PN data set, which realizes a $p$-value associated with the $\chi^2/\mathrm{dof}$ of $p \approx 5.7 \times 10^{-6}$. We would not expect to observe a $p$-value this small in any of the 16 rings. For example, Fig. H.3 shows some evidence for mild systematic uncertainties at low energies in the PN data, though these are captured through our spurious-signal formalism. We also note that there is some indication that the poor $\chi^2/\mathrm{dof}$ in PN Ring 3 arises from statistical fluctuations on scales much smaller than the detector

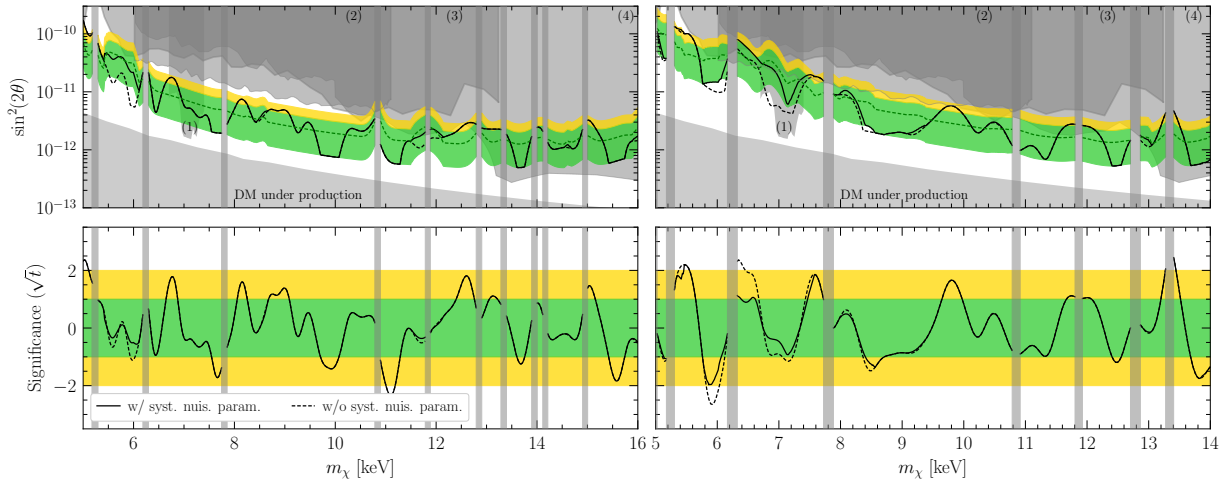



Figure H.5: As in Fig. 9.3, but for the MOS (left panel) and PN (right panel) analyses individually and with and without the spurious-signal nuisance parameter. The $1\sigma$ and $2\sigma$ expectations are shown only for the case with the spurious-signal nuisance parameter. The limits without the nuisance parameter are slightly stronger at low masses. The sharp variations in the expected sensitivity, especially visible in the PN results, arise from how the spurious-signal hyperparameter is determined through the sliding window procedure.

energy resolution; for example, down-binning that data set to bins of width 45 eV, which is still smaller by a factor of a few relative to the energy resolution across the full energy range, improves the $p$-value associated with the $\chi^2/\mathrm{dof}$ to $p \approx 4 \times 10^{-3}$. As an additional test, we compare the results obtained without the spurious-signal formalism in the joint analysis of the PN data with and without the inclusion of the data in Ring 3. These results are presented in Fig. H.6, and are qualitatively unchanged by the inclusion or exclusion of the PN Ring 3 data set. We also provide best-fit normalizations for our GP kernels, presented in Tab. H.4.

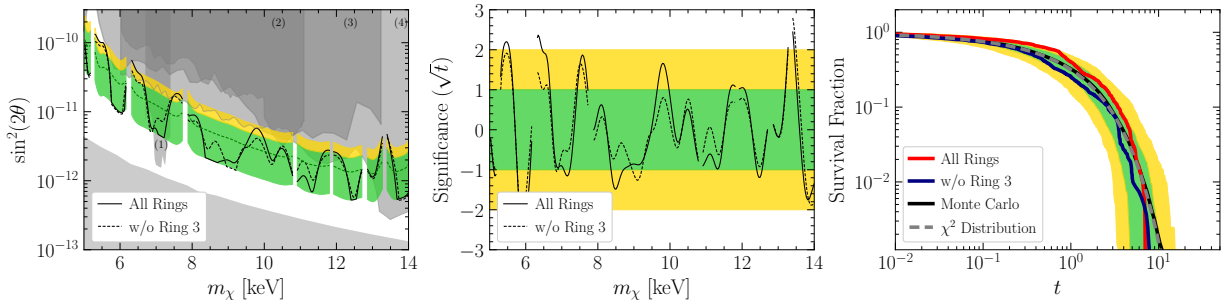



Figure H.6: A comparison of all results obtained in the joint analysis of PN data with and without the inclusion of Ring 3, which may be subject to systematic mismodeling. Note that for this comparison we do not profile over the spurious-signal nuisance parameter.

In Fig. H.7 we present the main result in Fig. 9.3 in terms of the DM lifetime instead of in

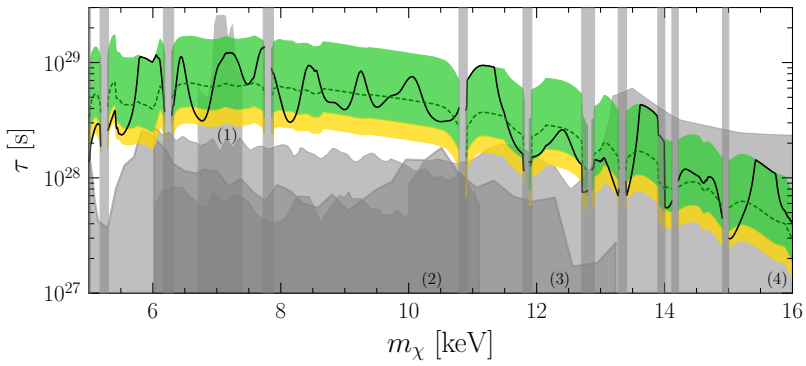



Figure H.7: As in Fig. 9.3, but interpreted as limits on the DM lifetime. This figure applies for DM whose decays produce a single mono-energetic photon at energy $m_\chi/2$. If the DM decay produces two photons (as in an axion model), then the lifetime limits are twice as strong.

terms of $\sin^2(2\theta)$. The result in Fig. H.7 is more general than in Fig. 9.3 since it holds for more general DM models beyond the sterile neutrino model. Note, however, that this figure applies to DM whose decays produces one mono-energetic photon at energy $m_\chi/2$. Axion-like models produce two photons during the decay, in which case the limits are twice as strong as those shown in Fig. H.7.

| Instrument | Ring 1 | Ring 2 | Ring 3 | Ring 4 | Ring 5 | Ring 6 | Ring 7 | Ring 8 |
|---|---|---|---|---|---|---|---|---|
| MOS [$\chi^2$/dof] | 1133.6/1093 | 1069.6/1096 | 1190.7/1091 | 1114.2/1096 | 1157.7/1098 | 1073.2/1097 | 1083.4/1095 | 1100.1/1095 |
| MOS [$p$-value] | 0.19 | 0.71 | 0.02 | 0.34 | 0.10 | 0.69 | 0.59 | 0.45 |
| PN [$\chi^2$/dof] | 860.7/894 | 845.2/897 | 1091.8/894 | 873.2/893 | 915.5/899 | 920.8/897 | 838.1/897 | 874.5/896 |
| PN [$p$-value] | 0.77 | 0.89 | $5.7 \times 10^{-6}$ | 0.65 | 0.34 | 0.25 | 0.92 | 0.68 |

Table H.3: The goodness-of-fit of the null model fit in each annulus for PN and MOS data sets as measured by the $\chi^2$ divided by the number of degrees of freedom (dof). With the exception of Ring 3 of the PN data set, this measure indicates an acceptable goodness-of-fit to the data under the null, as quantified through the $p$-value (see text for details).

| Instrument | Ring 1 | Ring 2 | Ring 3 | Ring 4 | Ring 5 | Ring 6 | Ring 7 | Ring 8 |
|---|---|---|---|---|---|---|---|---|
| MOS | $6.9 \times 10^{-3}$ | $1.6 \times 10^{-3}$ | $1.1 \times 10^{-3}$ | $3.4 \times 10^{-3}$ | $6.6 \times 10^{-4}$ | $1.9 \times 10^{-3}$ | $9.0 \times 10^{-4}$ | $1.5 \times 10^{-3}$ |
| PN | $2.3 \times 10^{-2}$ | $5.3 \times 10^{-3}$ | $5.7 \times 10^{-3}$ | $9.1 \times 10^{-3}$ | $2.0 \times 10^{-3}$ | $4.2 \times 10^{-3}$ | $2.0 \times 10^{-4}$ | $1.1 \times 10^{-2}$ |

Table H.4: The best fit normalization of the GP kernel for each ring in PN and MOS. We present $\sqrt{A_{\rm GP}}$ in units of photons/cm$^2$/s/keV for $A_{\rm GP}$ defined in (H.3).

Next, we show our results from the analyses to the individual MOS and PN annuli. In Figs. H.8 through H.15 we show the best-fit fluxes and significances (times the sign of the excess or deficit)

for the UXLs for all of the annuli and for both MOS and PN. Note that the shaded grey regions denote the masks that we use to avoid searching for UXLs in the direct vicinity of background lines included in the analyses.

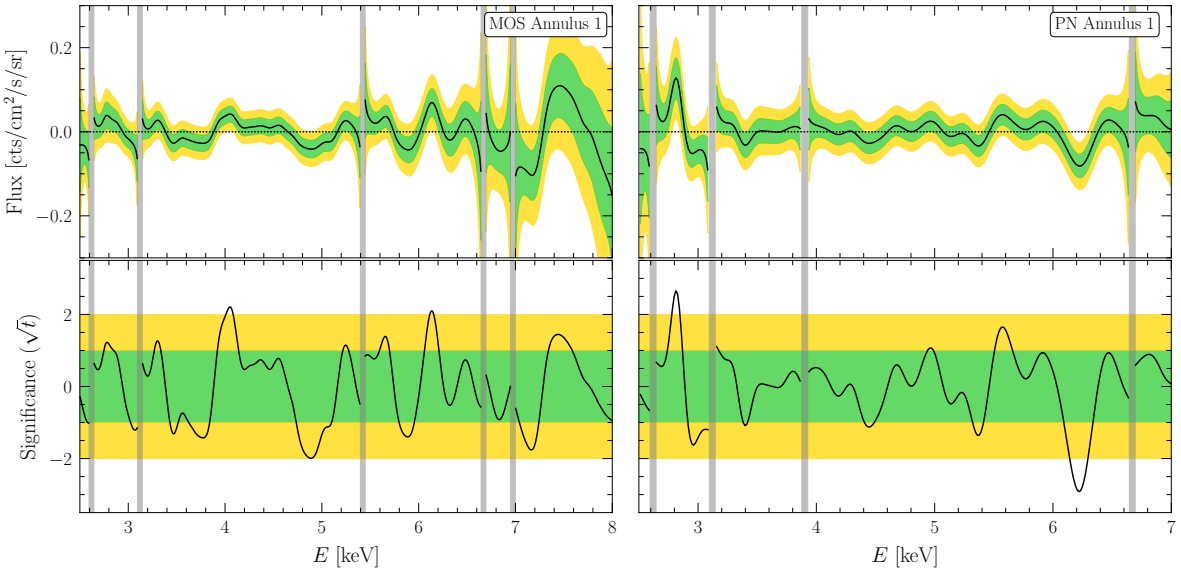



Figure H.8: (Upper Left) The best-fit signal flux, and 1 and $2\sigma$ uncertainties, as a function of the central UXL energy across our full energy range for the innermost MOS ring. (Lower Left) The corresponding significance in favor of the signal model, multiplied by the sign of the best fit UXL normalization at that energy, along with the $1/2\sigma$ expectations under the null hypothesis. (Right Panel) As in the left panel but for the innermost PN annulus.

The distribution of discovery TSs that we observe in the individual annuli all appear consistent with expectations from MC, as illustrated in Fig. H.16 for MOS and Fig. H.17 for PN. These figures illustrate the survival fractions of TSs, as in Fig. H.4, but at the level of the individual annuli instead of the joint analysis. Note that the MC expectations are constructed independently for each annulus and each data set. These results do not include the systematic nuisance parameter since that is only included at the level of the joint likelihood, after combining the results from all of the individual annuli.

## H.3 Synthetic signal tests for the fiducial analysis

In this section we verify that our analysis framework has the ability to discover real signals if they are present in the data. We do so by injecting a synthetic signal into the real data and analyzing the hybrid data set with our full analysis. We also demonstrate the full analysis as applied to fully synthetic data generated with varying injected signal strengths.

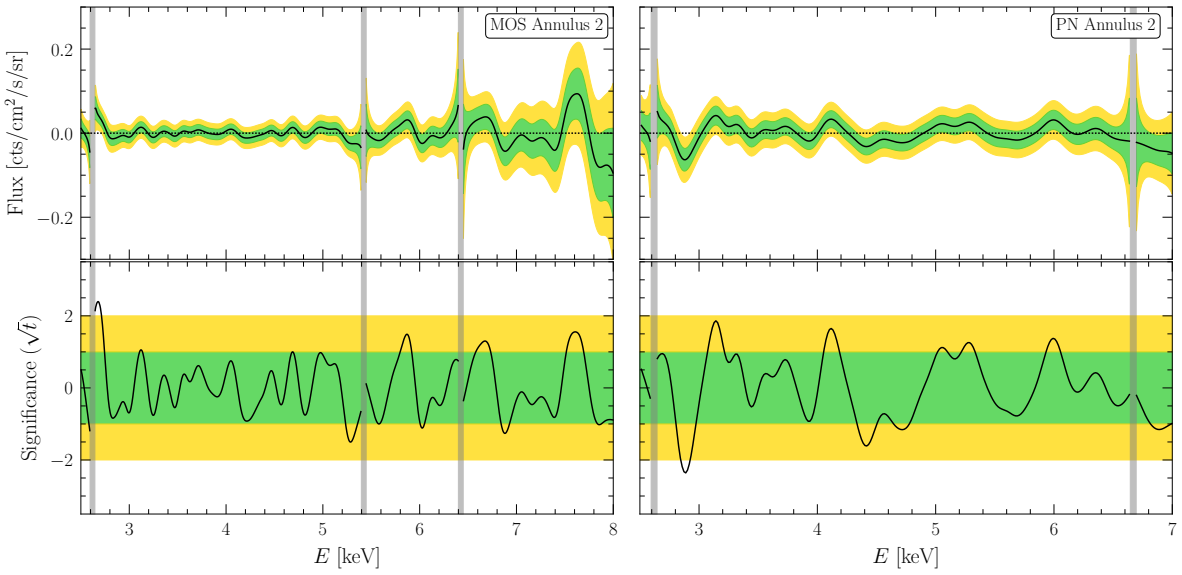



Figure H.9: As in Fig. H.8 but for annulus 2.

### H.3.1 Signal injection in real data

For injection tests in the real data, we chose a DM mass $m_\chi = 7.0$ keV and a mixing angle $\sin^2(2\theta) = 2.5 \times 10^{-11}$. We chose this mixing angle because we expect such a signal to be detected at approximately $5\sigma$ significance. We forward model this signal through the appropriate MOS and PN detector responses, draw Poisson counts, and then add these counts to the actual data sets. The results of the data analysis of the hybrid data are illustrated in Fig. H.18. In the top panel we show the 95% upper limits for MOS, PN, and the joint analysis, with and without the systematic nuisance parameter. Note that the injected signal is indicated by the red star. The upper limits weaken at the injected signal point, as expected, and do not exclude the injected signal coupling. In the second row we show the corresponding detection significances. The signal is detected at nearly $5\sigma$ in MOS alone and at around $2\sigma$ in PN. The systematics nuisance parameter slightly reduces the significance of the discovery, but by a minimal amount since we mask a 0.4 keV window around the test mass when determining the systematics nuisance parameter. In the third row we show how the discovery of the injected signal extends the survival function to higher TS values. Lastly, in the bottom row we show the 1, 2, and 3 $\sigma$ best-fit regions in the $m_\chi$-$\sin^2(2\theta)$ plane for the DM candidate. In red we mark the location of the injected signal, which is recovered appropriately.

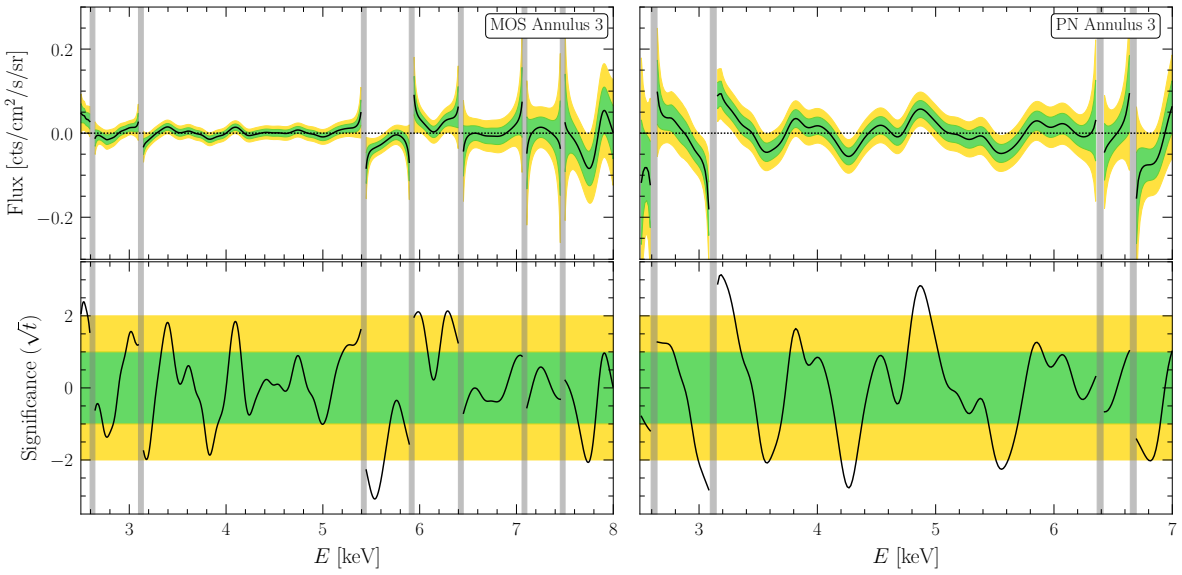



Figure H.10: As in Fig. H.8 but for annulus 3.

### H.3.2 Signal injection in synthetic background data

For injection tests on the real data, we first generate synthetic data according to the best-fit null models for each of the eight rings studied in MOS and PN data sets. We then inject a synthetic signal at a specified value of $\sin^2(2\theta)$ on top of the null-model realizations using the same procedure as applied for the signal injection on the real data and repeat our full analysis procedure in search of the injected signal with the exception that we do not apply a nuisance parameter tuning and correction. We perform 1000 independent realizations and analyses for each value of $\sin^2(2\theta$, and we repeat this procedure for 30 values of $\sin^2(2\theta)$ between $10^{-13}$ and $10^{-10}$ for two different neutrino masses: 7.0 keV and 11.5 keV. The results of the data analysis of the hybrid data are illustrated in Fig. H.19. In the top row, we show ensemble statistics for the 95% upper limits as a function of injected signal strength for the two neutrino masses studied in this test. In the bottom row, we show the ensemble statistics of the recovered detection test statistic as a function of the injected signal strength. The upper limits weaken with increasing injected signal strength without excluding the true value of the injected signal. Moreover, the detection test statistic smoothly increases as a function of increasing injected signal strength. Critically, at large injected signal strength, the test statistic safely exceeds $TS \approx 30$, which is the approximate threshold for a $5\sigma$ detection after correcting for the look-elsewhere effect.

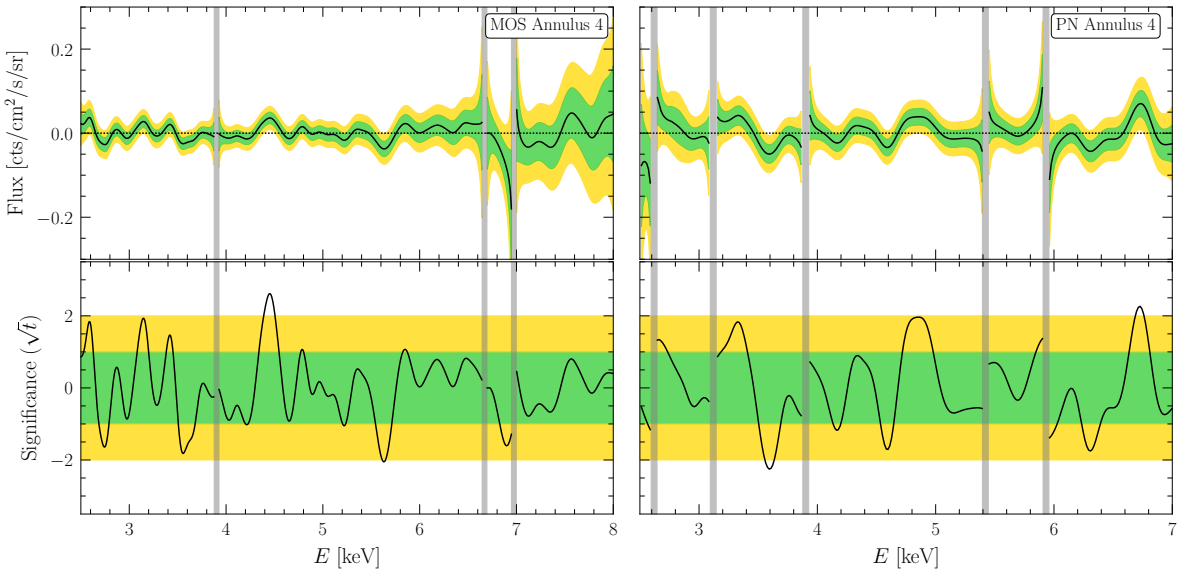



Figure H.11: As in Fig. H.8 but for annulus 4.

## H.4 Systematic Analysis Variations

Our fiducial result, which is illustrated in Fig. 9.3, made a number of physics-level and analysis-level choices. These choices are justified in the main text and the supplementary results of the proceeding sections of the SM. Still, it is worthwhile to consider how our results change for different physics and analysis assumptions and choices, as this gives an indication of the robustness of the limits and significances shown in Fig. 9.3.

### H.4.1 Alternate DM Density Profiles

In the main text, and in particular in Fig. 9.3, we adopted the conservative DM profile that was shown in Fig. 9.1. As already described, the present expectation is that in the absence of baryons, the DM halo is well described by an NFW profile. Baryons are then expected to contract this profile, increasing the DM density towards the GC, and potentially also introducing a core on top of this. For our fiducial analysis we conservatively assumed an uncontracted NFW halo, using the most conservative parameters determined within the 68% best fit region of [416]. In particular, we used an NFW profile with $r_s = 19.1$ kpc and normalized to a local DM density of $\rho_{\rm DM} = 0.29$ GeV/cm$^3$.

In Fig. H.20, we show our main results if instead we repeat the analysis for the best fit NFW profile determined in Ref. [416], which corresponds to $r_s = 15.6$ kpc and $\rho_{\rm DM} = 0.31$ GeV/cm$^3$, as well as showing results for the more realistic contracted profile. There is not a parametric form for the contracted profile, however, Ref. [416] provides a best fit model for the DM mass distribution,

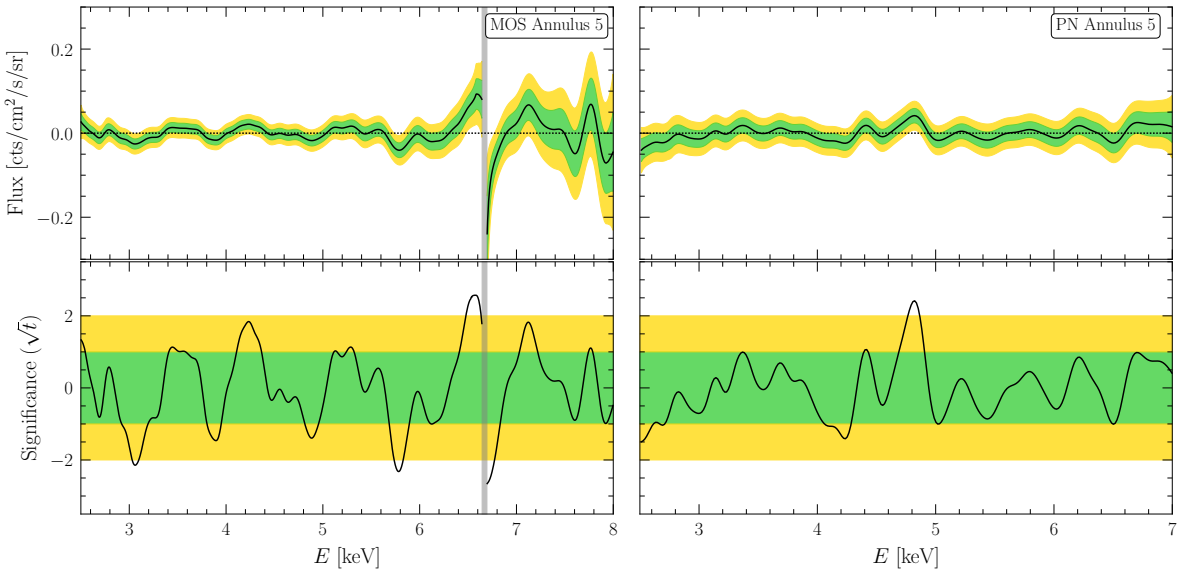



Figure H.12: As in Fig. H.8 but for annulus 5.

which we use to infer the density and then $D$-factor. The model only provides an estimate down to 1 kpc from the GC, within which we conservatively assume the density profile is completely cored.

As the figure demonstrates, adopting a more realistic contracted DM profile strengthens our limits by roughly a factor of 2. Importantly, however, changing the profile does not appreciably change the distribution of significance, and we continue to see no clear evidence for an UXL.

### H.4.2 Dependence on the GP model

For our fiducial analysis we use the GP kernel given in (H.3) with the choice $\sigma_E = 0.3$. This choice was made so that the residual background model has the ability to adjust on scales around one order of magnitude larger in scale than the energy resolution of the detectors, which are $\delta E/E \sim 0.03$. In this section we verify that our results do not depend in detail upon the particular value chosen.

First, we consider a small modification to our default analysis by taking $\sigma_E = 0.2$ and $\sigma_E = 0.4$. The results of these analyses are shown in Figs. H.21 and H.22. As a further modification of our GP modeling procedure, we repeat our analysis with the relative scale of our kernel promoted to a nuisance parameter that we independently profile in each annulus in both instruments between the range of $0.15$ and $0.9$. We report the resulting best-fit GP scales in Tab. H.5. Results for this analysis are shown in Fig. H.23. In those figures we show the 95% upper limits (upper panel), significances (middle panel), and survival fraction of significances (bottom panel). We give the results both with an without nuisance parameters. There is a slight trend where increasing $\sigma_E$

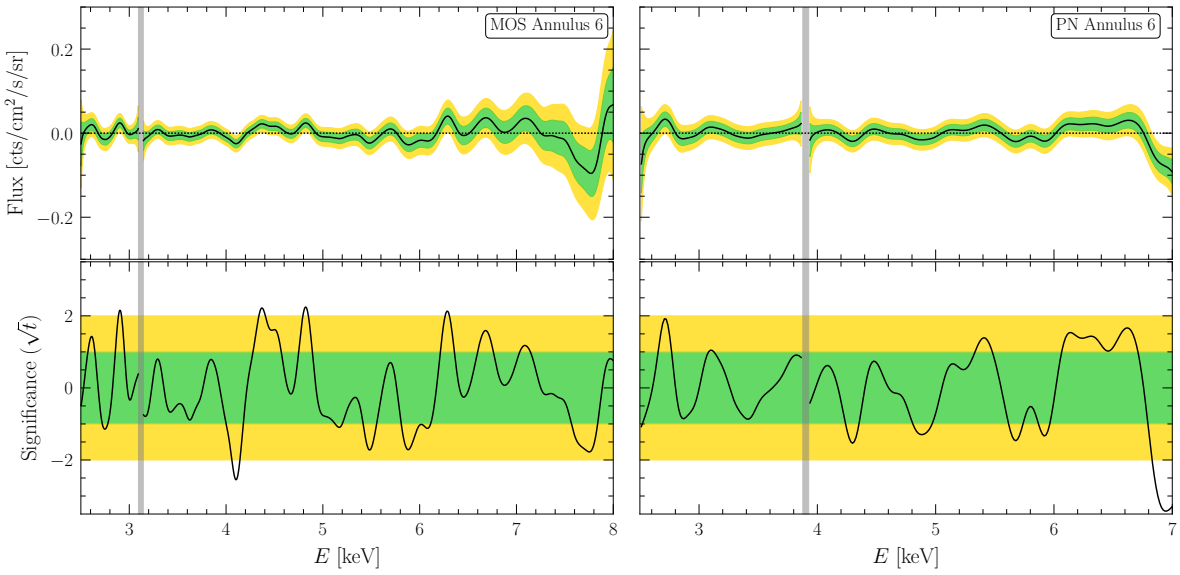



Figure H.13: As in Fig. H.8 but for annulus 6.

leads to a corresponding strengthening of the sensitivity, though this difference is minor compared to other choices in the analysis. In general, the results appear robust to the choice of $\sigma_E$.

| Instruments | Ring 1 | Ring 2 | Ring 3 | Ring 4 | Ring 5 | Ring 6 | Ring 7 | Ring 8 |
|---|---|---|---|---|---|---|---|---|
| MOS | 0.60 | 0.90 | 0.90 | 0.34 | 0.81 | 0.90 | 0.42 | 0.90 |
| PN | 0.77 | 0.84 | 0.59 | 0.90 | 0.66 | 0.28 | 0.90 | 0.54 |

Table H.5: The best-fit scale $\sigma_E$, determined under the null model, when this scale is treated as profiled nuisance parameter. In all cases except Ring 6 of PN data, the best-fit scale is larger than the scale of the kernel used in our fiducial analysis, indicating that our fiducial choice of $\sigma_E$ = 0.3 was conservative and endowed the GP model with sufficient flexibility.

Next, we consider changing the GP modeling more significantly by adopting an alternate kernel. In particular, we consider the standard (and stationary) double exponential kernel

$$K(E, E') = A_{\rm GP} \exp\left[-\frac{(E - E')^2}{2\sigma^2}\right], \tag{H.10}$$

which has the hyperparameter $\sigma^2$. Note that our fiducial kernel, given in (H.3), has the property whereby the correlation length increases with the energy resolution of the detector. The kernel in (H.10), on the other hand, has a fixed correlation length as a function of energy. In Figs. H.24 and H.25 we show the results of using the double exponential kernel with scale length $\sigma^2 = 0.5$ keV$^2$ and $\sigma^2 = 1.0$ keV$^2$, respectively. As with our fiducial kernel, in this case we also find that increasing $\sigma$ slightly increases the limits. However, the differences between the double-exponential

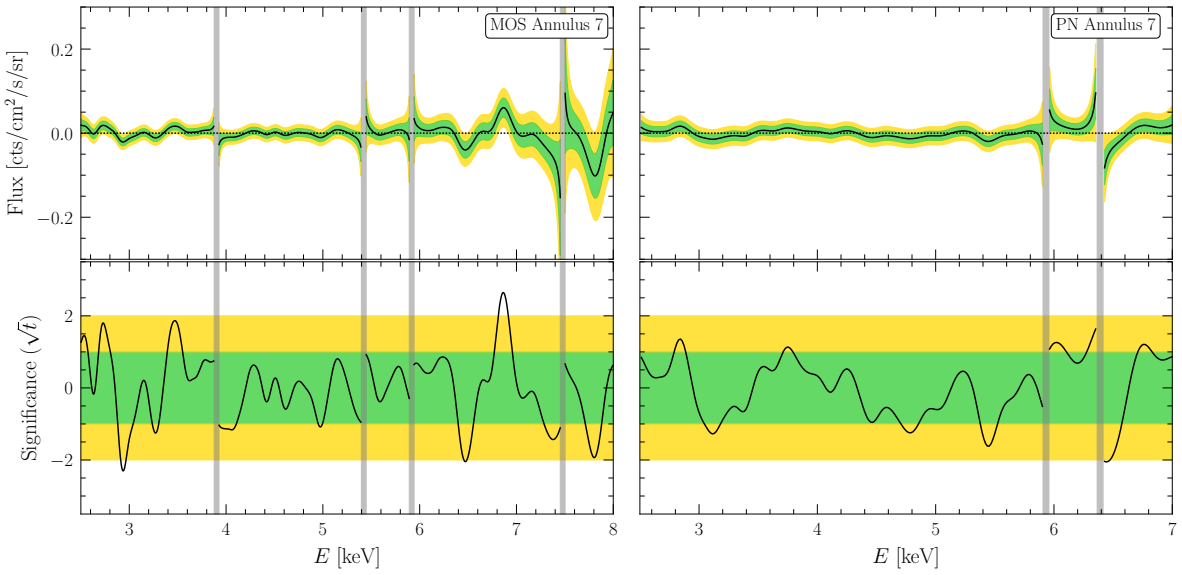



Figure H.14: As in Fig. H.8 but for annulus 7.

kernel results and our fiducial results are minor and most evident at high DM masses, $m_\chi$, where the two kernels predict the largest differences. In particular, we find no evidence for decaying DM with the alternate kernels and similar 95% upper limits. The systematic uncertainty associated with this choice is generally less than other aspects of the analysis such as our assumptions regarding the DM density profile.

A full comparison of the limits obtained under all the described GP kernel modeling choices is presented in Fig. H.26. The choice of GP kernel and treatment of its scale as a fixed or profiled parameter is shown to have a marginal impact on our limit-setting procedure.

It is worth noting that the upper limit in Fig. H.26 from the analysis where $\sigma_E$ is profiled over a nuisance parameter is, at some mass points, stronger than the upper limits in those analyses where $\sigma_E$ is fixed. This may seem counterintuitive, since the common assumption is that profiling over nuisance parameters will lead to weaker constraints than in analyses where those nuisance parameters are fixed. The key point, however, is that the previous sentence is only correct if the nuisance parameters are fixed at their best-fit values; as may be seen in *e.g.* Tab. H.5, our fixed-$\sigma_E$ analyses do not have $\sigma_E$ fixed at their best-fit values. Instead, our fixed $\sigma_E$ analyses fix this hyperparameter away from its best fit value such that the analysis is forced to give the GP model more freedom than it would otherwise want, allowing for more conservative limits in the fixed-$\sigma_E$ analyses.

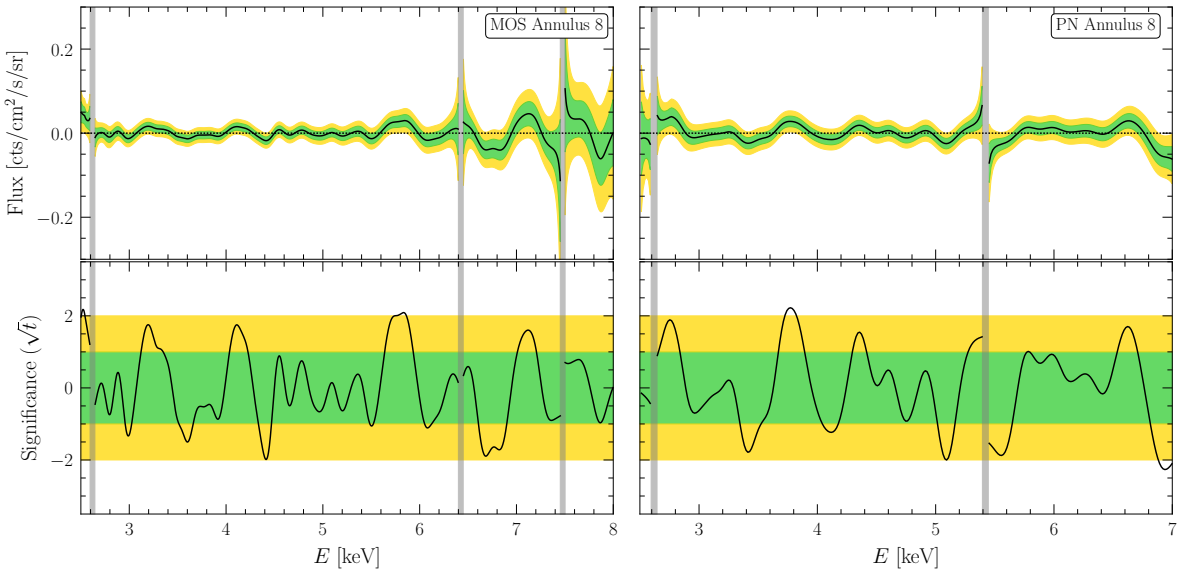



Figure H.15: As in Fig. H.8 but for annulus 8.

### H.4.3 Unmodeled lines in the vicinity of 3.5 keV

Our results have a significant impact on the decaying DM interpretation of the previously-observed 3.5 keV line from nearby galaxies and galaxy clusters [379–383]. Ref. [378] used a similar method to that presented in this work to argue that the non-observation of the UXL in *XMM-Newton* BSOs excluded the decaying DM origin of the 3.5 keV line. However, subsequent works [460, 461] questioned the validity of the results in [378] for three primary reasons: (i) the use of a narrow energy range, (ii) the possible importance of instrumental or astrophysical lines in the analysis region, (iii) the $D$-factor profile chosen with a local DM density of $0.4$ GeV/cm$^3$. These points were addressed extensively in the response [384], and we do not review the arguments here for how these points are addressed within the context of the analysis in [378].

Here we point out that the analysis in this work provides a probe of the decaying DM origin of the 3.5 keV line that is more robust to systematic uncertainties than [378] and that the null results from this work strongly disfavor the decaying DM interpretation of the 3.5 keV line. Ref. [378] performed a similar analysis to this work, but the analysis focused on the limited mass range from 6.7 to 7.4 keV. As in this work [378] used *XMM-Newton* blank sky data, with a comparable exposure time within the signal ROI to that in this work. As mentioned in the main text, Ref. [378] used a joint likelihood over individual exposures, as opposed to this work where we stack the data in rings and construct the joint likelihood in individual rings. Use of the ringed data facilitates our background subtraction and GP modeling procedures, in part because the number of counts in each ring is large enough that we may make the Gaussian approximation to the Poisson likelihood. This work also performs a more systematic accounting of astrophysical and instrumental lines that

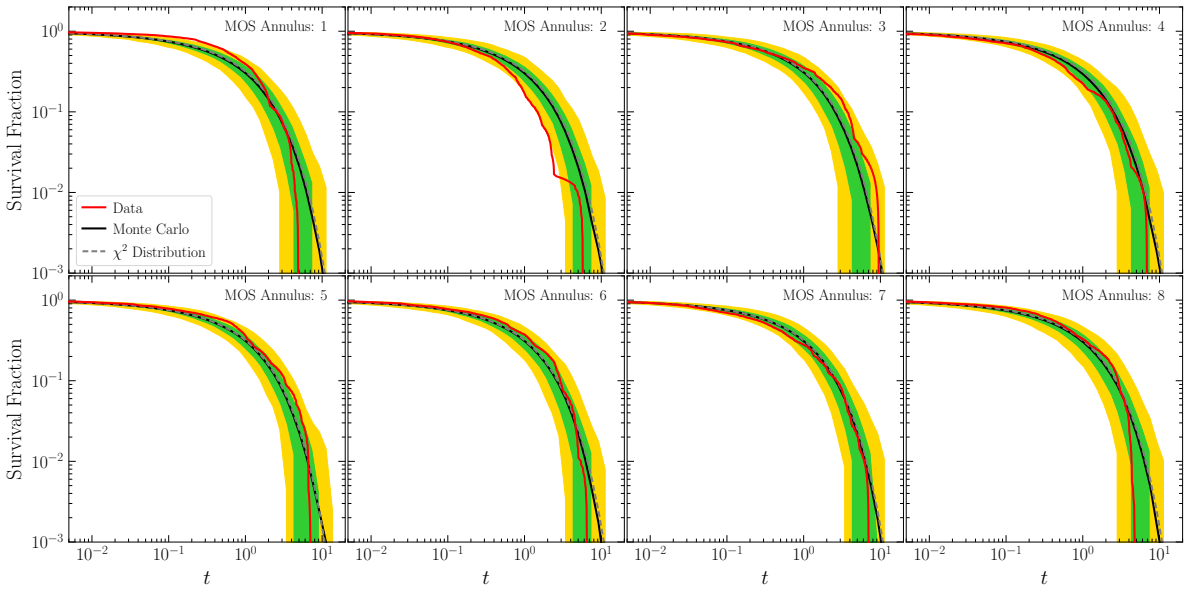



Figure H.16: As in Fig. H.4 but for the individual MOS annuli. Note that the systematic nuisance parameter has not been applied since that is only incorporated in the joint likelihood that combines the results from the individual annuli.

are not fully removed by the background subtraction process. Because we analyze a wide energy range in this analysis, we are able to use energy side-bands to determine the hyperparameter for the spurious-signal contribution to the likelihood, which accounts for residual mismodeling. Thus while the limit presented in this work in Fig. 9.3 is slightly weaker than the fiducial limit from [378], it is more robust to mismodeling. Furthermore, we use a more conservative $D$-factor profile in this work, though astrophysical uncertainties on the DM density profile are not sufficiently large to explain why a decaying DM signal would have appeared in nearby galaxies and clusters but not in this work (see [384] for a discussion of this point).

Still, in this section we investigate the potential for mismodeling in the vicinity of 3.5 keV. In particular, [462] argued that lines may be present near 3.32 and 3.68 keV in both the MOS and PN data. Note that in [378] these possible lines were tested for and their inclusion did not change the central conclusion of that work. Moreover, there is no robust evidence to-date for these lines in the MOS and PN data sets. For completeness, however, we investigate how the inclusion of these lines affects the results of the analysis in this work. Importantly, following our normal line-dropping procedure neither the 3.32 nor the 3.68 keV lines meet our criterion for inclusion in any of the rings for either MOS or PN. This itself serves as evidence for the non-importance of these line candidates on our conclusions. However, as a systematic test we perform an analysis where we include these two lines in all of our rings for both MOS and PN, while performing the normal line-dropping procedure for the rest of the background lines. We treat the amplitude of these lines

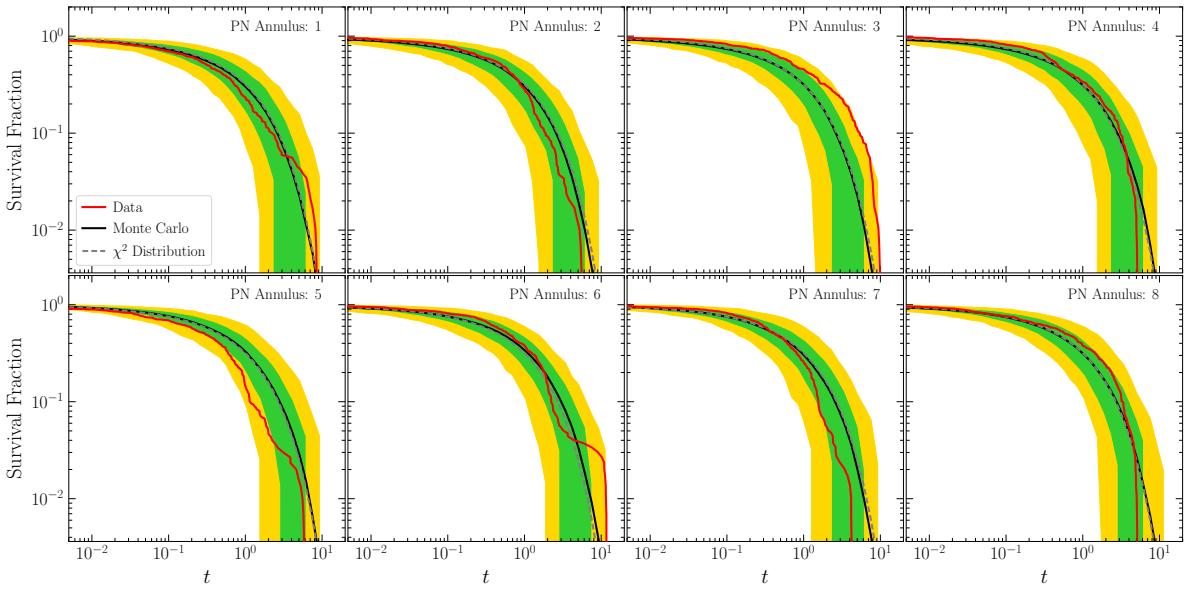



Figure H.17: As in Fig. H.16 but for the PN data sets.

as a nuisance parameter which is allowed to take arbitrarily large positive or negative values.

A summary of the full results of the analysis which includes these additional lines is provided in Fig. H.27. No new detections are made. The limits obtained by this analysis in the 6-8 keV range are compared with the results obtained in our fiducial analysis in Fig. H.28, which reveals small but unimportant changes in our limits.

### H.4.4 Analysis of Fully Stacked Data

In the main body, we divided our signal ROI into rings and modeled the flux independently in each ring. The motivation behind this choice was to incorporate spatial information into the analysis, particularly as we expect the flux of an actual DM decay signal to steadily increase towards the GC. Here we show the results of an alternative approach where instead of modeling the data ring-by-ring, we instead combine the data in the innermost three rings of the signal ROI and model that directly. We effectively are then left with a single combined ring, which we analyze using our fiducial procedure.

In Fig. H.29 we show the resulting limit in the case where we also subtract the background-ROI flux from the stacked signal region data. While there are small differences, the resulting sensitivity and limits from this simpler approach are in good qualitative agreement to those of our default analysis. In detail, the result here are slightly weaker, which is as expected because there is less information in the signal ROI (we use fewer rings and by stacking the spatial information is partially erased).

Next, in Fig. H.30 we repeat this procedure but without subtracting the background flux. The differences are now more noticeable - the expected and resulting limit undergoes larger fluctuations and there are several mildly significant excesses. This emphasizes the importance of the background subtraction procedure in simplifying the data, particularly around bright instrumental lines.

### H.4.5 Parametric Modeling without Background Subtraction

In this section, we detail an alternate analysis to the one presented in the main body of this paper and provide a comparison between the fiducial and alternate analysis in a representative example over the 8-9 keV mass range. The alternate framework uses the same data as used in our primary analysis. However, a more traditional approach is adopted for the background modeling. Firstly, we consider the unsubstracted data in each ring within the signal ROI. The flux within each annulus is then modeled as follows. The background and putative signal lines are treated identically to our fiducial approach, but the smooth background contribution is modeled parametrically using an unfolded second order polynomial, rather than with a GP model. The three parameters that define the quadratic background component are treated as nuisance parameters and profiled over. As the quadratic background has less freedom than the GP model, we restrict to a smaller energy range. Specifically, we determine the energy range by fitting a Gaussian to the detector response at a given putative signal energy, and we define our energy range to extend 5 standard deviations out from the signal energy in either direction. In the 8-9 keV DM mass range, this corresponds to an approximate energy range of 0.60 keV. Furthermore, background lines within 7 standard deviations of the signal energy are included in the model. Thus, in the 8-9 keV mass range, the only line included is the 4.52 keV instrumental line for all PN annuli. We do not include the systematic nuisance parameter modeling for this example.

While our background modeling is significantly different in this case, we find again that our results are qualitatively unchanged compared to our fiducial analysis. To provide a representative example, in Fig. H.31 we show the comparison between our fiducial analysis (without the systematics nuisance parameter to facilitate the comparison) and this alternate approach over the mass range 8-9 keV. As can be seen, the expected sensitivity of the two approaches is almost identical. This is a significant further demonstration that our specific choice of background model is not underpinning our sensitivity.

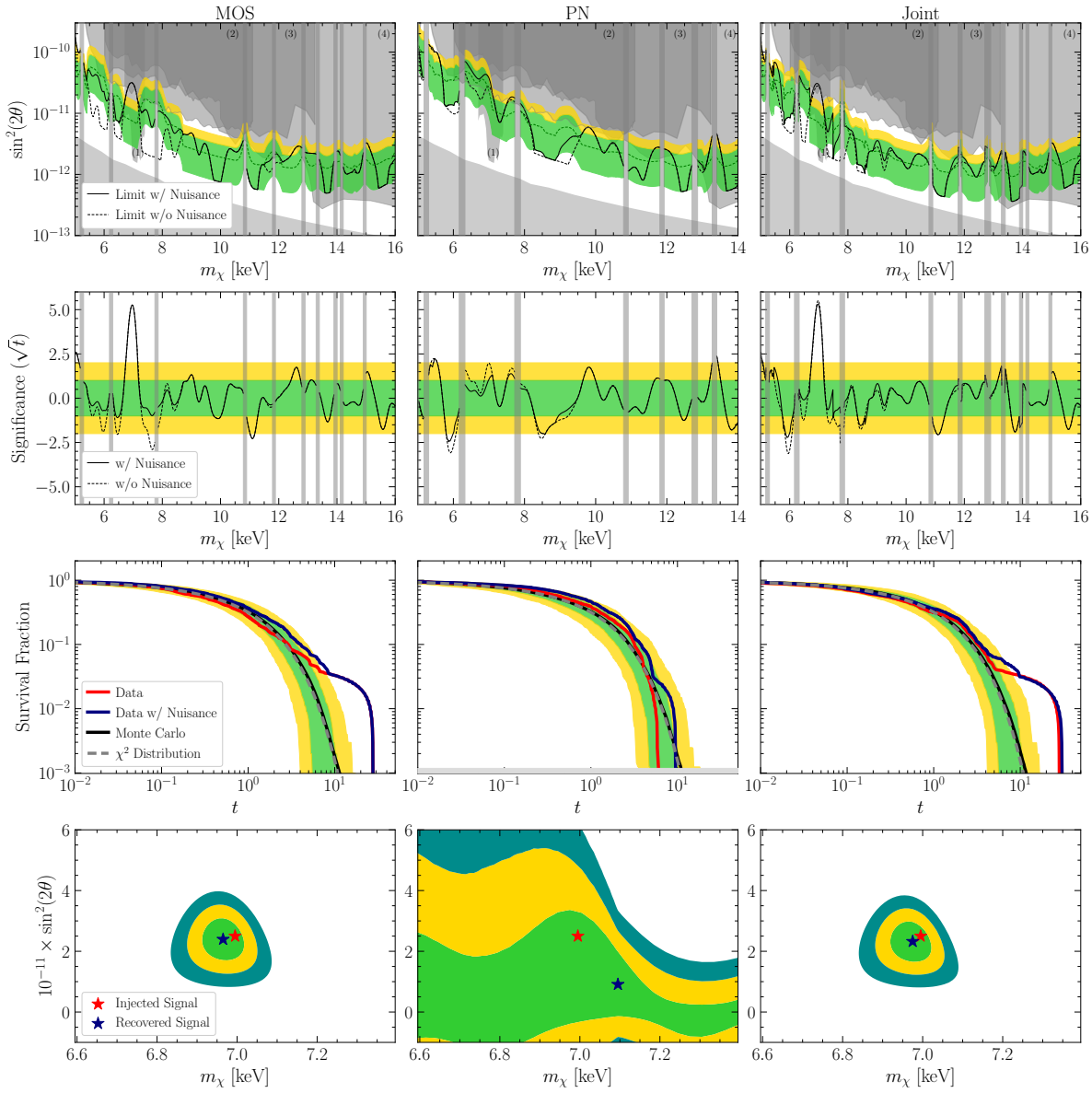



Figure H.18: The results of the analysis of the hybrid data that consists of the real MOS and PN data plus a synthetic DM signal. The DM signal is generated with mass $m_\chi = 7.0$ keV and mixing angle $\sin^2(2\theta) = 2.5 \times 10^{-11}$ as described in the text. The top, middle, and third rows are analogous to Figs. H.4 and H.5, but for the hybrid data set. The last row shows the 1, 2, and 3 $\sigma$ recovered parameter space for the signal in the mass and mixing angle plane. The best-fit recovered signal is indicated in dark blue, while the red star denotes the true value injected. The synthetic signal is appropriately recovered, adding confidence that our analysis procedure has the ability to detect real DM signals if present in the data.

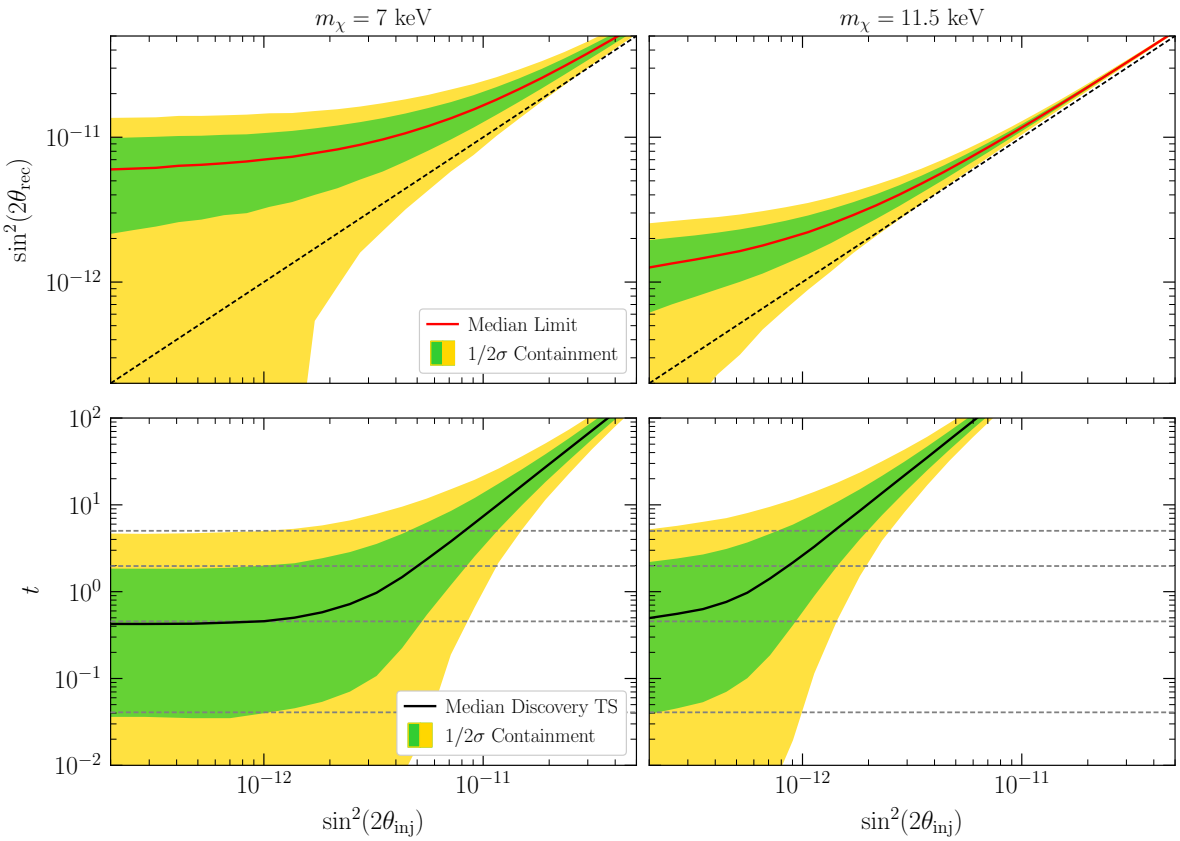



Figure H.19: (*Top Row*) In red, the median 95$^{\rm th}$ percentile upper limit on the recovered signal as a function of the injection signal strength at two neutrino masses evaluated on synthetic data. We additionally indicate the 1 and $2\sigma$ containment intervals for the ensemble of upper limits realized at each injected signal strength. Note that these upper limits are not power constrained. These results demonstrate that our analysis framework places robust upper limits that do not rule out an injected signal. (*Bottom Row*) In black, the median recovered detection test statistic for a signal injected in the synthetic data as a function of the injected signal strength, with the 1 and $2\sigma$ containment intervals also indicated. Under the null hypothesis, the detection test statistic should follow a $\chi^2$-distribution; the median and $1\sigma$ and $2\sigma$ percentile values of the $\chi^2$-distribution are indicated by dashed grey lines. These results demonstrate that our detection test statistic follows its theoretically expected distribution under the null hypothesis ($\sin^2(2\theta_{\rm inj}) = 0$) and that our analysis framework can robustly identify a signal which is present in the data. The results are smoothed with a Savitzky–Golay filter for clarity.

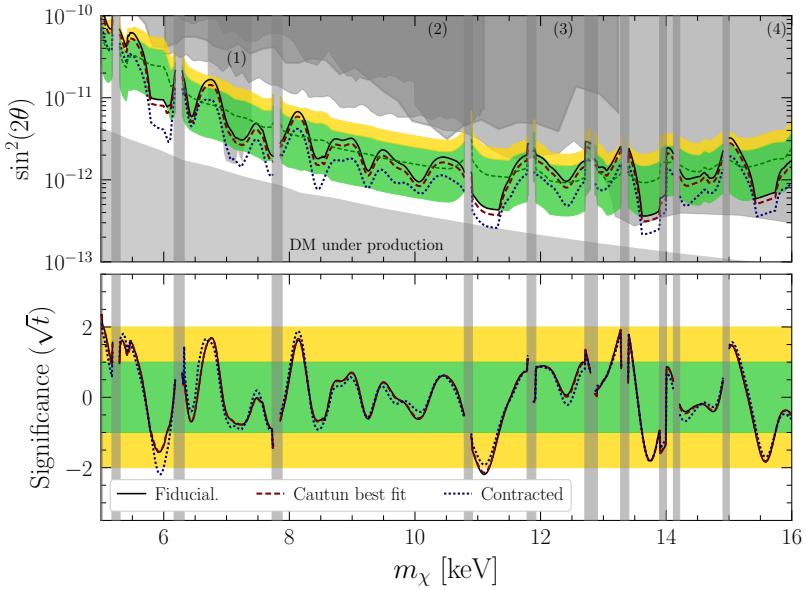



Figure H.20: As in Fig. 9.3, but for three different DM density profiles, all based upon Ref. [416]. In solid curve we show our fiducial results, corresponding to the uncontracted NFW profile with a conservative density. The dashed curve then shows our results using the best fit NFW profile, whereas in dashed we show the stronger limits that would be obtained with a contracted DM distribution. Details of the distributions are provided in the text.

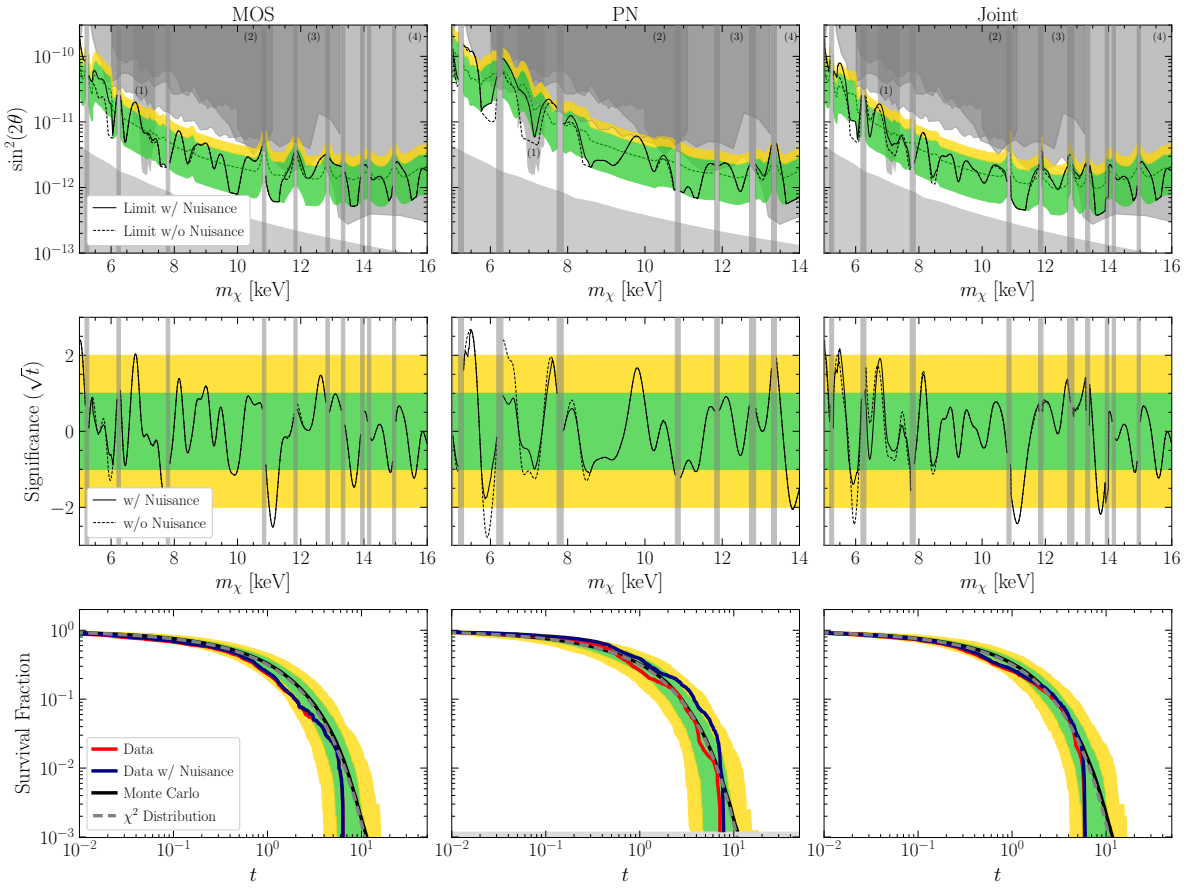



Figure H.21: The analogues of Figs. H.4 and H.5, but changing the kernel correlation length to $\sigma_E = 0.2$ (c.f. our fiducial value of $\sigma_E = 0.3$). The differences between the $\sigma_E = 0.2$ and 0.3 results are minor.

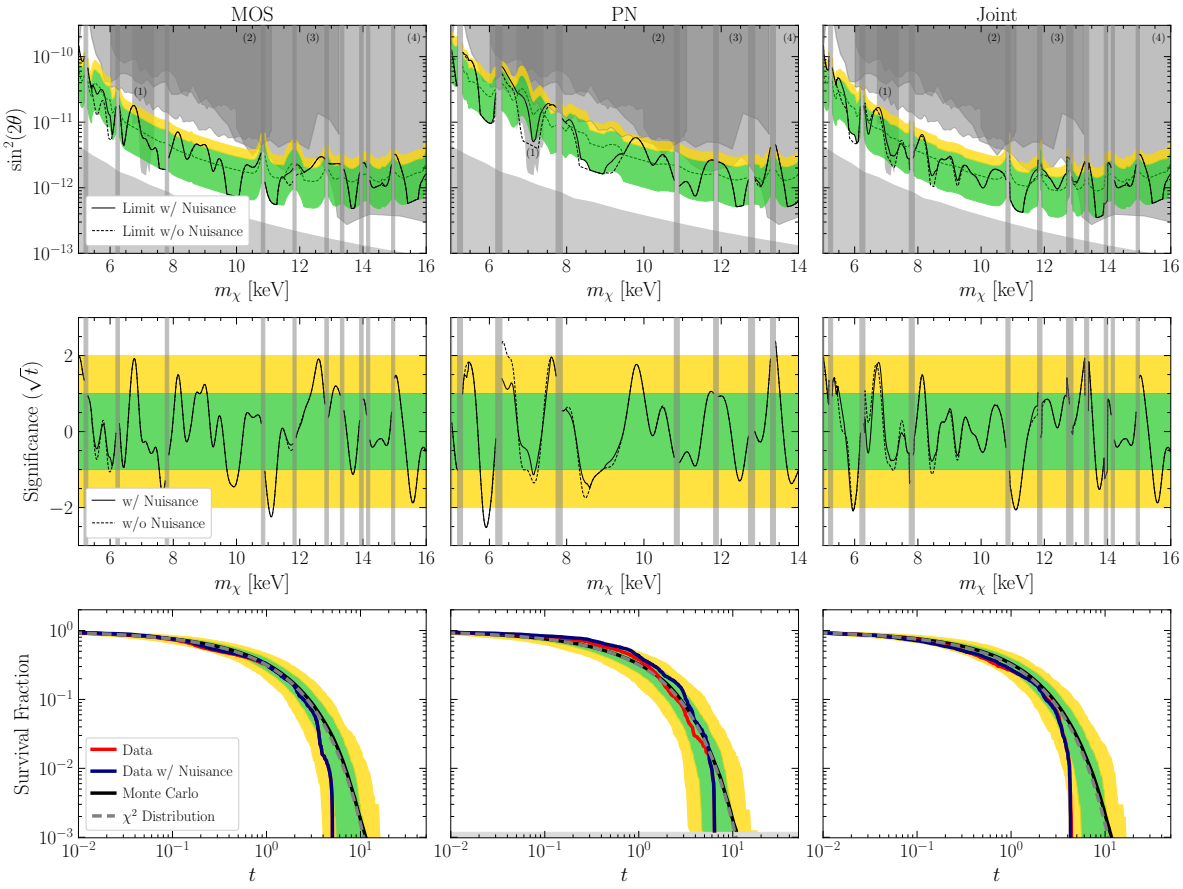



Figure H.22: As in Fig. H.21 but with $\sigma_E = 0.4$. The limit is slightly strengthened, although again the differences are not significant.

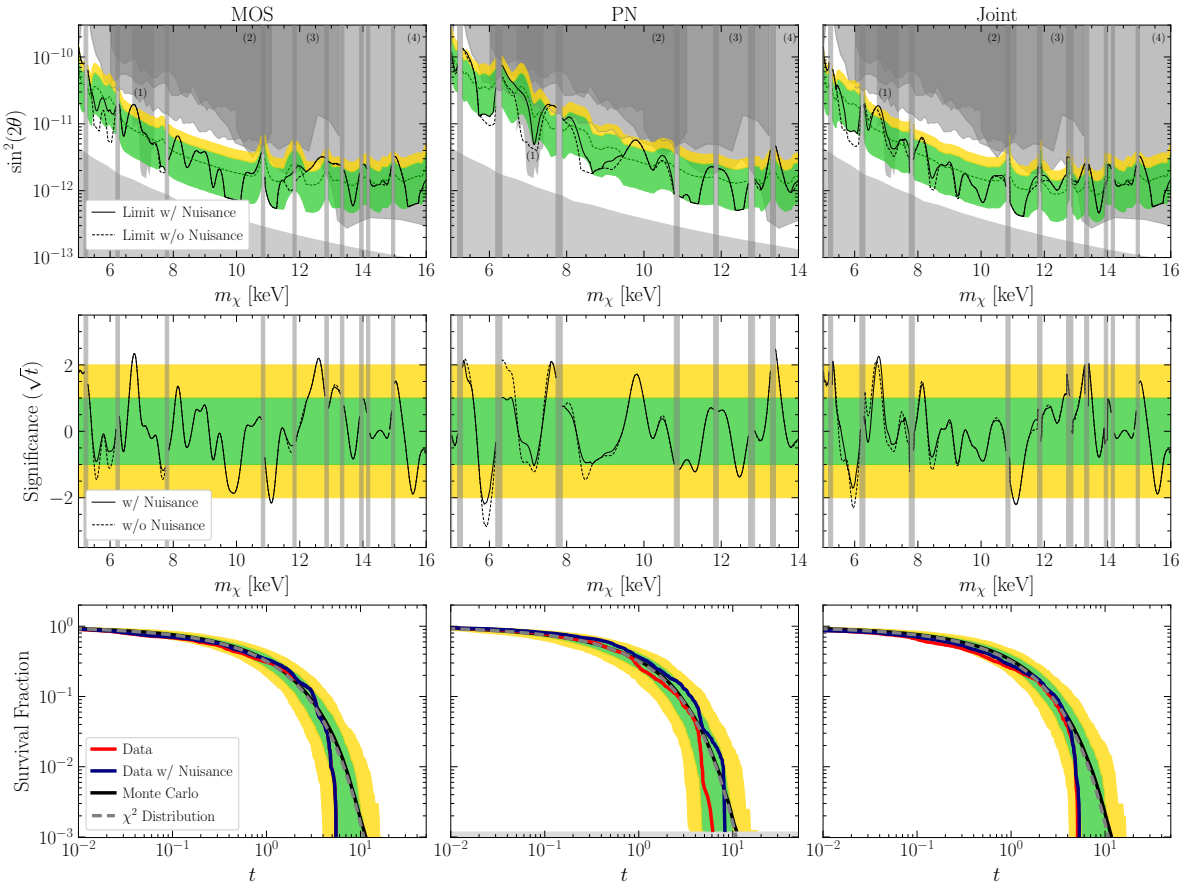



Figure H.23: As in Fig. H.21 but with $\sigma_E$ treated as a profiled nuisance parameter. The results demonstrate that even providing our background model this additional freedom does not have a significant impact on the limit.

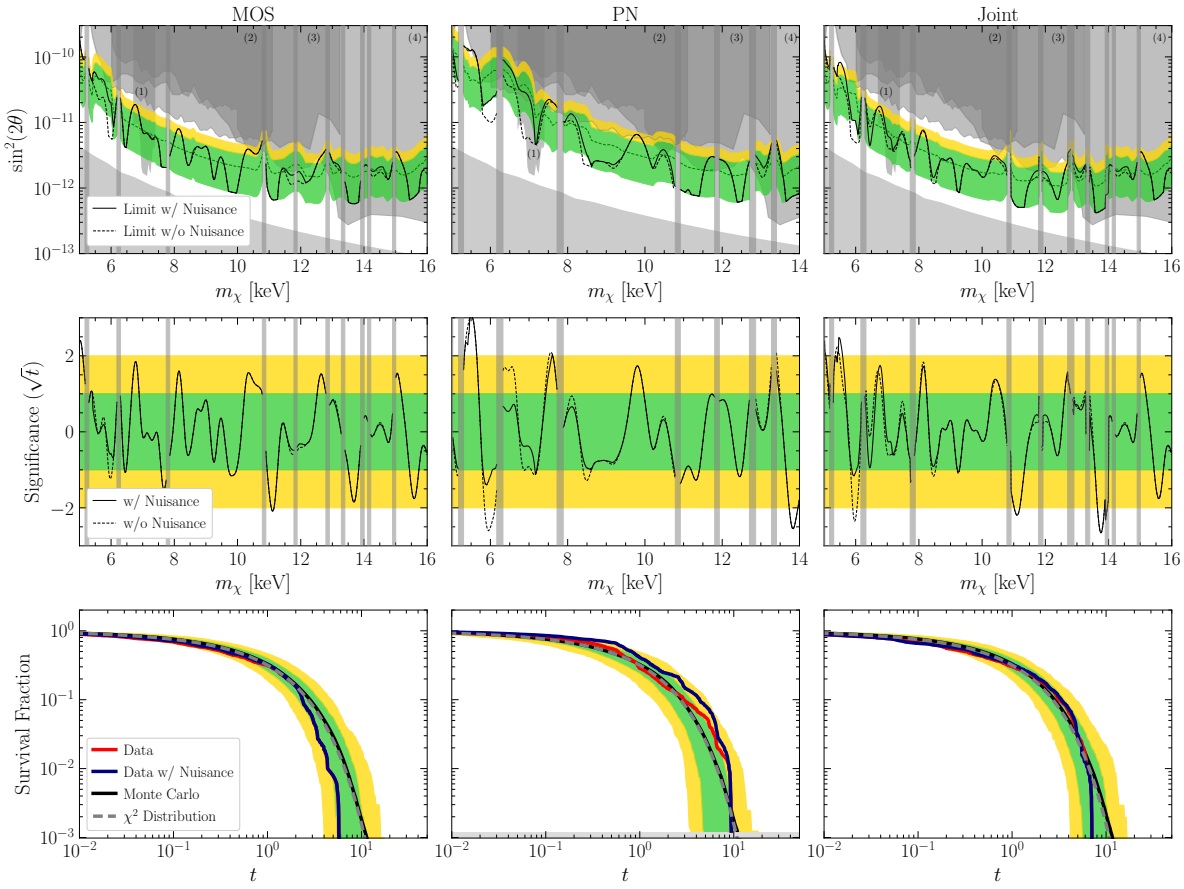



Figure H.24: As in Fig. H.21 but with the alternate GP kernel, in (H.10), with $\sigma^2 = 0.5$ keV$^2$.

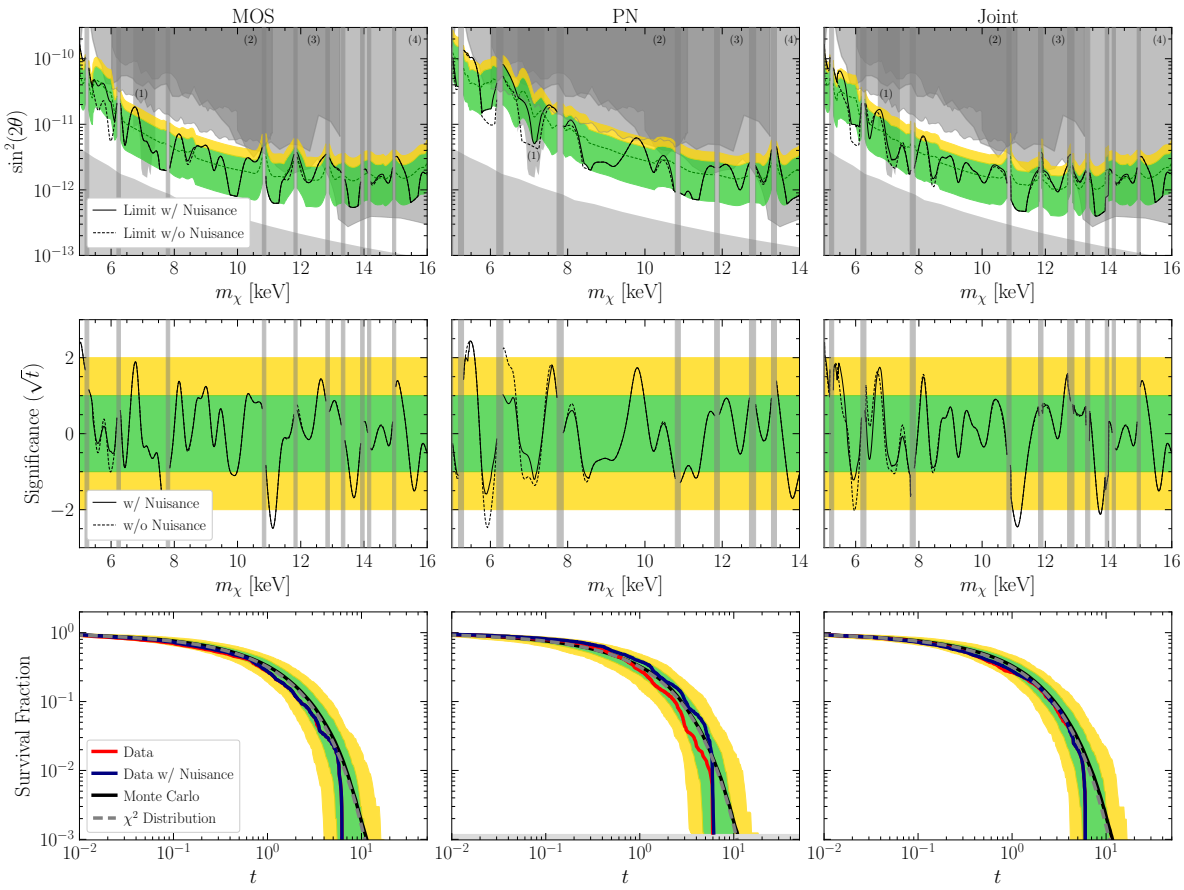



Figure H.25: As in Fig. H.24 but with $\sigma^2 = 1.0$ keV$^2$. Adopting a large scale length again slightly strengthens the limits, although again the systematic variation of our results with the kernel is relatively small.

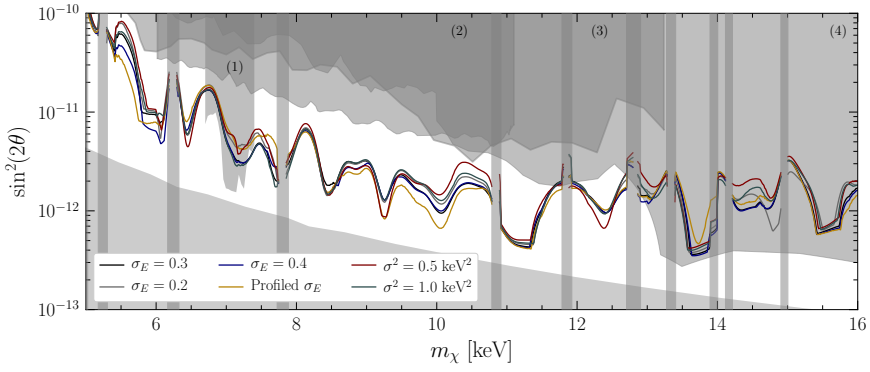



Figure H.26: A comparison of the limits obtained across the full mass range for each variation of the GP correlation-length hyperparameter considered. In particular we show results for variations of the relative-scale and fixed-scale kernels (denoted $\sigma_E$ and $\sigma^2$ respectively), as well as the relative-scale kernel where the scale profiled independently in each annulus.

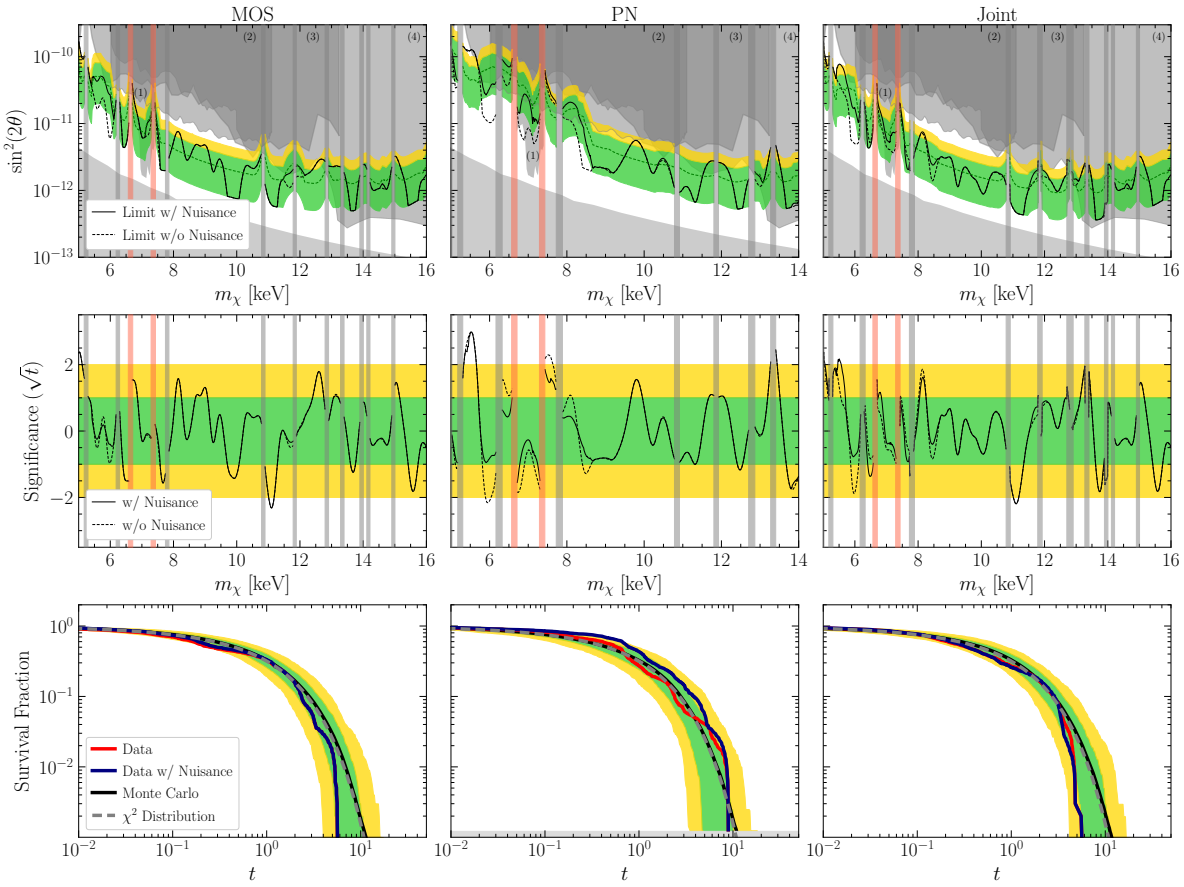



Figure H.27: As in Fig. H.21, but with the fiducial GP kernel at $\sigma_E = 0.3$ and the inclusion of 3.32 and 3.68 keV lines in all analyzed annuli. The newly masked region associated with these two lines is highlighted in light red.

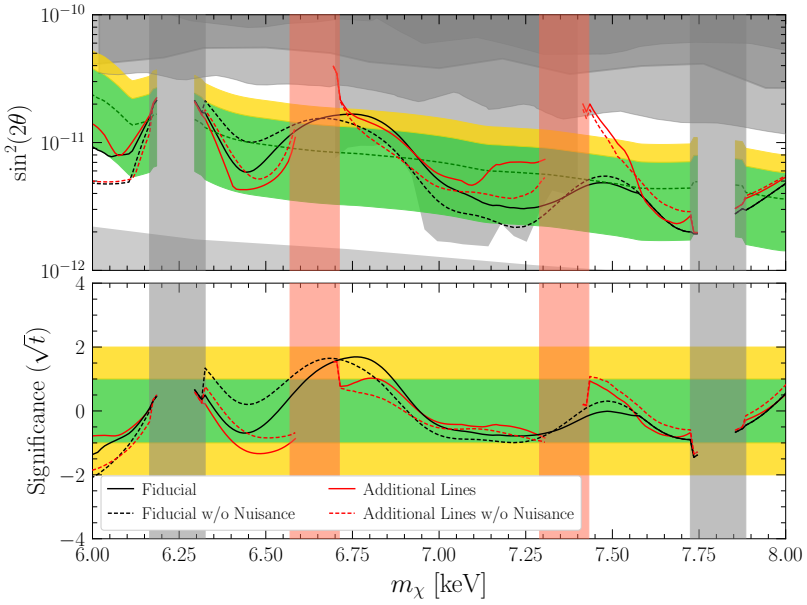



Figure H.28: A close inspection of the limits set in our fiducial analysis and the modified analysis that includes a 3.32 and 3.68 keV line in each annulus. We compare the limits set in these two analyses both with (solid lines) and without (dashed lines) the inclusion of our systematic nuisance parameter designed to test for and correct possible mismodeling.

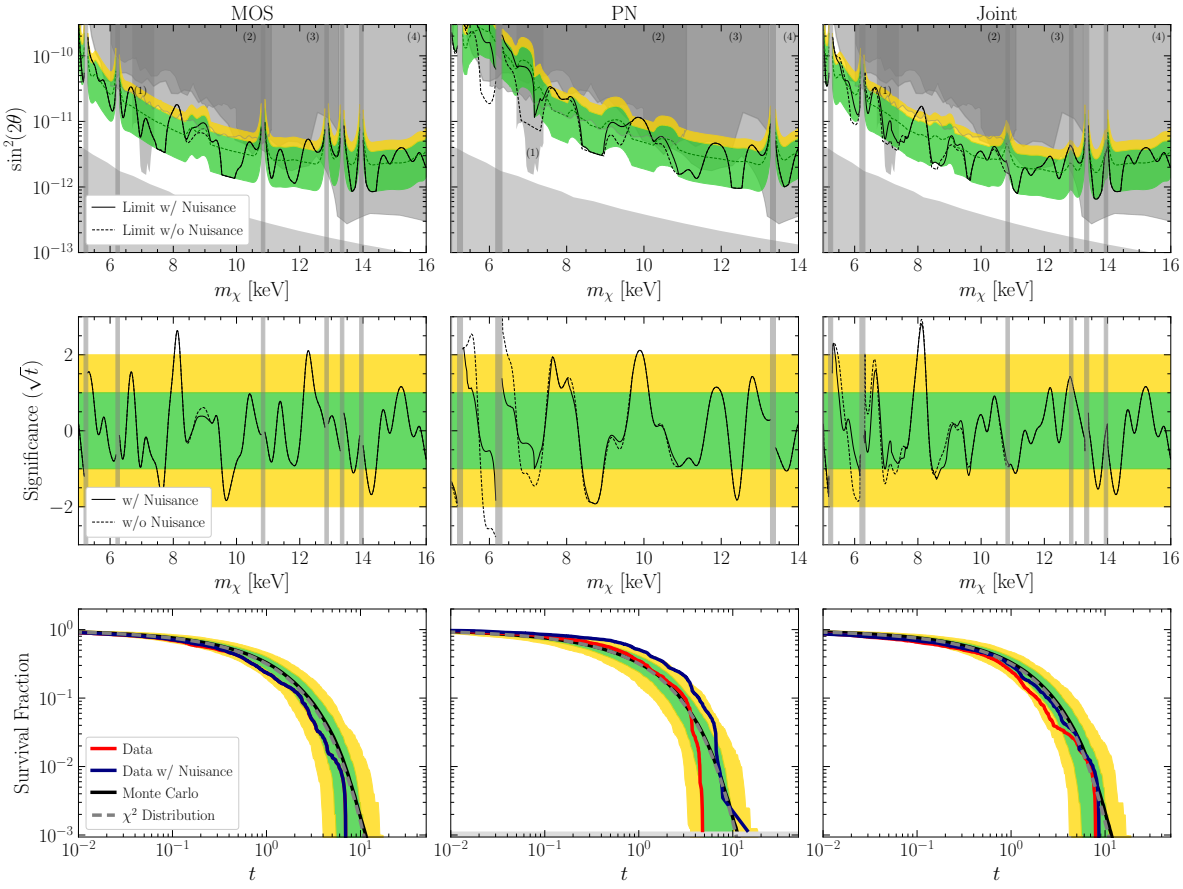



Figure H.29: The same results as presented in Figs. H.4 and H.5, however on a modified data set where instead of analyzing the signal ROI divided into eight individual rings, we stack the inner three rings into a single annulus. As in our primary approach, we subtract the background ROI flux from the signal-region data. The results are comparable to, although slightly weaker than, those from our fiducial approach, consistent with the reduced information available.

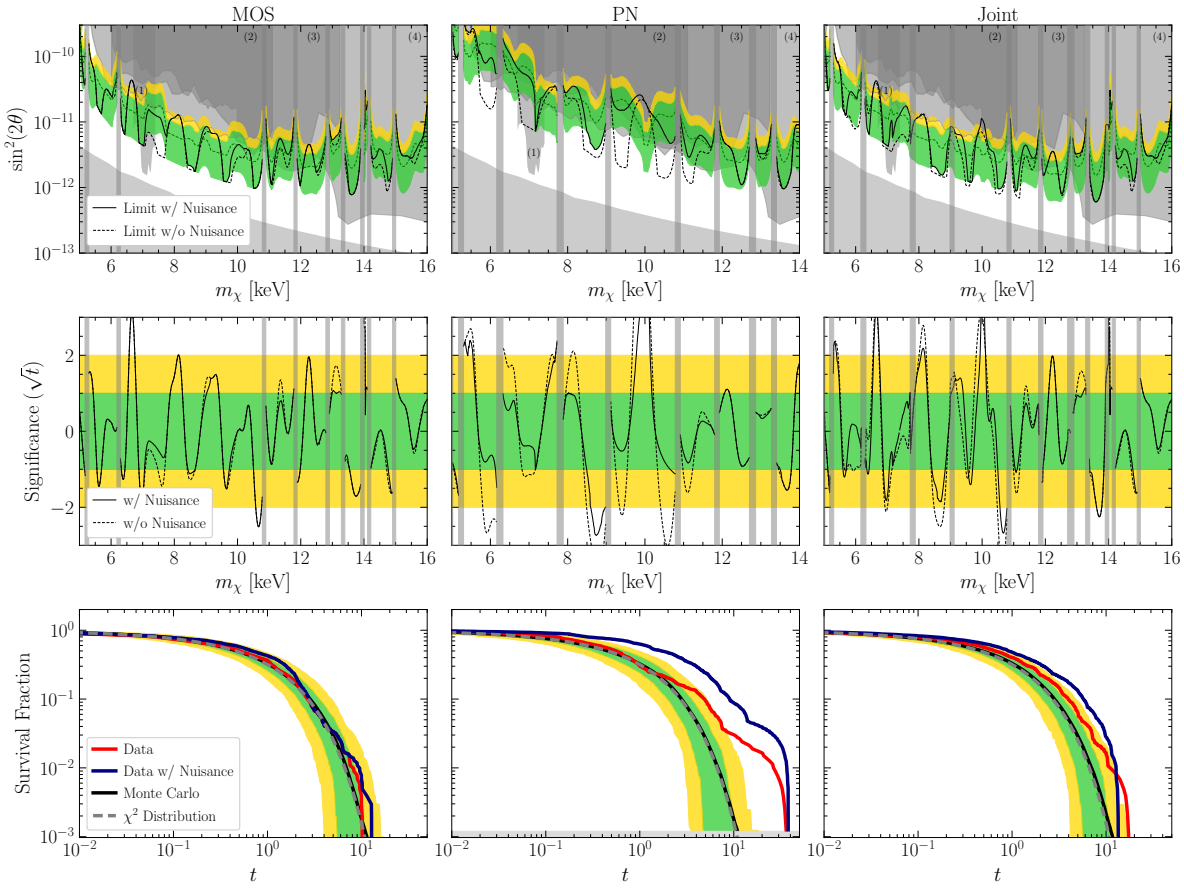



Figure H.30: As in Fig. H.29, however considering the stacked signal ROI without subtracting the background. The limit is noticeably worse, and several excesses appear, highlighting the importance of the background subtraction.

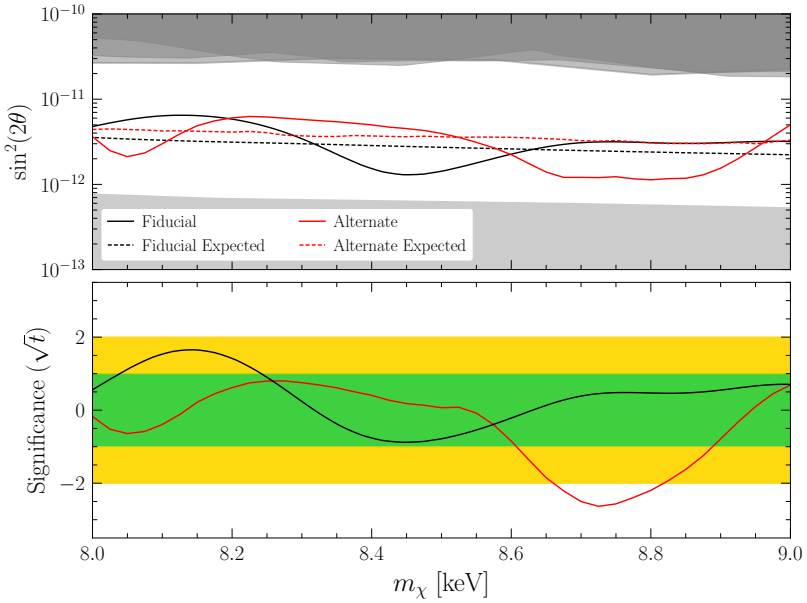



Figure H.31: Here we compare our fiducial results using a GP model, shown in black, to the result of an approach where the continuum background contribution is modeled with a second order polynomial, shown in red, as described in the text. Both results are shown without imposing a systematic nuisance parameter. While our fiducial approach uses the background-subtracted signal-ROI data, the alternate approach uses the un-subtracted data. We see that in both cases the expected and resulting limits are in qualitative agreement, demonstrating that our choice of GP modeling in our fiducial analysis does not drive the sensitivity of our results.

# BIBLIOGRAPHY


[1] Gianfranco Bertone and Dan Hooper. History of dark matter. *Rev. Mod. Phys.*, 90(4):045002, 2018.

[2] T. S. V. Albada, R. Sancisi, M. Petrou, and R. J. Tayler. Dark Matter in Spiral Galaxies [and Discussion]. *Phil. Trans. Roy. Soc. Lond. A*, 320(1556):447–464, 1986.

[3] Siddharth Mishra-Sharma. Extragalactic Searches for Dark Matter Annihilation. Other thesis, 9 2018.

[4] Y. Sofue, M. Honma, and T. Omodaka. Unified Rotation Curve of the Galaxy – Decomposition into de Vaucouleurs Bulge, Disk, Dark Halo, and the 9-kpc Rotation Dip –. *Publ. Astron. Soc. Jap.*, 61:227, 2009.

[5] Scott Dodelson. The Real Problem with MOND. *Int. J. Mod. Phys. D*, 20:2749–2753, 2011.

[6] M. Milgrom. A modification of the Newtonian dynamics as a possible alternative to the hidden mass hypothesis. *The Astrophysical Journal*, 270:365–370, July 1983.

[7] M. Milgrom. A modification of the Newtonian dynamics - Implications for galaxies. *The Astrophysical Journal*, 270:371–383, July 1983.

[8] M. Milgrom. A modification of the newtonian dynamics : implications for galaxy systems. *The Astrophysical Journal*, 270:384–389, July 1983.

[9] Mariangela Lisanti, Matthew Moschella, Nadav Joseph Outmezguine, and Oren Slone. Testing Dark Matter and Modifications to Gravity using Local Milky Way Observables. *Phys. Rev. D*, 100(8):083009, 2019.

[10] Constantinos Skordis, D. F. Mota, P. G. Ferreira, and C. Boehm. Large Scale Structure in Bekenstein's theory of relativistic Modified Newtonian Dynamics. *Phys. Rev. Lett.*, 96:011301, 2006.

[11] Constantinos Skordis. TOPICAL REVIEW: The tensor-vector-scalar theory and its cosmology. *Classical and Quantum Gravity*, 26(14):143001, July 2009.

[12] Benoit Famaey and Stacy McGaugh. Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions. *Living Rev. Rel.*, 15:10, 2012.

[13] Douglas Clowe, Marusa Bradac, Anthony H. Gonzalez, Maxim Markevitch, Scott W. Randall, Christine Jones, and Dennis Zaritsky. A direct empirical proof of the existence of dark matter. *Astrophys. J. Lett.*, 648:L109–L113, 2006.

[14] Marusa Bradac, Steven W. Allen, Tommaso Treu, Harald Ebeling, Richard Massey, R. Glenn Morris, Anja von der Linden, and Douglas Applegate. Revealing the properties of dark matter in the merging cluster MACSJ0025.4-1222. *Astrophys. J.*, 687:959, 2008.

[15] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020.

[16] Julien Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. 4 2011.

[17] Edward W. Kolb and Michael S. Turner. The Early Universe. *Front. Phys.*, 69:1–547, 1990.

[18] Scott Dodelson. *Modern Cosmology*. Academic Press, Amsterdam, 2003.

[19] Wayne Hu. Lecture Notes on CMB Theory: From Nucleosynthesis to Recombination. 2 2008.

[20] T. M. C. Abbott et al. Dark Energy Survey Year 3 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing. 5 2021.

[21] L. V. E. Koopmans, M. Barnabe, A. Bolton, M. Bradac, L. Ciotti, A. Congdon, O. Czoske, S. Dye, A. Dutton, A. Elliasdottir, E. Evans, C. D. Fassnacht, N. Jackson, C. Keeton, J. Lasio, L. Moustakas, M. Meneghetti, S. Myers, C. Nipoti, S. Suyu, G. van de Ven, S. Vegetti, O. Wucknitz, and H. S. Zhao. Strong Gravitational Lensing as a Probe of Gravity, Dark-Matter and Super-Massive Black Holes. In *astro2010: The Astronomy and Astrophysics Decadal Survey*, volume 2010, page 159, January 2009.

[22] Ethan O. Nadler, Simon Birrer, Daniel Gilman, Risa H. Wechsler, Xiaolong Du, Andrew Benson, Anna M. Nierenberg, and Tommaso Treu. Dark Matter Constraints from a Unified Analysis of Strong Gravitational Lenses and Milky Way Satellite Galaxies. 1 2021.

[23] Will J. Percival et al. The shape of the SDSS DR5 galaxy power spectrum. *Astrophys. J.*, 657:645–663, 2007.

[24] Shadab Alam et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Mon. Not. Roy. Astron. Soc.*, 470(3):2617–2652, 2017.

[25] K. Bechtol et al. Eight New Milky Way Companions Discovered in First-year Dark Energy Survey Data. *The Astrophysical Journal*, 807(1):50, July 2015.

[26] A. Drlica-Wagner et al. Eight Ultra-faint Galaxy Candidates Discovered in Year Two of the Dark Energy Survey. *Astrophys. J.*, 813(2):109, 2015.

[27] Sergey E. Koposov, Vasily Belokurov, Gabriel Torrealba, and N. Wyn Evans. Beasts of the Southern Wild: Discovery of nine Ultra Faint satellites in the vicinity of the Magellanic Clouds. *Astrophys. J.*, 805(2):130, 2015.

[28] J. I. Read. The Local Dark Matter Density. *J. Phys.*, G41:063101, 2014.

[29] Leszek Roszkowski, Enrico Maria Sessolo, and Sebastian Trojanowski. WIMP dark matter candidates and searches—current status and future prospects. *Rept. Prog. Phys.*, 81(6):066201, 2018.

[30] Francesca Chadha-Day, John Ellis, and David J. E. Marsh. Axion Dark Matter: What is it and Why Now? 5 2021.

[31] Kevork N. Abazajian. Sterile neutrinos in cosmology. 2017.

[32] Marco Fabbrichesi, Emidio Gabrielli, and Gaia Lanfranchi. The Dark Photon. 5 2020.

[33] Anson Hook. TASI Lectures on the Strong CP Problem and Axions. *PoS*, TASI2018:004, 2019.

[34] C. Abel et al. Measurement of the permanent electric dipole moment of the neutron. *Phys. Rev. Lett.*, 124:081803, Feb 2020.

[35] R. J. Crewther, P. Di Vecchia, G. Veneziano, and Edward Witten. Chiral Estimate of the Electric Dipole Moment of the Neutron in Quantum Chromodynamics. *Phys. Lett. B*, 88:123, 1979. [Erratum: Phys.Lett.B 91, 487 (1980)].

[36] Michael Dine, Laurel Stephenson Haskins, Lorenzo Ubaldi, and Di Xu. Some Remarks on Anthropic Approaches to the Strong CP Problem. *JHEP*, 05:171, 2018.

[37] S. Borsanyi, M. Dierigl, Z. Fodor, S. D. Katz, S. W. Mages, D. Nogradi, J. Redondo, A. Ringwald, and K. K. Szabo. Axion cosmology, lattice QCD and the dilute instanton gas. *Phys. Lett. B*, 752:175–181, 2016.

[38] R. D. Peccei and Helen R. Quinn. CP Conservation in the Presence of Instantons. *Phys. Rev. Lett.*, 38:1440–1443, 1977.

[39] R. D. Peccei and Helen R. Quinn. Constraints Imposed by CP Conservation in the Presence of Instantons. *Phys. Rev.*, D16:1791–1797, 1977.

[40] Jihn E. Kim. Weak Interaction Singlet and Strong CP Invariance. *Phys. Rev. Lett.*, 43:103, 1979.

[41] Michael Dine, Willy Fischler, and Mark Srednicki. A Simple Solution to the Strong CP Problem with a Harmless Axion. *Phys. Lett.*, B104:199, 1981.

[42] A. R. Zhitnitsky. On Possible Suppression of the Axion Hadron Interactions. (In Russian). *Sov. J. Nucl. Phys.*, 31:260, 1980. [Yad. Fiz.31,497(1980)].

[43] Steven Weinberg. A New Light Boson? *Phys.Rev.Lett.*, 40:223–226, 1978.

[44] Frank Wilczek. Problem of Strong P and T Invariance in the Presence of Instantons. *Phys.Rev.Lett.*, 40:279–282, 1978.

[45] L. F. Abbott and P. Sikivie. A Cosmological Bound on the Invisible Axion. *Phys. Lett.*, B120:133–136, 1983.

[46] John Preskill, Mark B. Wise, and Frank Wilczek. Cosmology of the Invisible Axion. *Phys. Lett.*, B120:127–132, 1983.

[47] Ann E. Nelson. Naturally Weak CP Violation. *Phys. Lett. B*, 136:387–391, 1984.

[48] S. M. Barr. Solving the strong CP problem without the peccei-quinn symmetry. *Phys. Rev. Lett.*, 53:329–332, Jul 1984.

[49] Daniel R. Nelson, George Tamminga Fleming, and Gregory W. Kilcup. Is strong CP due to a massless up quark? *Phys. Rev. Lett.*, 90:021601, 2003.

[50] Anson Hook. Anomalous solutions to the strong CP problem. *Phys. Rev. Lett.*, 114(14):141801, 2015.

[51] Peter Svrcek and Edward Witten. Axions In String Theory. *JHEP*, 06:051, 2006.

[52] Asimina Arvanitaki, Savas Dimopoulos, Sergei Dubovsky, Nemanja Kaloper, and John March-Russell. String Axiverse. *Phys. Rev.*, D81:123530, 2010.

[53] Anne Ernst, Luca Di Luzio, Andreas Ringwald, and Carlos Tamarit. Axion properties in GUTs. *PoS*, CORFU2018:054, 2019.

[54] Raymond T. Co, Lawrence J. Hall, and Keisuke Harigaya. Predictions for Axion Couplings from ALP Cogenesis. *JHEP*, 01:172, 2021.

[55] Peter W. Graham, Igor G. Irastorza, Steven K. Lamoreaux, Axel Lindner, and Karl A. van Bibber. Experimental Searches for the Axion and Axion-Like Particles. *Ann. Rev. Nucl. Part. Sci.*, 65:485–514, 2015.

[56] Igor G. Irastorza and Javier Redondo. New experimental approaches in the search for axion-like particles. *Prog. Part. Nucl. Phys.*, 102:89–159, 2018.

[57] M. Tanabashi et al. Review of Particle Physics. *Phys. Rev. D*, 98(3):030001, 2018.

[58] Dmitry Budker, Peter W. Graham, Micah Ledbetter, Surjeet Rajendran, and Alex Sushkov. Proposal for a Cosmic Axion Spin Precession Experiment (CASPEr). *Phys. Rev.*, X4(2):021030, 2014.

[59] Jonathan Ouellet and Zachary Bogorad. Solutions to Axion Electrodynamics in Various Geometries. *Phys. Rev. D*, 99(5):055010, 2019.

[60] Jeff A. Dror, Hitoshi Murayama, and Nicholas L. Rodd. Cosmic axion background. *Phys. Rev. D*, 103(11):115004, 2021.

[61] P. Sikivie. Experimental tests of the "invisible" axion. *Phys. Rev. Lett.*, 51:1415–1417, Oct 1983.

[62] Yuri Nickolaevich Gnedin and S. V. Krasnikov. Polariometric effects associated with the detection of Goldstone bosons in stars and galaxies. *Sov. Phys. JETP*, 75:933–937, 1992.

[63] Yu N. Gnedin. Astrophysical consequences of the existence of goldstone bosons. *Astron. Astrophys. Trans.*, 5(1-4):163–175, 1994.

[64] Stephen L. Adler, J. Gamboa, F. Mendez, and J. Lopez-Sarrion. Axions and 'Light Shining Through a Wall': A Detailed Theoretical Analysis. *Annals Phys.*, 323:2851–2872, 2008.

[65] Katelin Schutz, Tongyan Lin, Benjamin R. Safdi, and Chih-Liang Wu. Constraining a Thin Dark Matter Disk with Gaia. *Phys. Rev. Lett.*, 121(8):081101, 2018.

[66] Malte Buschmann, Benjamin R. Safdi, and Katelin Schutz. The Galactic potential and dark matter density from angular stellar accelerations. 3 2021.

[67] N. Du et al. A Search for Invisible Axion Dark Matter with the Axion Dark Matter Experiment. *Phys. Rev. Lett.*, 120(15):151301, 2018.

[68] K. M. Backes et al. A quantum-enhanced search for dark matter axions. *Nature*, 590(7845):238–242, 2021.

[69] J. Choi, S. Ahn, B. R. Ko, S. Lee, and Y. K. Semertzidis. CAPP-8TB: Axion Dark Matter Search Experiment around 6.7 $\mu$eV. 7 2020.

[70] Allen Caldwell, Gia Dvali, Béla Majorovits, Alexander Millar, Georg Raffelt, Javier Redondo, Olaf Reimann, Frank Simon, and Frank Steffen. Dielectric Haloscopes: A New Way to Detect Axion Dark Matter. *Phys. Rev. Lett.*, 118(9):091801, 2017.

[71] Yonatan Kahn, Benjamin R. Safdi, and Jesse Thaler. Broadband and Resonant Approaches to Axion Dark Matter Detection. *Phys. Rev. Lett.*, 117(14):141801, 2016.

[72] Jonathan L. Ouellet et al. First Results from ABRACADABRA-10 cm: A Search for Sub-$\mu$eV Axion Dark Matter. *Phys. Rev. Lett.*, 122(12):121802, 2019.

[73] Chiara P. Salemi et al. The search for low-mass axion dark matter with ABRACADABRA-10cm. 2 2021.

[74] Alexander V. Gramolin, Deniz Aybas, Dorian Johnson, Janos Adam, and Alexander O. Sushkov. Sensitivity enhancement for a light axion dark matter search with magnetic material. 11 2018.

[75] Alexander V. Gramolin, Deniz Aybas, Dorian Johnson, Janos Adam, and Alexander O. Sushkov. Search for axion-like dark matter with ferromagnets. 3 2020.

[76] Maximiliano Silva-Feaver et al. Design Overview of the DM Radio Pathfinder Experiment. *IEEE Trans. Appl. Supercond.*, 27(4):1400204, 2016.

[77] Luc Demortier and Louis Lyons. Testing Hypotheses in Particle Physics: Plots of $p_0$ Versus $p_1$. *arXiv e-prints*, page arXiv:1408.6123, August 2014.

[78] S. S. Wilks. The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses. *Annals Math. Statist.*, 9(1):60–62, 1938.

[79] Malte Buschmann, Joshua W. Foster, and Benjamin R. Safdi. Early-Universe Simulations of the Cosmological Axion. *Phys. Rev. Lett.*, 124(16):161103, 2020.

[80] Joshua W. Foster, Nicholas L. Rodd, and Benjamin R. Safdi. Revealing the Dark Matter Halo with Axion Direct Detection. *Phys. Rev. D*, 97(12):123006, 2018.

[81] Joshua W. Foster, Yonatan Kahn, Rachel Nguyen, Nicholas L. Rodd, and Benjamin R. Safdi. Dark Matter Interferometry. *Phys. Rev. D*, 103(7):076018, 2021.

[82] Jonathan L. Ouellet et al. Design and implementation of the ABRACADABRA-10 cm axion dark matter search. *Phys. Rev. D*, 99(5):052012, 2019.

[83] Joshua W. Foster, Yonatan Kahn, Oscar Macias, Zhiquan Sun, Ralph P. Eatough, Vladislav I. Kondratiev, Wendy M. Peters, Christoph Weniger, and Benjamin R. Safdi. Green Bank and Effelsberg Radio Telescope Searches for Axion Dark Matter Conversion in Neutron Star Magnetospheres. *Phys. Rev. Lett.*, 125(17):171301, 2020.

[84] Christopher Dessert, Joshua W. Foster, and Benjamin R. Safdi. X-ray Searches for Axions from Super Star Clusters. *Phys. Rev. Lett.*, 125(26):261102, 2020.

[85] Joshua W. Foster, Marius Kongsore, Christopher Dessert, Yujin Park, Nicholas L. Rodd, Kyle Cranmer, and Benjamin R. Safdi. Deep search for decaying dark matter with xmm-newton blank-sky observations. *Phys. Rev. Lett.*, 127:051101, Jul 2021.

[86] Michael Dine and Willy Fischler. The Not So Harmless Axion. *Phys. Lett.*, B120:137–141, 1983.

[87] Mikhail A. Shifman, A. I. Vainshtein, and Valentin I. Zakharov. Can Confinement Ensure Natural CP Invariance of Strong Interactions? *Nucl. Phys.*, B166:493, 1980.

[88] Mark Srednicki. Axion Couplings to Matter. 1. CP Conserving Parts. *Nucl. Phys.*, B260:689–700, 1985.

[89] David J. E. Marsh. Axion Cosmology. *Phys. Rept.*, 643:1–79, 2016.

[90] C. J. Hogan and M. J. Rees. AXION MINICLUSTERS. *Phys. Lett.*, B205:228–230, 1988.

[91] Edward W. Kolb and Igor I. Tkachev. Axion miniclusters and Bose stars. *Phys. Rev. Lett.*, 71:3051–3054, 1993.

[92] Edward W. Kolb and Igor I. Tkachev. Nonlinear axion dynamics and formation of cosmological pseudosolitons. *Phys. Rev.*, D49:5040–5051, 1994.

[93] Edward W. Kolb and Igor I. Tkachev. Large amplitude isothermal fluctuations and high density dark matter clumps. *Phys. Rev.*, D50:769–773, 1994.

[94] Kathryn M. Zurek, Craig J. Hogan, and Thomas R. Quinn. Astrophysical Effects of Scalar Dark Matter Miniclusters. *Phys. Rev.*, D75:043511, 2007.

[95] Jonas Enander, Andreas Pargner, and Thomas Schwetz. Axion minicluster power spectrum and mass function. *JCAP*, 1712(12):038, 2017.

[96] Alejandro Vaquero, Javier Redondo, and Julia Stadler. Early seeds of axion miniclusters. *JCAP*, 04:012, 2019.

[97] Peter Tinyakov, Igor Tkachev, and Konstantin Zioutas. Tidal streams from axion miniclusters and direct axion searches. *JCAP*, 1601(01):035, 2016.

[98] Sacha Davidson and Thomas Schwetz. Rotating Drops of Axion Dark Matter. *Phys. Rev.*, D93(12):123509, 2016.

[99] Malcolm Fairbairn, David J. E. Marsh, Jérémie Quevillon, and Simon Rozier. Structure formation and microlensing with axion miniclusters. *Phys. Rev.*, D97(8):083502, 2018.

[100] Jeff A. Dror, Harikrishnan Ramani, Tanner Trickle, and Kathryn M. Zurek. Pulsar Timing Probes of Primordial Black Holes and Subhalos. *Phys. Rev. D*, 100(2):023003, 2019.

[101] M. S. Pshirkov and S. B. Popov. Conversion of Dark matter axions to photons in magnetospheres of neutron stars. *J. Exp. Theor. Phys.*, 108:384–388, 2009.

[102] Fa Peng Huang, Kenji Kadota, Toyokazu Sekiguchi, and Hiroyuki Tashiro. Radio telescope search for the resonant conversion of cold dark matter axions from the magnetized astrophysical sources. *Phys. Rev.*, D97(12):123001, 2018.

[103] Anson Hook, Yonatan Kahn, Benjamin R. Safdi, and Zhiquan Sun. Radio Signals from Axion Dark Matter Conversion in Neutron Star Magnetospheres. *Phys. Rev. Lett.*, 121(24):241102, 2018.

[104] Benjamin R. Safdi, Zhiquan Sun, and Alexander Y. Chen. Detecting Axion Dark Matter with Radio Lines from Neutron Star Populations. *Phys. Rev. D*, 99(12):123021, 2019.

[105] Yang Bai and Yuta Hamada. Detecting Axion Stars with Radio Telescopes. *Phys. Lett.*, B781:187–194, 2018.

[106] Malte Buschmann, Joshua W. Foster, Bejamin R. Safdi, and Kathryn M. Zurek. Radio bursts from axion minihalo encounters with neutron stars. To appear.

[107] T. M. Shokair et al. Future Directions in the Microwave Cavity Search for Dark Matter Axions. *Int. J. Mod. Phys.*, A29:1443004, 2014.

[108] B. M. Brubaker et al. First results from a microwave cavity axion search at 24 $\mu$eV. *Phys. Rev. Lett.*, 118(6):061302, 2017.

[109] S. Al Kenany et al. Design and operational experience of a microwave cavity axion detector for the 20–100 $\mu$eV range. *Nucl. Instrum. Meth.*, A854:11–24, 2017.

[110] B. M. Brubaker, L. Zhong, S. K. Lamoreaux, K. W. Lehnert, and K. A. van Bibber. HAYSTAC axion search analysis procedure. *Phys. Rev. D*, 96(12):123008, 2017.

[111] Saptarshi Chaudhuri, Peter W. Graham, Kent Irwin, Jeremy Mardon, Surjeet Rajendran, and Yue Zhao. Radio for hidden-photon dark matter detection. *Phys. Rev.*, D92(7):075012, 2015.

[112] Vincent B.. Klaer and Guy D.. Moore. The dark-matter axion mass. *JCAP*, 1711(11):049, 2017.

[113] Richard Lynn Davis. Cosmic Axions from Cosmic Strings. *Phys. Lett.*, B180:225–230, 1986.

[114] R. L. Davis and E. P. S. Shellard. DO AXIONS NEED INFLATION? *Nucl. Phys.*, B324:167–186, 1989.

[115] R. A. Battye and E. P. S. Shellard. Axion string constraints. *Phys. Rev. Lett.*, 73:2954–2957, 1994. [Erratum: Phys. Rev. Lett.76,2203(1996)].

[116] Olivier Wantz and E. P. S. Shellard. Axion Cosmology Revisited. *Phys. Rev.*, D82:123508, 2010.

[117] Takashi Hiramatsu, Masahiro Kawasaki, Toyokazu Sekiguchi, Masahide Yamaguchi, and Jun'ichi Yokoyama. Improved estimation of radiated axions from cosmological axionic strings. *Phys. Rev.*, D83:123531, 2011.

[118] Masahiro Kawasaki, Ken'ichi Saikawa, and Toyokazu Sekiguchi. Axion dark matter from topological defects. *Phys. Rev.*, D91(6):065014, 2015.

[119] Guillermo Ballesteros, Javier Redondo, Andreas Ringwald, and Carlos Tamarit. Standard Model—axion—seesaw—Higgs portal inflation. Five problems of particle physics and cosmology solved in one stroke. *JCAP*, 1708(08):001, 2017.

[120] Luca Visinelli, Sebastian Baum, Javier Redondo, Katherine Freese, and Frank Wilczek. Dilute and dense axion stars. *Phys. Lett.*, B777:64–72, 2018.

[121] Takashi Hiramatsu, Masahiro Kawasaki, Ken'ichi Saikawa, and Toyokazu Sekiguchi. Production of dark matter axions from collapse of string-wall systems. *Phys. Rev.*, D85:105020, 2012. [Erratum: Phys. Rev.D86,089902(2012)].

[122] Giovanni Grilli di Cortona, Edward Hardy, Javier Pardo Vega, and Giovanni Villadoro. The QCD axion, precisely. *JHEP*, 01:034, 2016.

[123] Sz. Borsanyi et al. Calculation of the axion mass based on high-temperature lattice quantum chromodynamics. *Nature*, 539(7627):69–71, 2016.

[124] Marco Gorghetto, Edward Hardy, and Giovanni Villadoro. Axions from Strings: the Attractive Solution. *JHEP*, 07:151, 2018.

[125] Leesa Fleury and Guy D. Moore. Axion dark matter: strings and their cores. *JCAP*, 1601:004, 2016.

[126] Mustafa A. Amin and David Shirokoff. Flat-top oscillons in an expanding universe. *Phys. Rev.*, D81:085045, 2010.

[127] Malte Buschmann, Joshua W. Foster, and Benjamin R. Safdi. Early-Universe Simulations of the Cosmological Axion: Supplementary Data, June 2019.

[128] N. Aghanim et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.*, 641:A6, 2020.

[129] Anne M Green. Astrophysical uncertainties on the local dark matter distribution and direct detection experiments. *J. Phys.*, G44(8):084001, 2017.

[130] Stephen J. Asztalos et al. Large scale microwave cavity search for dark matter axions. *Phys. Rev.*, D64:092003, 2001.

[131] S. J. Asztalos, G. Carosi, C. Hagmann, D. Kinion, K. van Bibber, M. Hotz, L. J Rosenberg, G. Rybka, J. Hoskins, J. Hwang, P. Sikivie, D. B. Tanner, R. Bradley, and J. Clarke. SQUID-Based Microwave Cavity Search for Dark-Matter Axions. *Phys. Rev. Lett.*, 104:041301, Jan 2010.

[132] Woohyun Chung. CULTASK, The Coldest Axion Experiment at CAPP/IBS in Korea. *PoS*, CORFU2015:047, 2016.

[133] Béla Majorovits and Javier Redondo. MADMAX: A new Dark Matter Axion Search using a Dielectric Haloscope. In *12th Patras Workshop on Axions, WIMPs and WISPs (AXION-WIMP 2016) Jeju Island, South Korea, June 20-24, 2016*, 2016.

[134] Alexander J. Millar, Georg G. Raffelt, Javier Redondo, and Frank D. Steffen. Dielectric Haloscopes to Search for Axion Dark Matter: Theoretical Foundations. *JCAP*, 1701(01):061, 2017.

[135] Ara N. Ioannisian, Narine Kazarian, Alexander J. Millar, and Georg G. Raffelt. Axion-photon conversion caused by dielectric interfaces: quantum field calculation. *JCAP*, 1709(09):005, 2017.

[136] Y. V. Stadnik and V. V. Flambaum. Axion-induced effects in atoms, molecules, and nuclei: Parity nonconservation, anapole moments, electric dipole moments, and spin-gravity and spin-axion momentum couplings. *Phys. Rev.*, D89(4):043522, 2014.

[137] C. Abel et al. Search for Axionlike Dark Matter through Nuclear Spin Precession in Electric and Magnetic Fields. *Phys. Rev.*, X7(4):041034, 2017.

[138] Ben T. McAllister, Graeme Flower, Justin Kruger, Eugene N. Ivanov, Maxim Goryachev, Jeremy Bourhill, and Michael E. Tobar. The ORGAN Experiment: An axion haloscope above 15 GHz. *Phys. Dark Univ.*, 18:67–72, 2017.

[139] Giuseppe Ruoso, Augusto Lombardi, Antonello Ortolan, Ruggero Pengo, Caterina Braggio, Giovanni Carugno, Carmelo Sebastiano Gallo, and Clive C. Speake. The QUAX proposal: a search of galactic axion with magnetic materials. *J. Phys. Conf. Ser.*, 718(4):042051, 2016.

[140] R. Barbieri, C. Braggio, G. Carugno, C. S. Gallo, A. Lombardi, A. Ortolan, R. Pengo, G. Ruoso, and C. C. Speake. Searching for galactic axions through magnetized media: the QUAX proposal. *Phys. Dark Univ.*, 15:135–141, 2017.

[141] Nicolò Crescini, Caterina Braggio, Giovanni Carugno, Paolo Falferi, Antonello Ortolan, and Giuseppe Ruoso. The QUAX-$g_p$ $g_s$ experiment to search for monopole-dipole Axion interaction. *Nucl. Instrum. Meth.*, A842:109–113, 2017.

[142] V. Anastassopoulos et al. Towards a medium-scale axion helioscope and haloscope. *JINST*, 12(11):P11019, 2017.

[143] Oliver K. Baker, Michael Betz, Fritz Caspers, Joerg Jaeckel, Axel Lindner, Andreas Ringwald, Yannis Semertzidis, Pierre Sikivie, and Konstantin Zioutas. Prospects for Searching Axion-like Particle Dark Matter with Dipole, Toroidal and Wiggler Magnets. *Phys. Rev.*, D85:035018, 2012.

[144] Peter W. Graham and Surjeet Rajendran. Axion Dark Matter Detection with Cold Molecules. *Phys. Rev.*, D84:055013, 2011.

[145] Peter W. Graham and Surjeet Rajendran. New Observables for Direct Detection of Axion Dark Matter. *Phys. Rev.*, D88:035023, 2013.

[146] Dieter Horns, Joerg Jaeckel, Axel Lindner, Andrei Lobanov, Javier Redondo, and Andreas Ringwald. Searching for WISPy Cold Dark Matter with a Dish Antenna. *JCAP*, 1304:016, 2013.

[147] P. Sikivie, N. Sullivan, and D. B. Tanner. Proposal for Axion Dark Matter Detection Using an LC Circuit. *Phys. Rev. Lett.*, 112(13):131301, 2014.

[148] Dieter Horns, Axel Lindner, Andrei Lobanov, and Andreas Ringwald. WISPers from the Dark Side: Radio Probes of Axions and Hidden Photons. In *9th Patras Workshop on Axions, WIMPs & WISPs (PATRAS13) Mainz, Germany, June 24-28, 2013*, 2013.

[149] Christian Beck. Possible resonance effect of axionic dark matter in Josephson junctions. *Phys. Rev. Lett.*, 111:231801, 2013.

[150] Christian Beck. Axion mass estimates from resonant Josephson junctions. *Phys. Dark Univ.*, 7-8:6–11, 2015.

[151] Jooyoo Hong, Jihn E. Kim, Soonkeon Nam, and Yannis Semertzidis. Calculations of resonance enhancement factor in axion-search tube-experiments. 2014.

[152] B. M. Roberts, Y. V. Stadnik, V. A. Dzuba, V. V. Flambaum, N. Leefer, and D. Budker. Limiting P-odd interactions of cosmic fields with electrons, protons and neutrons. *Phys. Rev. Lett.*, 113:081601, 2014.

[153] B. M. Roberts, Y. V. Stadnik, V. A. Dzuba, V. V. Flambaum, N. Leefer, and D. Budker. Parity-violating interactions of cosmic fields with atoms, molecules, and nuclei: Concepts and calculations for laboratory searches and extracting limits. *Phys. Rev.*, D90(9):096005, 2014.

[154] Y. V. Stadnik and V. V. Flambaum. Searching for dark matter and variation of fundamental constants with laser and maser interferometry. *Phys. Rev. Lett.*, 114:161301, 2015.

[155] Christopher T. Hill. Axion Induced Oscillating Electric Dipole Moments. *Phys. Rev.*, D91(11):111702, 2015.

[156] Christopher T. Hill. Axion Induced Oscillating Electric Dipole Moment of the Electron, 2015.

[157] Ben T. McAllister, Stephen R. Parker, and Michael E. Tobar. Axion Dark Matter Coupling to Resonant Photons via Magnetic Field. *Phys. Rev. Lett.*, 116(16):161804, 2016. [Erratum: Phys. Rev. Lett.117,no.15,159901(2016)].

[158] Naoto Yokoi and Eiji Saitoh. Stimulated Emission of Dark Matter Axion from Condensed Matter Excitations. *JHEP*, 01:022, 2018.

[159] Stephon Alexander and Robert Sims. Detecting axions via induced electron spin precession. *Phys. Rev. D*, 98(1):015011, 2018.

[160] ChunJun Cao and Ariel Zhitnitsky. Axion detection via Topological Casimir Effect. *Phys. Rev.*, D96(1):015013, 2017.

[161] Ben T. McAllister, Graeme Flower, Lucas E. Tobar, and Michael E. Tobar. Tunable Supermode Dielectric Resonators for Axion Dark-Matter Haloscopes. *Phys. Rev. Applied*, 9(1):014028, 2018.

[162] Peter W. Graham, David E. Kaplan, Jeremy Mardon, Surjeet Rajendran, William A. Terrano, Lutz Trahms, and Thomas Wilkason. Spin Precession Experiments for Light Axionic Dark Matter. *Phys. Rev. D*, 97(5):055006, 2018.

[163] M. S. Safronova, D. Budker, D. DeMille, Derek F. Jackson Kimball, A. Derevianko, and C. W. Clark. Search for New Physics with Atoms and Molecules. *Rev. Mod. Phys.*, 90(2):025008, 2018.

[164] Christian Beck. Possible resonance effect of dark matter axions in SNS Josephson junctions. In *2017 European Physical Society Conference on High Energy Physics (EPS-HEP 2017) Venice, Italy, July 5-12, 2017*, 2017.

[165] Asimina Arvanitaki, Savas Dimopoulos, and Ken Van Tilburg. Resonant absorption of bosonic dark matter in molecules. *Phys. Rev. X*, 8(4):041001, 2018.

[166] Jihn E. Kim. Constraints on very light axions from cavity experiments. *Phys. Rev.*, D58:055006, 1998.

[167] C. Patrignani et al. Review of Particle Physics. *Chin. Phys.*, C40(10):100001, 2016.

[168] Luca Di Luzio, Federico Mescia, and Enrico Nardi. Redefining the Axion Window. *Phys. Rev. Lett.*, 118(3):031801, 2017.

[169] Prateek Agrawal, JiJi Fan, Matthew Reece, and Lian-Tao Wang. Experimental Targets for Photon Couplings of the QCD Axion. *JHEP*, 02:006, 2018.

[170] R. H. Dicke. The Measurement of Thermal Radiation at Microwave Frequencies. *Rev. Sci. Instrum.*, 17:268, 1946.

[171] H. Peng et al. Cryogenic cavity detector for a large scale cold dark-matter axion search. *Nucl. Instrum. Meth.*, A444:569–583, 2000.

[172] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J.*, C71:1554, 2011. [Erratum: Eur. Phys. J.C73,2501(2013)].

[173] Katherine Freese, Paolo Gondolo, and Heidi Jo Newberg. Detectability of weakly interacting massive particles in the Sagittarius dwarf tidal stream. *Phys. Rev.*, D71:043516, 2005.

[174] Chris W. Purcell, Andrew R. Zentner, and Mei-Yu Wang. Dark Matter Direct Search Rates in Simulations of the Milky Way and Sagittarius Stream. *JCAP*, 1208:027, 2012.

[175] Christopher Savage, Katherine Freese, and Paolo Gondolo. Annual Modulation of Dark Matter in the Presence of Streams. *Phys. Rev.*, D74:043531, 2006.

[176] Ciaran A. J. O'Hare and Anne M. Green. Directional detection of dark matter streams. *Phys. Rev.*, D90(12):123511, 2014.

[177] Alexander J. Millar, Javier Redondo, and Frank D. Steffen. Dielectric haloscopes: sensitivity to the axion dark matter velocity. *JCAP*, 1710(10):006, 2017.

[178] J. V. Sloan et al. Limits on axion–photon coupling or on local axion density: Dependence on models of the Milky Way's dark halo. *Phys. Dark Univ.*, 14:95–102, 2016.

[179] Stephen J. Asztalos et al. Experimental constraints on the axion dark matter halo density. *Astrophys. J.*, 571:L27–L30, 2002.

[180] Stephen J. Asztalos et al. An Improved RF cavity search for halo axions. *Phys. Rev.*, D69:011101, 2004.

[181] Fu-Sin Ling, Pierre Sikivie, and Stuart Wick. Diurnal and annual modulation of cold dark matter signals. *Phys. Rev.*, D70:123503, 2004.

[182] J. D. Vergados and Y. Semertzidis. Axionic dark matter signatures in various halo models. *Nucl. Phys.*, B915:10–18, 2017.

[183] Ciaran A. J. O'Hare and Anne M. Green. Axion astronomy with microwave cavity experiments. *Phys. Rev.*, D95(6):063017, 2017.

[184] Samuel K. Lee, Mariangela Lisanti, Annika H. G. Peter, and Benjamin R. Safdi. Effect of Gravitational Focusing on Annual Modulation in Dark-Matter Direct-Detection Experiments. *Phys. Rev. Lett.*, 112(1):011301, 2014.

[185] Jonah Herzog-Arbeitman, Mariangela Lisanti, Piero Madau, and Lina Necib. Empirical Determination of Dark Matter Velocities using Metal-Poor Stars. *Phys. Rev. Lett.*, 120(4):041102, 2018.

[186] Jonah Herzog-Arbeitman, Mariangela Lisanti, and Lina Necib. The Metal-Poor Stellar Halo in RAVE-TGAS and its Implications for the Velocity Distribution of Dark Matter. *JCAP*, 04:052, 2018.

[187] Marco Battaglieri et al. US Cosmic Visions: New Ideas in Dark Matter 2017: Community Report. In *U.S. Cosmic Visions: New Ideas in Dark Matter*, 7 2017.

[188] Edward John Daw. *A search for halo axions*. PhD thesis, MIT, 1998.

[189] Leanne D. Duffy, P. Sikivie, D. B. Tanner, Stephen J. Asztalos, C. Hagmann, D. Kinion, L. J Rosenberg, K. van Bibber, D. B. Yu, and R. F. Bradley. A high resolution search for dark-matter axions. *Phys. Rev.*, D74:012006, 2006.

[190] P. Sikivie. Experimental Tests of the Invisible Axion. *Phys. Rev. Lett.*, 51:1415–1417, 1983. [Erratum: Phys. Rev. Lett.52,695(1984)].

[191] HC Seton, DM Bussell, JMS Hutchison, and DJ Lurie. Use of a DC SQUID receiver preamplifier in a low field MRI system. *Applied Superconductivity, IEEE Transactions on*, 5(2):3218–3221, 1995.

[192] HC Seton, JMS Hutchison, and DM Bussell. Gradiometer pick-up coil design for a low field SQUID-MRI system. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 8(2):116–120, 1999.

[193] Katherine Freese, Mariangela Lisanti, and Christopher Savage. Colloquium: Annual modulation of dark matter. *Rev. Mod. Phys.*, 85:1561–1581, 2013.

[194] Mark Vogelsberger, A. Helmi, Volker Springel, Simon D. M. White, Jie Wang, Carlos S. Frenk, Adrian Jenkins, A. D. Ludlow, and Julio F. Navarro. Phase-space structure in the local dark matter distribution and its signature in direct detection experiments. *Mon. Not. Roy. Astron. Soc.*, 395:797–811, 2009.

[195] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Power-Constrained Limits. 2011.

[196] Frederik Beaujean, Allen Caldwell, and Olaf Reimann. Is the bump significant? An axion-search example. *Eur. Phys. J. C*, 78(9):793, 2018.

[197] Daniel Bowring (ADMX). ADMX - the Axion Dark Matter eXperiment. APS-DPF 2017, 2017.

[198] I. Stern. ADMX Status. *PoS*, ICHEP2016:198, 2016.

[199] Michael T. Hotz. *A SQUID-Based RF Cavity Search for Dark Matter Axions*. PhD thesis, Washon U., 2013.

[200] Sangjun Lee, Sung Woo Youn, and Y. K. Semertzidis. Comment on "Axion Dark Matter Coupling to Resonant Photons via Magnetic Field". 6 2016.

[201] Samuel K. Lee, Mariangela Lisanti, and Benjamin R. Safdi. Dark-Matter Harmonics Beyond Annual Modulation. *JCAP*, 1311:033, 2013.

[202] F. Feroz, M. P. Hobson, and M. Bridges. MultiNest: an efficient and robust Bayesian inference tool for cosmology and particle physics. *Mon. Not. Roy. Astron. Soc.*, 398:1601–1614, 2009.

[203] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astron. Astrophys.*, 564:A125, 2014.

[204] Malte Buschmann, Joachim Kopp, Benjamin R. Safdi, and Chih-Liang Wu. Stellar Wakes from Dark Matter Subhalos. *Phys. Rev. Lett.*, 120(21):211101, 2018.

[205] Michal Maciejewski, Mark Vogelsberger, Simon D. M. White, and Volker Springel. Bound and unbound substructures in Galaxy-scale Dark Matter haloes. *Mon. Not. Roy. Astron. Soc.*, 415:2475, 2011.

[206] Mariangela Lisanti and David N. Spergel. Dark Matter Debris Flows in the Milky Way. *Phys. Dark Univ.*, 1:155–161, 2012.

[207] Marc Kamionkowski and Savvas M. Koushiappas. Galactic substructure and direct detection of dark matter. *Phys. Rev.*, D77:103509, 2008.

[208] J. I. Read, G. Lake, O. Agertz, and Victor P. Debattista. Thin, thick and dark discs in LCDM. *Mon. Not. Roy. Astron. Soc.*, 389:1041–1057, 2008.

[209] T. Bruch, J. Read, L. Baudis, and G. Lake. Detecting the Milky Way's Dark Disk. *Astrophys. J.*, 696:920–923, 2009.

[210] J. I. Read, L. Mayer, A. M. Brooks, F. Governato, and G. Lake. A dark matter disc in three cosmological simulations of Milky Way mass galaxies. *MNRAS*, 397:44–51, July 2009.

[211] C. W. Purcell, J. S. Bullock, and M. Kaplinghat. The Dark Disk of the Milky Way. *ApJ*, 703:2275–2284, October 2009.

[212] Bijunath R. Patla, Robert J. Nemiroff, Dieter H. H. Hoffmann, and Konstantin Zioutas. Flux Enhancement of Slow-moving Particles by Sun or Jupiter: Can they be Detected on Earth? *Astrophys. J.*, 780:158, 2014.

[213] Sergio Bertolucci, Konstantin Zioutas, Sebastian Hofmann, and Marios Maroudas. The Sun and its Planets as detectors for invisible matter. *Phys. Dark Univ.*, 17:13–21, 2017.

[214] K. Zioutas et al. Search for axions in streaming dark matter. 3 2017.

[215] Ann E. Nelson and Jakub Scholtz. Dark Light, Dark Matter and the Misalignment Mechanism. *Phys. Rev. D*, 84:103501, 2011.

[216] M. Ahlers, H. Gies, J. Jaeckel, J. Redondo, and A. Ringwald. Light from the hidden sector. *Phys. Rev. D*, 76:115005, 2007.

[217] Joerg Jaeckel and Andreas Ringwald. A Cavity Experiment to Search for Hidden Sector Photons. *Phys. Lett. B*, 659:509–514, 2008.

[218] S. Lee, S. Ahn, J. Choi, B.R. Ko, and Y.K. Semertzidis. Axion Dark Matter Search around 6.7 $\mu$eV. *Phys. Rev. Lett.*, 124(10):101802, 2020.

[219] L. Zhong et al. Results from phase 1 of the HAYSTAC microwave cavity axion experiment. *Phys. Rev. D*, 97(9):092001, 2018.

[220] Saptarshi Chaudhuri, Kent D. Irwin, Peter W. Graham, and Jeremy Mardon. Optimal Electromagnetic Searches for Axion and Hidden-Photon Dark Matter. 4 2019.

[221] Saptarshi Chaudhuri, Kent Irwin, Peter W. Graham, and Jeremy Mardon. Fundamental Limits of Electromagnetic Axion and Hidden-Photon Dark Matter Searches: Part I - The Quantum Limit. 3 2018.

[222] William DeRocco and Anson Hook. Axion interferometry. *Phys. Rev. D*, 98(3):035021, 2018.

[223] Ippei Obata, Tomohiro Fujita, and Yuta Michimura. Optical Ring Cavity Search for Axion Dark Matter. *Phys. Rev. Lett.*, 121(16):161301, 2018.

[224] Hongwan Liu, Brodi D. Elwood, Matthew Evans, and Jesse Thaler. Searching for Axion Dark Matter with Birefringent Cavities. *Phys. Rev. D*, 100(2):023548, 2019.

[225] Stefan Knirck, Alexander J. Millar, Ciaran A.J. O'Hare, Javier Redondo, and Frank D. Steffen. Directional axion detection. *JCAP*, 11:051, 2018.

[226] Matthew Lawson, Alexander J. Millar, Matteo Pancaldi, Edoardo Vitagliano, and Frank Wilczek. Tunable axion plasma haloscopes. *Phys. Rev. Lett.*, 123(14):141802, 2019.

[227] Stefan Knirck, Jan Schütte-Engel, Alexander Millar, Javier Redondo, Olaf Reimann, Andreas Ringwald, and Frank Steffen. A First Look on 3D Effects in Open Axion Haloscopes. *JCAP*, 08:026, 2019.

[228] Andrei Derevianko. Detecting dark-matter waves with a network of precision-measurement tools. *Phys. Rev. A*, 97(4):042506, 2018.

[229] D. Lynden-Bell and R. M. Lynden-Bell. Ghostly streams from the formation of the Galaxy's halo. *Monthly Notices of the Royal Astronomical Society*, 275(2):429–442, 07 1995.

[230] V Belokurov, D Erkal, N W Evans, S E Koposov, and A J Deason. Co-formation of the disc and the stellar halo. *Monthly Notices of the Royal Astronomical Society*, 478(1):611–619, 06 2018.

[231] Amina Helmi, Carine Babusiaux, Helmer H. Koppelman, Davide Massari, Jovan Veljanoski, and Anthony G. A. Brown. The merger that led to the formation of the Milky Way's inner stellar halo and thick disk. *Nature*, 563(7729):85–88, October 2018.

[232] M. Malnou, D.A. Palken, B.M. Brubaker, Leila R. Vale, Gene C. Hilton, and K.W. Lehnert. Squeezed vacuum used to accelerate the search for a weak classical signal. *Phys. Rev. X*, 9(2):021023, 2019.

[233] Samuel K. Lee and Annika H.G. Peter. Probing the Local Velocity Distribution of WIMP Dark Matter with Directional Detectors. *JCAP*, 04:029, 2012.

[234] K.A. van Bibber and S.D. Kinion. Experimental searches for galactic halo axions. *Phil. Trans. Roy. Soc. Lond. A*, 361:2553–2567, 2003.

[235] Heidi Jo Newberg et al. Sagittarius tidal debris 90 kpc from the Galactic center. *Astrophys. J. Lett.*, 596:L191–L194, 2003.

[236] Brian Yanny et al. A Low latitude halo stream around the Milky Way. *Astrophys. J.*, 588:824, 2003. [Erratum: Astrophys.J. 605, 575–577 (2004)].

[237] Steven R. Majewski, M.F. Skrutskie, Martin D. Weinberg, and James C. Ostheimer. A 2mass all-sky view of the Sagittarius dwarf galaxy: I. Morphology of the Sagittarius core and tidal arms. *Astrophys. J.*, 599:1082–1115, 2003.

[238] Lina Necib, Bryan Ostdiek, Mariangela Lisanti, Timothy Cohen, Marat Freytsis, Shea Garrison-Kimmel, Philip F. Hopkins, Andrew Wetzel, and Robyn Sanderson. Evidence for a Vast Prograde Stellar Stream in the Solar Vicinity. *Nature Astron.*, 4(11):1078–1083, 2020.

[239] Ciaran A.J. O'Hare, Christopher McCabe, N. Wyn Evans, GyuChul Myeong, and Vasily Belokurov. Dark matter hurricane: Measuring the S1 stream with dark matter detectors. *Phys. Rev. D*, 98(10):103006, 2018.

[240] G.C. Myeong, N.W. Evans, V. Belokurov, N.C. Amorisco, and S. Koposov. Halo Substructure in the SDSS-Gaia Catalogue : Streams and Clumps. *Mon. Not. Roy. Astron. Soc.*, 475(2):1537–1548, 2018.

[241] G. C. Myeong, N. W. Evans, V. Belokurov, J. L. Sand ers, and S. E. Koposov. Discovery of new retrograde substructures: the shards of $\omega$ Centauri? *MNRAS*, 478(4):5449–5459, August 2018.

[242] Til Piffl et al. The RAVE survey: the Galactic escape speed and the mass of the Milky Way. *Astron. Astrophys.*, 562:A91, 2014.

[243] Ralph Schönrich, James Binney, and Walter Dehnen. Local kinematics and the local standard of rest. *MNRAS*, 403(4):1829–1833, April 2010.

[244] Michal Maciejewski, Mark Vogelsberger, Simon D. M. White, and Volker Springel. Bound and unbound substructures in Galaxy-scale dark matter haloes. *MNRAS*, 415(3):2475–2484, August 2011.

[245] Farhan Feroz and M.P. Hobson. Multimodal nested sampling: an efficient and robust alternative to MCMC methods for astronomical data analysis. *Mon. Not. Roy. Astron. Soc.*, 384:449, 2008.

[246] F. Feroz, M.P. Hobson, E. Cameron, and A.N. Pettitt. Importance Nested Sampling and the MultiNest Algorithm. *Open J. Astrophys.*, 2(1):10, 2019.

[247] J. Buchner, A. Georgakakis, K. Nandra, L. Hsu, C. Rangel, M. Brightman, A. Merloni, M. Salvato, J. Donley, and D. Kocevski. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *A&A*, 564:A125, April 2014.

[248] Riccardo Catena and Piero Ullio. A novel determination of the local dark matter density. *JCAP*, 1008:004, 2010.

[249] Fabio Iocco, Miguel Pato, Gianfranco Bertone, and Philippe Jetzer. Dark Matter distribution in the Milky Way: microlensing and dynamical constraints. *JCAP*, 11:029, 2011.

[250] Zeeshan Ahmed et al. Quantum Sensing for High Energy Physics. In *First workshop on Quantum Sensing for High Energy Physics*, 3 2018.

[251] Luca Di Luzio, Andreas Ringwald, and Carlos Tamarit. Axion mass prediction from minimal grand unification. *Phys. Rev. D*, 98(9):095011, 2018.

[252] Max Tegmark, Anthony Aguirre, Martin Rees, and Frank Wilczek. Dimensionless constants, cosmology and other dark matters. *Phys. Rev.*, D73:023505, 2006.

[253] D. Horns and M. Meyer. Indications for a pair-production anomaly from the propagation of VHE gamma-rays. *JCAP*, 1202:033, 2012.

[254] Manuel Meyer, Dieter Horns, and Martin Raue. First lower limits on the photon-axion-like particle coupling from very high energy gamma-ray observations. *Phys. Rev.*, D87(3):035027, 2013.

[255] Alessandro De Angelis, Marco Roncadelli, and Oriana Mansutti. Evidence for a new light spin-zero boson from cosmological gamma-ray propagation? *Phys. Rev. D*, 76:121301, 2007.

[256] Clare Burrage, Anne-Christine Davis, and Douglas J. Shaw. Active Galactic Nuclei Shed Light on Axion-like-Particles. *Phys. Rev. Lett.*, 102:201101, 2009.

[257] C. Gatti et al. The Klash Proposal: Status and Perspectives. In *14th Patras Workshop on Axions, WIMPs and WISPs*, 11 2018.

[258] Stephen J. Asztalos et al. Large scale microwave cavity search for dark matter axions. *Phys. Rev. D*, 64:092003, 2001.

[259] S. J. Asztalos et al. A SQUID-based microwave cavity search for dark-matter axions. *Phys. Rev. Lett.*, 104:041301, 2010.

[260] C. Hagmann, P. Sikivie, N. S. Sullivan, and D. B. Tanner. Results from a search for cosmic axions. *Phys. Rev.*, D42:1297–1300, 1990.

[261] S. De Panfilis, A. C. Melissinos, B. E. Moskowitz, J. T. Rogers, Y. K. Semertzidis, Walter Wuensch, H. J. Halama, A. G. Prodell, W. B. Fowler, and F. A. Nezrick. Limits on the Abundance and Coupling of Cosmic Axions at 4.5-Microev ¡ m(a) ¡ 5.0-Microev. *Phys. Rev. Lett.*, 59:839, 1987.

[262] C. Hagmann, D. Kinion, W. Stoeffl, K. van Bibber, E. Daw, H. Peng, Leslie J Rosenberg, J. LaVeigne, P. Sikivie, N. S. Sullivan, D. B. Tanner, F. Nezrick, Michael S. Turner, D. M. Moltz, J. Powell, and N. A. Golubev. Results from a high-sensitivity search for cosmic axions. *Phys. Rev. Lett.*, 80:2043–2046, Mar 1998.

[263] V. Anastassopoulos et al. New CAST Limit on the Axion-Photon Interaction. *Nature Phys.*, 13:584–590, 2017.

[264] Hooman Davoudiasl, Dan Hooper, and Samuel D. McDermott. Inflatable dark matter. *Physical Review Letters*, 116(3), Jan 2016.

[265] Peter W. Graham and Adam Scherlis. Stochastic axion scenario. *Phys. Rev. D*, 98(3):035017, 2018.

[266] Prateek Agrawal, Gustavo Marques-Tavares, and Wei Xue. Opening up the QCD axion window. *JHEP*, 03:049, 2018.

[267] Luca Visinelli and Paolo Gondolo. Axion cold dark matter in non-standard cosmologies. *Phys. Rev. D*, 81:063508, 2010.

[268] Raymond T. Co, Francesco D'Eramo, and Lawrence J. Hall. Supersymmetric axion grand unified theories and their predictions. *Phys. Rev.*, D94(7):075001, 2016.

[269] Marco Farina, Duccio Pappadopulo, Fabrizio Rompineve, and Andrea Tesi. The photophilic QCD axion. *JHEP*, 01:095, 2017.

[270] Super conducting systems inc. http://www.superconductingsystems.com.

[271] S. Henry, M. Pipe, A. Cottle, C. Clarke, U. Divakar, and A. Lynch. Characterisation of superconducting capillaries for magnetic shielding of twisted-wire pairs in a neutron electric dipole moment experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 763:155–162, 2014.

[272] M. Frigo and S.G. Johnson. The design and implementation of fftw3. *Proceedings of the IEEE*, 93(2):216–231, 2005.

[273] Paul J. McMillan and James J. Binney. The uncertainty in galactic parameters. *Monthly Notices of the Royal Astronomical Society*, 402(2):934–940, Feb 2010.

[274] Derek F. Jackson Kimball et al. Overview of the Cosmic Axion Spin Precession Experiment (CASPEr). *Springer Proc. Phys.*, 245:105–121, 2020.

[275] M. Silva-Feaver, S. Chaudhuri, H. Cho, C. Dawson, P. Graham, K. Irwin, S. Kuenstner, D. Li, J. Mardon, H. Moseley, R. Mule, A. Phipps, S. Rajendran, Z. Steffen, and B. Young. Design overview of DM Radio pathfinder experiment. *IEEE Transactions on Applied Superconductivity*, 27(4):1–4, June 2017.

[276] Alexander V Gramolin, Deniz Aybas, Dorian Johnson, Janos Adam, and Alexander O Sushkov. Search for axion-like dark matter with ferromagnets. *Nature Physics*, 2020.

[277] COMSOL Multiphysics® v. 5.4. www.comsol.com. COMSOL AB, Stockholm, Sweden.

[278] Magnicon. http://www.magnicon.com/.

[279] M. Ackermann et al. Search for Gamma-ray Spectral Lines with the Fermi Large Area Telescope and Dark Matter Implications. *Phys. Rev. D*, 88:082002, 2013.

[280] Andrea Albert, German A. Gomez-Vargas, Michael Grefe, Carlos Munoz, Christoph Weniger, Elliott D. Bloom, Eric Charles, Mario N. Mazziotta, and Aldo Morselli. Search for 100 MeV to 10 GeV $\gamma$-ray lines in the Fermi-LAT data and implications for gravitino dark matter in $\mu\nu$SSM. *JCAP*, 10:023, 2014.

[281] M. Ackermann et al. Updated search for spectral lines from Galactic dark matter interactions with pass 8 data from the Fermi Large Area Telescope. *Phys. Rev. D*, 91(12):122002, 2015.

[282] Dark Matter New Initiatives FY2019. https://science.osti.gov/-/media/hep/pdf/Awards/Dark-Matter-New-Initiatives-FY-2019--List-of-Awards.pdf?la=en&hash=7134EDEA489651A3097DC8DB8C59069384A4538E.

[283] J. L. Ouellet et al. Probing the QCD Axion with DMRadio-m$^3$. *Snowmass 2021 Letter of Interest*, CF2(217), 2020.

[284] S. Chaudhuri et al. DMRadio-GUT: Probing GUT-scale QCD Axion Dark Matter. *Snowmass 2021 Letter of Interest*, CF2(219), 2020.

[285] A. F. Leder et al. Magnet R&D for Low-Mass Axion Searches. *Snowmass 2021 Letter of Interest*, AF5(244), 2020.

[286] S. E. Kuenstner et al. Radio Frequency Quantum Upconverters: Precision Metrology for Fundamental Physics. *Snowmass 2021 Letter of Interest*, IF1(193), 2020.

[287] Richard A. Battye, Bjoern Garbrecht, Jamie I. McDonald, Francesco Pace, and Sankarshana Srinivasan. Dark matter axion detection in the radio/mm-waveband. *Phys. Rev. D*, 102(2):023504, 2020.

[288] Mikaël Leroy, Marco Chianese, Thomas D. P. Edwards, and Christoph Weniger. Radio Signal of Axion-Photon Conversion in Neutron Stars: A Ray Tracing Analysis. *Phys. Rev. D*, 101(12):123003, 2020.

[289] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[290] Walter Wuensch, S. De Panfilis-Wuensch, Y. K. Semertzidis, J. T. Rogers, A. C. Melissinos, H. J. Halama, B. E. Moskowitz, A. G. Prodell, W. B. Fowler, and F. A. Nezrick. Results of a Laboratory Search for Cosmic Axions and Other Weakly Coupled Light Particles. *Phys. Rev.*, D40:3153, 1989.

[291] T. Braine et al. Extended Search for the Invisible Axion with the Axion Dark Matter Experiment. *Phys. Rev. Lett.*, 124(10):101303, 2020.

[292] Richard M. Prestage, Marty Bloss, Joe Brandt, Hong Chen, Ray Creager, Paul Demorest, John Ford, Glenn Jones, Amanda Kepley, Adam Kobelski, Paul Marganian, Melinda Mello, David McMahon, Randy McCullough, Jason Ray, D. Anish Roshi, Dan Werthimer, and Mark Whitehead. The versatile GBT astronomical spectrometer (VEGAS): Current status and future plans. In *2015 URSI-USNC Radio Science Meeting*, page 4, July 2015.

[293] P. Lazarus, R. Karuppusamy, E. Graikou, R. N. Caballero, D. J. Champion, K. J. Lee, J. P. W. Verbiest, and M. Kramer. Prospects for high-precision pulsar timing with the new Effelsberg PSRIX backend. *MNRAS*, 458(1):868–880, May 2016.

[294] Bob Garwood, Paul Marganian, Jim Braatz, and Ron Maddalena. `tGBTIDL` data analysis software, May 2005.

[295] W. van Straten and M. Bailes. DSPSR: Digital Signal Processing Software for Pulsar Astronomy. *Publications of the Astron. Soc. of Australia*, 28:1–14, January 2011.

[296] Jo Bovy and Scott Tremaine. On the local dark matter density. *Astrophys. J.*, 756:89, 2012.

[297] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. The Structure of cold dark matter halos. *Astrophys. J.*, 462:563–575, 1996.

[298] Julio F. Navarro, Carlos S. Frenk, and Simon D. M. White. A Universal density profile from hierarchical clustering. *Astrophys. J.*, 490:493–508, 1997.

[299] David L. Kaplan and M. H. van Kerkwijk. A Coherent timing solution for the nearby isolated neutron star RX J0720.4-3125. *Astrophys. J.*, 628:L45–L48, 2005.

[300] D. L. Kaplan and M. H. van Kerkwijk. Constraining the Spin-down of the Nearby Isolated Neutron Star RX J0806.4-4123, and Implications for the Population of Nearby Neutron Stars. *Astrophys. J.*, 705:798–808, 2009.

[301] Malte Buschmann, Raymond T. Co, Christopher Dessert, and Benjamin R. Safdi. Axion Emission Can Explain a New Hard X-Ray Excess from Nearby Isolated Neutron Stars. *Phys. Rev. Lett.*, 126(2):021102, 2021.

[302] P. Goldreich and W. H. Julian. Pulsar Electrodynamics. *Ap. J.*, 157:869, August 1969.

[303] C. Dessert et al. *Supplementary Data*, 2020. https://github.com/bsafdi/axionSSC.

[304] Philip F Hopkins et al. FIRE-2 Simulations: Physics versus Numerics in Galaxy Formation. *Mon. Not. Roy. Astron. Soc.*, 480(1):800–863, 2018.

[305] A. Weltman et al. Fundamental Physics with the Square Kilometre Array. *Publ. Astron. Soc. Austral.*, 37:e002, 2020.

[306] Rendong Nan, Di Li, Chengjin Jin, Qiming Wang, Lichun Zhu, Wenbai Zhu, Haiyan Zhang, Youling Yue, and Lei Qian. The Five-Hundred Aperture Spherical Radio Telescope (fast) Project. *International Journal of Modern Physics D*, 20(6):989–1024, January 2011.

[307] H. Primakoff. Photoproduction of neutral mesons in nuclear electric fields and the mean life of the neutral meson. *Phys. Rev.*, 81:899, 1951.

[308] Georg G. Raffelt. Astrophysical axion bounds diminished by screening effects. *Phys. Rev. D*, 33:897–909, Feb 1986.

[309] Adrian Ayala, Inma Domínguez, Maurizio Giannotti, Alessandro Mirizzi, and Oscar Straniero. Revisiting the bound on axion-photon coupling from Globular Clusters. *Phys. Rev. Lett.*, 113(19):191302, 2014.

[310] Fiona A. Harrison et al. The Nuclear Spectroscopic Telescope Array (NuSTAR) High-Energy X-Ray Mission. *Astrophys. J.*, 770:103, 2013.

[311] Georg G. Raffelt. Axion Constraints From White Dwarf Cooling Times. *Phys. Lett.*, 166B:402–406, 1986.

[312] J. Isern, E. Garcia-Berro, S. Torres, and S. Catalan. Axions and the cooling of white dwarf stars. *Astrophys. J.*, 682:L109, 2008.

[313] J. Isern, S. Catalan, E. Garcia-Berro, and S. Torres. Axions and the white dwarf luminosity function. *J. Phys. Conf. Ser.*, 172:012005, 2009.

[314] J. Isern, E. Garcia-Berro, L. G. Althaus, and A. H. Corsico. Axions and the pulsation periods of variable white dwarfs revisited. *Astron. Astrophys.*, 512:A86, 2010.

[315] Marcelo M. Miller Bertolami, Brenda E. Melendez, Leandro G. Althaus, and Jordi Isern. Revisiting the axion bounds from the Galactic white dwarf luminosity function. *JCAP*, 1410(10):069, 2014.

[316] Javier Redondo. Solar axion flux from the axion-electron coupling. *JCAP*, 1312:008, 2013.

[317] Nicolás Viaux, Márcio Catelan, Peter B. Stetson, Georg Raffelt, Javier Redondo, Aldo A. R. Valcarce, and Achim Weiss. Neutrino and axion bounds from the globular cluster M5 (NGC 5904). *Phys. Rev. Lett.*, 111:231301, 2013.

[318] Maurizio Giannotti, Igor Irastorza, Javier Redondo, and Andreas Ringwald. Cool WISPs for stellar cooling excesses. *JCAP*, 1605(05):057, 2016.

[319] Maurizio Giannotti, Igor G. Irastorza, Javier Redondo, Andreas Ringwald, and Ken'ichi Saikawa. Stellar Recipes for Axion Hunters. *JCAP*, 1710(10):010, 2017.

[320] Jean-François Fortin and Kuver Sinha. Constraining Axion-Like-Particles with Hard X-ray Emission from Magnetars. *JHEP*, 06:048, 2018.

[321] Christopher Dessert, Andrew J. Long, and Benjamin R. Safdi. X-ray Signatures of Axion Conversion in Magnetic White Dwarf Stars. *Phys. Rev. Lett.*, 123(6):061104, 2019.

[322] Christopher Dessert, Joshua W. Foster, and Benjamin R. Safdi. Hard X-ray Excess from the Magnificent Seven Neutron Stars. *Astrophys. J.*, 904(1):42, 2020.

[323] Eric D. Carlson. Pseudoscalar conversion and X-rays from stars. *Physics Letters B*, 344:245–251, Feb 1995.

[324] Maurizio Giannotti. Fermi-LAT and NuSTAR as Stellar Axionscopes. In *13th Patras Workshop on Axions, WIMPs and WISPs*, pages 23–27, 2018.

[325] M. Meyer, M. Giannotti, A. Mirizzi, J. Conrad, and M.A. Sánchez-Conde. Fermi Large Area Telescope as a Galactic Supernovae Axionscope. *Phys. Rev. Lett.*, 118(1):011103, 2017.

[326] Georg G. Raffelt. Plasmon decay into low-mass bosons in stars. *Phys. Rev. D*, 37:1356–1359, Mar 1988.

[327] Bill Paxton, Lars Bildsten, Aaron Dotter, Falk Herwig, Pierre Lesaffre, and Frank Timmes. Modules for Experiments in Stellar Astrophysics (MESA). *ApJS*, 192(1):3, January 2011.

[328] Bill Paxton, Matteo Cantiello, Phil Arras, Lars Bildsten, Edward F. Brown, Aaron Dotter, Christopher Mankovich, M. H. Montgomery, Dennis Stello, F. X. Timmes, and Richard Townsend. Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars. *ApJS*, 208(1):4, September 2013.

[329] Francisco Najarro, Donald F. Figer, D.John Hillier, and Rolf P. Kudritzki. Metallicity in the Galactic Center. The Arches cluster. *Astrophys. J. Lett.*, 611:L105–L108, 2004.

[330] Francisco Najarro, Don F. Figer, D.John Hillier, T.R. Geballe, and Rolf P. Kudritzki. Metallicity in the Galactic Center: The Quintuplet cluster. *Astrophys. J.*, 691:1816–1827, 2009.

[331] I. Hunter, D.J. Lennon, P.L. Dufton, C. Trundle, S. Simon-Diaz, S.J. Smartt, R.S.I. Ryans, and C.J. Evans. The VLT-FLAMES survey of massive stars: Atmospheric parameters and rotational velocity distributions for B-type stars in the Magellanic Clouds. *Astron. Astrophys.*, 479:541, 2008.

[332] Ines Brott et al. Rotating Massive Main-Sequence Stars II: Simulating a Population of LMC early B-type Stars as a Test of Rotational Mixing. *Astron. Astrophys.*, 530:A116, 2011.

[333] Pavel Kroupa. On the variation of the initial mass function. *Mon. Not. Roy. Astron. Soc.*, 322:231, 2001.

[334] C. Weidner and J. S. Vink. The masses, and the mass discrepancy of o-type stars. *Astron. Astrophys.*, 524:A98, Nov 2010.

[335] W.-R. Hamann, G. Graefener, and Adriane Liermann. The Galactic WN stars: Spectral analyses with line-blanketed model atmospheres versus stellar evolution models with and without rotation. *Astron. Astrophys.*, 457:1015, 2006.

[336] J. S. Clark, M. E. Lohr, L. R. Patrick, F. Najarro, H. Dong, and D. F. Figer. An updated stellar census of the quintuplet cluster. *Astron. Astrophys.*, 618:A2, Oct 2018.

[337] Mojgan Aghakhanloo, Jeremiah W. Murphy, Nathan Smith, John Parejko, Mariangelly Díaz-Rodríguez, Maria R. Drout, Jose H. Groh, Joseph Guzman, and Keivan G. Stassun. Inferring the parallax of Westerlund 1 from Gaia DR2. *MNRAS*, 492(2):2497–2509, February 2020.

[338] Ben Davies and Emma R. Beasor. The distances to star clusters hosting Red Supergiants: $\chi$ Per, NGC 7419, and Westerlund 1. *MNRAS*, 486(1):L10–L14, June 2019.

[339] J.S. Clark, B. W. Ritchie, and I. Negueruela. A vlt/flames survey for massive binaries in westerlund 1. vii. cluster census. *Astron. Astrophys.*, Nov 2019.

[340] A. E. Piatti, E. Bica, and J. J. Claria. Fundamental parameters of the highly reddened young open clusters Westerlund 1 and 2. *A&AS*, 127:423–432, February 1998.

[341] Georg Raffelt and Leo Stodolsky. Mixing of the Photon with Low Mass Particles. *Phys. Rev.*, D37:1237, 1988.

[342] Q. Daniel Wang, Hui Dong, and Cornelia Lang. The interplay between star formation and the nuclear environment of our Galaxy: deep X-ray observations of the Galactic centre Arches and Quintuplet clusters. *MNRAS*, 371(1):38–54, September 2006.

[343] Ronnie Jansson and Glennys R. Farrar. A New Model of the Galactic Magnetic Field. *ApJ*, 757(1):14, Sep 2012.

[344] Ronnie Jansson and Glennys R. Farrar. The galactic magnetic field. *The Astrophysical Journal*, 761(1):L11, nov 2012.

[345] J. M. Yao, R. N. Manchester, and N. Wang. A new electron-density model for estimation of pulsar and frb distances. *The Astrophysical Journal*, 835(1):29, Jan 2017.

[346] James M. Cordes and T. J. W. Lazio. NE2001. 1. A New model for the galactic distribution of free electrons and its fluctuations. 2002.

[347] M. S. Pshirkov, P. G. Tinyakov, P. P. Kronberg, and K. J. Newton-McGee. Deriving the Global Structure of the Galactic Magnetic Field from Faraday Rotation Measures of Extragalactic Sources. *ApJ*, 738(2):192, September 2011.

[348] Vladislav V. Izmodenov and Dmitry B. Alexashov. Magnitude and direction of the local interstellar magnetic field inferred from Voyager 1 and 2 interstellar data and global heliospheric model. *A&A*, 633:L12, January 2020.

[349] E. J. Zirnstein, J. Heerikhuisen, H. O. Funsten, G. Livadiotis, D. J. McComas, and N. V. Pogorelov. Local Interstellar Magnetic Field Determined from the Interstellar Boundary Explorer Ribbon. *ApJL*, 818(1):L18, February 2016.

[350] Priscilla C. Frisch, B. G. Andersson, Andrei Berdyugin, Herbert O. Funsten, Antonio M. Magalhaes, David J. McComas, Vilppu Piirola, Nathan A. Schwadron, Jonathan D. Slavin, and Sloane J. Wiktorowicz. Comparisons of the Interstellar Magnetic Field Directions Obtained from the IBEX Ribbon and Interstellar Polarizations. *ApJ*, 724(2):1473–1479, December 2010.

[351] M. Salvati. The local Galactic magnetic field in the direction of Geminga. *A&A*, 513:A28, April 2010.

[352] Katia Ferriere. Interstellar magnetic fields in the Galactic center region. *Astron. Astrophys.*, 505:1183, 2009.

[353] Roland M. Crocker, David I. Jones, Fulvio Melia, Jürgen Ott, and Raymond J. Protheroe. A lower limit of 50 microgauss for the magnetic field near the galactic centre. *Nature*, 463(7277):65–67, Jan 2010.

[354] R. M. Crocker, D. I. Jones, F. Aharonian, C. J. Law, F. Melia, T. Oka, and J. Ott. Wild at heart: the particle astrophysics of the galactic centre. *Monthly Notices of the Royal Astronomical Society*, 413(2):763–788, Mar 2011.

[355] Mehmet Guenduez, Julia Becker Tjus, Katia Ferrière, and Ralf-Jürgen Dettmar. A novel analytical model of the magnetic field configuration in the Galactic Center. *Astron. Astrophys.*, 644:A71, 2020.

[356] Kaya Mori et al. NuSTAR Hard X-ray Survey of the Galactic Center Region. I. Hard X-ray Morphology and Spectroscopy of the Diffuse Emission. *Astrophys. J.*, 814(2):94, 2015.

[357] JaeSub Hong et al. NuSTAR Hard X-ray Survey of the Galactic Center Region II: X-ray Point Sources. *Astrophys. J.*, 825(2):132, 2016.

[358] J. K. Blackburn, R. A. Shaw, H. E. Payne, J. J. E. Hayes, and Heasarc. FTOOLS: A general package of software to manipulate FITS files, December 1999.

[359] D. Porquet, J. Rodriguez, S. Corbel, P. Goldoni, R.S. Warwick, A. Goldwurm, and A. Decourchelle. Xmm-newton study of the persistent x-ray source 1e1743.1-2843 located in the galactic center direction. *Astron. Astrophys.*, 406:299–304, 2003.

[360] Daniel R. Wik et al. NuSTAR Observations of the Bullet Cluster: Constraints on Inverse Compton Emission. *Astrophys. J.*, 792(1):48, 2014.

[361] A. Borghese et al. The multi-outburst activity of the magnetar in Westerlund I. *Mon. Not. Roy. Astron. Soc.*, 484(3):2931–2943, 2019.

[362] Kristin K. Madsen, Finn E. Christensen, William W. Craig, Karl W. Forster, Brian W. Grefenstette, Fiona A. Harrison, Hiromasa Miyasaka, and Vikram Rana. Observational Artifacts of NuSTAR: Ghost Rays and Stray Light. *arXiv e-prints*, page arXiv:1711.02719, November 2017.

[363] Hongjun An, Kristin K. Madsen, Niels J. Westergaard, Steven E. Boggs, Finn E. Christensen, William W. Craig, Charles J. Hailey, Fiona A. Harrison, Daniel K. Stern, and William W. Zhang. In-flight PSF calibration of the NuSTAR hard X-ray optics. *Proc. SPIE Int. Soc. Opt. Eng.*, 9144:91441Q, 2014.

[364] Christopher S. Reynolds, M. C. David Marsh, Helen R. Russell, Andrew C. Fabian, Robyn Smith, Francesco Tombesi, and Sylvain Veilleux. Astrophysical limits on very light axion-like particles from chandra grating spectroscopy of NGC 1275. *The Astrophysical Journal*, 890(1):59, feb 2020.

[365] Maxim Libanov and Sergey Troitsky. On the impact of magnetic-field models in galaxy clusters on constraints on axion-like particles from the lack of irregularities in high-energy spectra of astrophysical sources. *Phys. Lett.*, B802:135252, 2020.

[366] Alexandre Payez, Carmelo Evoli, Tobias Fischer, Maurizio Giannotti, Alessandro Mirizzi, and Andreas Ringwald. Revisiting the SN1987A gamma-ray limit on ultralight axion-like particles. *JCAP*, 1502(02):006, 2015.

[367] Georg G. Raffelt. Astrophysical methods to constrain axions and other novel particle phenomena. *Phys. Rept.*, 198:1–113, 1990.

[368] Georg G. Raffelt. Astrophysical axion bounds. *Lect. Notes Phys.*, 741:51–71, 2008. [,51(2006)].

[369] Jae Hyeok Chang, Rouven Essig, and Samuel D. McDermott. Revisiting Supernova 1987A Constraints on Dark Photons. *JHEP*, 01:107, 2017.

[370] Nitsan Bar, Kfir Blum, and Guido D'Amico. Is there a supernova bound on axions? *Phys. Rev. D*, 101(12):123025, 2020.

[371] M. Giannotti, B. Grefenstette, A. Mirizzi, M. Nynka, K. Perez, B. Roach, O. Straiero, and M. Xiao. Constraints on light axions from a hard x-ray observation of betelgeuse. To appear 2020.

[372] Scott Dodelson and Lawrence M. Widrow. Sterile-neutrinos as dark matter. *Phys. Rev. Lett.*, 72:17–20, 1994.

[373] Xiang-Dong Shi and George M. Fuller. A New dark matter candidate: Nonthermal sterile neutrinos. *Phys. Rev. Lett.*, 82:2832–2835, 1999.

[374] Alexander Kusenko. Sterile neutrinos, dark matter, and the pulsar velocities in models with a Higgs singlet. *Phys. Rev. Lett.*, 97:241301, 2006.

[375] Palash B. Pal and Lincoln Wolfenstein. Radiative decays of massive neutrinos. *Phys. Rev. D*, 25:766–773, Feb 1982.

[376] Tsutomu Yanagida. Horizontal Symmetry and Masses of Neutrinos. *Prog. Theor. Phys.*, 64:1103, 1980.

[377] Rabindra N. Mohapatra and Goran Senjanovic. Neutrino Mass and Spontaneous Parity Violation. *Phys. Rev. Lett.*, 44:912, 1980. [,231(1979)].

[378] Christopher Dessert, Nicholas L. Rodd, and Benjamin R. Safdi. The dark matter interpretation of the 3.5-keV line is inconsistent with blank-sky observations. *Science*, 367:1465, 2020.

[379] Esra Bulbul, Maxim Markevitch, Adam Foster, Randall K. Smith, Michael Loewenstein, and Scott W. Randall. Detection of An Unidentified Emission Line in the Stacked X-ray spectrum of Galaxy Clusters. *Astrophys. J.*, 789:13, 2014.

[380] Alexey Boyarsky, Oleg Ruchayskiy, Dmytro Iakubovskyi, and Jeroen Franse. Unidentified Line in X-Ray Spectra of the Andromeda Galaxy and Perseus Galaxy Cluster. *Phys. Rev. Lett.*, 113:251301, 2014.

[381] O. Urban, N. Werner, S. W. Allen, A. Simionescu, J. S. Kaastra, and L. E. Strigari. A Suzaku Search for Dark Matter Emission Lines in the X-ray Brightest Galaxy Clusters. *Mon. Not. Roy. Astron. Soc.*, 451(3):2447–2461, 2015.

[382] Tesla E. Jeltema and Stefano Profumo. Discovery of a 3.5 keV line in the Galactic Centre and a critical look at the origin of the line across astronomical targets. *Mon. Not. Roy. Astron. Soc.*, 450(2):2143–2152, 2015.

[383] Nico Cappelluti, Esra Bulbul, Adam Foster, Priyamvada Natarajan, Megan C. Urry, Mark W. Bautz, Francesca Civano, Eric Miller, and Randall K. Smith. Searching for the 3.5 keV Line in the Deep Fields with Chandra: the 10 Ms observations. *Astrophys. J.*, 854(2):179, 2018.

[384] Christopher Dessert, Nicholas L. Rodd, and Benjamin R. Safdi. Response to a comment on Dessert et al. "The dark matter interpretation of the 3.5 keV line is inconsistent with blank-sky observations". 6 2020.

[385] Shunsaku Horiuchi, Philip J. Humphrey, Jose Onorbe, Kevork N. Abazajian, Manoj Kaplinghat, and Shea Garrison-Kimmel. Sterile neutrino dark matter bounds from galaxies of the Local Group. *Phys. Rev.*, D89(2):025017, 2014.

[386] D. Malyshev, A. Neronov, and D. Eckert. Constraints on 3.55 keV line emission from stacked observations of dwarf spheroidal galaxies. *Phys. Rev.*, D90:103506, 2014.

[387] Michael E. Anderson, Eugene Churazov, and Joel N. Bregman. Non-Detection of X-Ray Emission From Sterile Neutrinos in Stacked Galaxy Spectra. *Mon. Not. Roy. Astron. Soc.*, 452(4):3905–3923, 2015.

[388] Takayuki Tamura, Ryo Iizuka, Yoshitomo Maeda, Kazuhisa Mitsuda, and Noriko Y. Yamasaki. An X-ray Spectroscopic Search for Dark Matter in the Perseus Cluster with Suzaku. *Publ. Astron. Soc. Jap.*, 67:23, 2015.

[389] Tesla E. Jeltema and Stefano Profumo. Deep XMM Observations of Draco rule out at the 99% Confidence Level a Dark Matter Decay Origin for the 3.5 keV Line. *Mon. Not. Roy. Astron. Soc.*, 458(4):3592–3596, 2016.

[390] F. A. Aharonian et al. $Hitomi$ constraints on the 3.5 keV line in the Perseus galaxy cluster. *Astrophys. J.*, 837(1):L15, 2017.

[391] A. Gewering-Peine, D. Horns, and J. H. M. M. Schmitt. A sensitive search for unknown spectral emission lines in the diffuse X-ray background with XMM-Newton. *JCAP*, 1706(06):036, 2017.

[392] Oleg Ruchayskiy, Alexey Boyarsky, Dmytro Iakubovskyi, Esra Bulbul, Dominique Eckert, Jeroen Franse, Denys Malyshev, Maxim Markevitch, and Andrii Neronov. Searching for decaying dark matter in deep XMM–Newton observation of the Draco dwarf spheroidal. *Mon. Not. Roy. Astron. Soc.*, 460(2):1390–1398, 2016.

[393] A. Neronov, Denys Malyshev, and Dominique Eckert. Decaying dark matter search with NuSTAR deep sky observations. *Phys. Rev. D*, 94(12):123504, 2016.

[394] Kerstin Perez, Kenny C. Y. Ng, John F. Beacom, Cora Hersh, Shunsaku Horiuchi, and Roman Krivonos. Almost closing the $\nu$MSM sterile neutrino dark matter window with NuSTAR. *Phys. Rev. D*, 95(12):123002, 2017.

[395] Brandon M. Roach, Kenny C.Y. Ng, Kerstin Perez, John F. Beacom, Shunsaku Horiuchi, Roman Krivonos, and Daniel R. Wik. NuSTAR Tests of Sterile-Neutrino Dark Matter: New Galactic Bulge Observations and Combined Impact. *Phys. Rev. D*, 101(10):103011, 2020.

[396] S. Riemer-Sørensen et al. Dark matter line emission constraints from NuSTAR observations of the Bullet Cluster. *Astrophys. J.*, 810(1):48, 2015.

[397] Kenny C.Y. Ng, Brandon M. Roach, Kerstin Perez, John F. Beacom, Shunsaku Horiuchi, Roman Krivonos, and Daniel R. Wik. New Constraints on Sterile Neutrino Dark Matter from $NuSTAR$ M31 Observations. *Phys. Rev. D*, 99:083005, 2019.

[398] Tetsutaro Higaki, Kwang Sik Jeong, and Fuminobu Takahashi. The 7 keV axion dark matter and the X-ray line signal. *Phys. Lett. B*, 733:25–31, 2014.

[399] Alexander Kusenko, Michael Loewenstein, and Tsutomu T. Yanagida. Moduli dark matter and the search for its decay line using Suzaku X-ray telescope. *Phys. Rev. D*, 87(4):043508, 2013.

[400] See /github.com/bsafdi/XMM_BSO_DATA for the list of observations used and processed data products.

[401] K.M. Gorski, Eric Hivon, A.J. Banday, B.D. Wandelt, F.K. Hansen, M. Reinecke, and M. Bartelman. HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere. *Astrophys. J.*, 622:759–771, 2005.

[402] Georges Aad et al. Measurement of Higgs boson production in the diphoton decay channel in pp collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector. *Phys. Rev. D*, 90(11):112015, 2014.

[403] Sivaram Ambikasaran, Daniel Foreman-Mackey, Leslie Greengard, David W. Hogg, and Michael O'Neil. Fast direct methods for gaussian processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):252–265, 2016.

[404] Meghan Frate, Kyle Cranmer, Saarik Kalia, Alexander Vandenberg-Rodes, and Daniel Whiteson. Modeling Smooth Backgrounds and Generic Localized Signals with Gaussian Processes. 9 2017.

[405] Palash B. Pal and Lincoln Wolfenstein. Radiative Decays of Massive Neutrinos. *Phys. Rev. D*, 25:766, 1982.

[406] Mark Vogelsberger, Federico Marinacci, Paul Torrey, and Ewald Puchwein. Cosmological Simulations of Galaxy Formation. *Nature Rev. Phys.*, 2(1):42–66, 2020.

[407] Oleg Y. Gnedin, Andrey V. Kravtsov, Anatoly A. Klypin, and Daisuke Nagai. Response of dark matter halos to condensation of baryons: Cosmological simulations and improved adiabatic contraction model. *Astrophys. J.*, 616:16–26, 2004.

[408] Matthieu Schaller, Carlos S. Frenk, Richard G. Bower, Tom Theuns, Adrian Jenkins, Joop Schaye, Robert A. Crain, Michelle Furlong, Claudio Dalla Vecchia, and I.G. McCarthy. Baryon effects on the internal structure of $\Lambda$CDM haloes in the EAGLE simulations. *Mon. Not. Roy. Astron. Soc.*, 451(2):1247–1267, 2015.

[409] Qirong Zhu, Federico Marinacci, Moupiya Maji, Yuexing Li, Volker Springel, and Lars Hernquist. Baryonic impact on the dark matter distribution in Milky Way-sized galaxies and their satellites. *Mon. Not. Roy. Astron. Soc.*, 458(2):1559–1580, 2016.

[410] Aaron A. Dutton, Andrea V. Macciò, Avishai Dekel, Liang Wang, Gregory S. Stinson, Aura Obreja, Arianna Di Cintio, Chris B. Brook, Tobias Buck, and Xi Kang. NIHAO IX: The role of gas inflows and outflows in driving the contraction and expansion of cold dark matter haloes. *Mon. Not. Roy. Astron. Soc.*, 461(3):2658–2675, 2016.

[411] Mark R. Lovell et al. The fraction of dark matter within galaxies from the IllustrisTNG simulations. *Mon. Not. Roy. Astron. Soc.*, 481(2):1950–1975, 2018.

[412] T.K. Chan, D. Kereš, J. Oñorbe, P.F. Hopkins, A.L. Muratov, C. A. Faucher-Giguère, and E. Quataert. The impact of baryonic physics on the structure of dark matter haloes: the view from the FIRE cosmological simulations. *Mon. Not. Roy. Astron. Soc.*, 454(3):2981–3001, 2015.

[413] Pol Mollitor, Emmanuel Nezri, and Romain Teyssier. Baryonic and dark matter distribution in cosmological simulations of spiral galaxies. *Mon. Not. Roy. Astron. Soc.*, 447(2):1353–1369, 2015.

[414] Matthieu Portail, Ortwin Gerhard, Christopher Wegg, and Melissa Ness. Dynamical modelling of the galactic bulge and bar: the Milky Way's pattern speed, stellar and dark matter mass distribution. *Mon. Not. Roy. Astron. Soc.*, 465(2):1621–1644, February 2017.

[415] Alexandres Lazar, James S. Bullock, Michael Boylan-Kolchin, T. K. Chan, Philip F. Hopkins, Andrew S. Graus, Andrew Wetzel, Kareem El-Badry, Coral Wheeler, Maria C. Straight, Dušan Kereš, Claude-André Faucher-Giguère, Alex Fitts, and Shea Garrison-Kimmel. A dark matter profile to model diverse feedback-induced core sizes of ΛCDM haloes. *MNRAS*, 497(2):2393–2417, July 2020.

[416] Marius Cautun, Alejandro Benítez-Llambay, Alis J. Deason, Carlos S. Frenk, Azadeh Fattahi, Facundo A. Gómez, Robert J. J. Grand, Kyle A. Oman, Julio F. Navarro, and Christine M. Simpson. The milky way total mass profile as inferred from Gaia DR2. *MNRAS*, 494(3):4291–4313, April 2020.

[417] Anna-Christina Eilers, David W. Hogg, Hans-Walter Rix, and Melissa K. Ness. The Circular Velocity Curve of the Milky Way from 5 to 25 kpc. *Astrophysical Journal*, 871(1):120, January 2019.

[418] Laura L. Watkins, Roeland P. van der Marel, Sangmo Tony Sohn, and N. Wyn Evans. Evidence for an Intermediate-mass Milky Way from Gaia DR2 Halo Globular Cluster Motions. *Astrophysical Journal*, 873(2):118, March 2019.

[419] Thomas M. Callingham, Marius Cautun, Alis J. Deason, Carlos S. Frenk, Wenting Wang, Facundo A. Gómez, Robert J. J. Grand, Federico Marinacci, and Ruediger Pakmor. The mass of the Milky Way from satellite dynamics. *MNRAS*, 484(4):5453–5467, April 2019.

[420] Pablo F. de Salas and Axel Widmark. Dark matter local density determination: recent observations and future prospects. 12 2020.

[421] A.D. Dolgov, S.H. Hansen, S. Pastor, S.T. Petcov, G.G. Raffelt, and D.V. Semikoz. Cosmological bounds on neutrino degeneracy improved by flavor oscillations. *Nucl. Phys. B*, 632:363–382, 2002.

[422] Pasquale D. Serpico and Georg G. Raffelt. Lepton asymmetry and primordial nucleosynthesis in the era of precision cosmology. *Phys. Rev. D*, 71:127301, 2005.

[423] Alexey Boyarsky, Oleg Ruchayskiy, and Mikhail Shaposhnikov. The Role of sterile neutrinos in cosmology and astrophysics. *Ann. Rev. Nucl. Part. Sci.*, 59:191–214, 2009.

[424] A. Drlica-Wagner et al. Milky Way Satellite Census. I. The Observational Selection Function for Milky Way Satellites in DES Y3 and Pan-STARRS DR1. *Astrophys. J.*, 893:1, 2020.

[425] E. O. Nadler et al. Milky Way Satellite Census. III. Constraints on Dark Matter Properties from Observations of Milky Way Satellite Galaxies. *Phys. Rev. Lett.*, 126:091101, 2021.

[426] Dominic Sicilian, Nico Cappelluti, Esra Bulbul, Francesca Civano, Massimo Moscetti, and Christopher S. Reynolds. Probing the Milky Way's Dark Matter Halo for the 3.5 keV Line. *Astrophys. J.*, 905(2):146, 2020.

[427] X. Barcons, K. Nandra, D. Barret, J.W. den Herder, A.C. Fabian, L. Piro, and M.G. Watson. Athena: the X-ray observatory to study the hot and energetic Universe. *J. Phys. Conf. Ser.*, 610(1):012008, 2015.

[428] Science with the X-ray Imaging and Spectroscopy Mission (XRISM). 3 2020.

[429] R. Adler. *The Geometry of Random Fields*. Society for Industrial and Applied Mathematics, 2010.

[430] Edward W. Kolb and Igor I. Tkachev. Femtolensing and picolensing by axion miniclusters. *Astrophys. J.*, 460:L25–L28, 1996.

[431] M. C. Huang and P. Sikivie. The Structure of Axionic Domain Walls. *Phys. Rev. D*, 32:1560, 1985.

[432] Ralph B. D'Agostino. An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2):341–348, 1971.

[433] Ralph D'Agostino and E. S. Pearson. Tests for departure from normality. Empirical results for the distributions of $b^2$ and $\sqrt{b1}$. *Biometrika*, 60(3):613–622, 12 1973.

[434] David J. E. Marsh. Axions and ALPs: a very short introduction. In *Proceedings, 13th Patras Workshop on Axions, WIMPs and WISPs, (PATRAS 2017): Thessaloniki, Greece, 15 May 2017 - 19, 2017*, pages 59–74, 2018.

[435] Raymond T. Co, Francesco D'Eramo, and Lawrence J. Hall. Gravitino or Axino Dark Matter with Reheat Temperature as high as $10^{16}$ GeV. *JHEP*, 03:005, 2017.

[436] D. Drung, C. Assmann, J. Beyer, A. Kirste, M. Peters, F. Ruede, and T. Schurig. Highly sensitive and easy-to-use squid sensors. *IEEE Transactions on Applied Superconductivity*, 17:699–704, 2007.

[437] https://www.alazartech.com/.

[438] R. A. Perley and B. J. Butler. Integrated Polarization Properties of 3C48, 3C138, 3C147, and 3C286. *The Astrophysical Journal Supplement Series*, 206(2):16, Jun 2013.

[439] R. A. Perley and B. J. Butler. An accurate flux density scale from 50 MHz to 50 GHz. *The Astrophysical Journal Supplement Series*, 230(1):7, may 2017.

[440] A. W. Hotan, W. van Straten, and R. N. Manchester. PSRCHIVE and PSRFITS: An Open Approach to Radio Pulsar Data Storage and Analysis. *Publications of the Astron. Soc. of Australia*, 21:302–309, 2004.

[441] R. M. Crocker, D. I. Jones, F. Aharonian, C. J. Law, F. Melia, T. Oka, and J. Ott. Wild at Heart: the particle astrophysics of the Galactic Centre. *MNRAS*, 413(2):763–788, May 2011.

[442] W. Reich, P. Reich, and E. Fuerst. The Effelsberg 21 CM radio continuum survey of the-Galacticplane between L = 357 deg. and L = 95.5 deg. *Astronomy and Astrophysics, Supplement*, 83:539, June 1990.

[443] Malcolm Gray. *Maser molecules*, pages 156–185. Cambridge Astrophysics. Cambridge University Press, 2012.

[444] R. P. Norris and R. S. Booth. Observations of OH masers in W3 OH. *MNRAS*, 195:213–226, April 1981.

[445] Haihua Qiao, Juan Li, Zhiqiang Shen, Xi Chen, and Xingwu Zheng. The catalogues and mid-infrared environment of interstellar OH masers. *MNRAS*, 441(4):3137–3147, July 2014.

[446] J. L. Caswell. Positions of hydroxyl masers at 1665 and 1667 MHz. *MNRAS*, 297(1):215–235, June 1998.

[447] Martin McCall and Dan Censor. Relativity and mathematical tools: Waves in moving media. *American Journal of Physics*, 75(12):1134–1140, 2007.

[448] E. Glebbeek, E. Gaburov, S. E. de Mink, O. R. Pols, and S. F. Portegies Zwart. The evolution of runaway stellar collision products. *Astronomy & Astrophysics*, 497(1):255–264, Feb 2009.

[449] F. Yusef-Zadeh. The Origin of the Galactic center nonthermal radio filaments: Young stellar clusters. *Astrophys. J.*, 598:325–333, 2003.

[450] J. S. Clark, M. E. Lohr, L. R. Patrick, and F. Najarro. The arches cluster revisited. iii. an addendum to the stellar census, 2018.

[451] Roman Krivonos, Maica Clavel, JaeSub Hong, Kaya Mori, Gabriele Ponti, Juri Poutanen, Farid Rahoui, John Tomsick, and Sergey Tsygankov. NuSTAR and XMM–Newton observations of the Arches cluster in 2015: fading hard X-ray emission from the molecular cloud. *Mon. Not. Roy. Astron. Soc.*, 468(3):2822–2835, 2017.

[452] Roman A. Krivonos et al. First hard X-ray detection of the non-thermal emission around the Arches cluster: morphology and spectral studies with NuSTAR. *Astrophys. J.*, 781:107, 2014.

[453] Fabrice Martins, R. Genzel, D.J. Hillier, F. Eisenhauer, T. Paumard, S. Gillessen, T. Ott, and S. Trippe. Stellar and wind properties of massive stars in the central parsec of the Galaxy. *Astron. Astrophys.*, 468:233, 2007.

[454] "Users Guide to the XMM-Newton Science Analysis System", Issue 14.0, 2018 (ESA: XMM-Newton SOC).

[455] D. H. Lumb, R. S. Warwick, M. Page, and A. De Luca. X-ray background measurements with xmm-newton epic. *Astron. Astrophys.*, 389:93, 2002.

[456] de Plaa, J., Werner, N., Simionescu, A., Kaastra, J. S., Grange, Y. G., and Vink, J. Cold fronts and multi-temperature structures in the core of abell 2052. *A&A*, 523:A81, 2010.

[457] Leccardi, A. and Molendi, S. Radial temperature profiles for a large sample of galaxy clusters observed with xmm-newton*. *A&A*, 486(2):359–373, 2008.

[458] Adam Foster, Randall K. Smith, Nancy S. Brickhouse, and Xiaohong Cui. AtomDB and PyAtomDB: Atomic Data and Modelling Tools for High Energy and Non-Maxwellian Plasmas. In *American Astronomical Society Meeting Abstracts #227*, volume 227 of *American Astronomical Society Meeting Abstracts*, page 211.08, January 2016.

[459] Katsuji Koyama et al. Iron and Nickel Line Diagnostics for the Galactic Center Diffuse Emission. *Publ. Astron. Soc. Jap.*, 59:245, 2007.

[460] Alexey Boyarsky, Denys Malyshev, Oleg Ruchayskiy, and Denys Savchenko. Technical comment on the paper of Dessert et al. "The dark matter interpretation of the 3.5 keV line is inconsistent with blank-sky observations". 4 2020.

[461] Kevork N. Abazajian. Technical Comment on "The dark matter interpretation of the 3.5-keV line is inconsistent with blank-sky observations". 4 2020.

[462] A. Boyarsky, D. Iakubovskyi, O. Ruchayskiy, and D. Savchenko. Surface brightness profile of the 3.5 keV line in the Milky Way halo. 12 2018.