Precision Cosmology with Inaccurate Data: Assessing and Addressing Systematic Errors in Large-Scale Structure

by

Noah J. Weaverdyck

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 Dragan Huterer, Chair
Professor Fred C. Adams
Professor Dante Amidei
Professor August E. Evrard
Professor Christopher J. Miller

Noah J. Weaverdyck
nweaverd@umich.edu
ORCID: 0000-0001-9382-5199

Acknowledgements

I would not be where I am today if it weren't for countless other people who have supported me along the way. I surely underestimate how much of my success is thanks to the efforts and sacrifices of others, in addition to a large dose of luck. While it is impossible to name them all (and most will never even read this), I will try to recognize some of these important people here.

I am enormously grateful to my scientific advisor Dragan Huterer, whose enthusiastic encouragement has been invaluable, and who created a collaborative research environment that fostered personal and professional growth. I am thankful to Jessie Muir for the patience, council and friendship she provided me throughout my PhD studies, as well as Anqi Chen and the other cosmology graduate students who contributed to an exciting and welcoming research environment. Thanks to Gus Evrard and Dan Amidei for striving to improve pedagogy and for taking the time to discuss the world beyond physics, and to the rest of my coauthors for contributing to the studies in this work.

This work would not have happened without the slew of IT (Brandon), student services (Lauren and Chrissy), and building and custodial staff (Karen, Paul, Georgia, Tracy, and more) whose work tends to be undervalued but was instrumental in enabling the research in this dissertation.

I am also aware that in many respects, the broader social systems that have afforded me privileged access to resources all too often function to disenfranchise and subjugate others. I performed most of this work at the University of Michigan, which was funded by and founded on Anishinaabeg (including Odawa, Ojibwe, and Boodewadomi) and Wyandot lands ceded in coercive historical treaties, just as almost all property in the United States has been obtained through the forceful dispossession of indigenous peoples. Physicists and astronomers often fail to acknowledge the colonial dynamics that have fueled our research and which we have an obligation to interrogate, especially as it relates to affirming the sovereignty of indigenous peoples today.

I am appreciative of the Graduate Employees Organization and all of the individuals who selflessly work to organize for a more equitable, safe and supportive work environment at the

University of Michigan. Through them, I have gained a deep appreciation of the power of organized labor and the importance of solidarity in combating the oppressive structures and atomizing nature of late-stage capitalism, which manifest in numerous ways in the modern University system, not least in this last year during the COVID-19 pandemic.

More broadly, from the ravages of anthropogenic climate change to the impending collapse of the world's ecosystems, it is undeniable that we are facing crises unprecedented in human history. The fact that these imminent threats are human-driven and solvable with existing technology provides enormous opportunity, and a grave responsibility unlike any generation before us. I could not have made it through my PhD if it weren't for the hope and empowerment provided by the brilliant and passionate individuals organizing in groups like GEO, the Climate Action Movement at U-M, and Science for the People. They have provided me invaluable insight into theories of change and the appropriate role of science and scientists in society. In this truly historic moment, I could not in good conscience have worked on the research contained herein without having these channels to build power and work toward a more hopeful future.

My friends and family in Ann Arbor deserve my eternal gratitude for the love and support they have given me over the last six years. The camaraderie of Ceren and Eunice (along with the sustenance of Pita Kebab Grill) was vital during my first year, when Jackson problem sets made me especially question my choice to return to academia after a four-year hiatus. Similarly am I indebted to Anna and Joel, who always welcomed me warmly whenever I swung by unannounced at 11:40 PM after taking the last bus home. Through my partner Sasha, I have developed a fuller understanding of the world and grown into a fuller version of myself, for which I will always be grateful. Finally, thanks to my parents Shelly and Curtis for fostering in me an intense curiosity about the world, and providing a strong foundation of support from which to strive for self-actualization. I have always known that I can count on them.

Oh and of course this work is only possible thanks to a whole lot of open source software that people have shared with the world, including `Python` [1], `Anaconda` [2], `Matplotlib` [3], `CAMB` [4], `CLASS` [5], `Halofit` [6], `HMCode` [7], `Numpy` [8] `Scipy` [9], `Seaborn` [10], `Healpy` [11], `Healpix` [12], `NaMaster` [13], `Scikit-Learn` [14], and `Statsmodels` [15]. Thanks for putting communal progress over personal profit — we're all better off because of it.

Preface

This dissertation is based on research I undertook at the University of Michigan from 2016 – 2020. The three studies presented are ones in which I took a leading role; all have been published in academic journals.

Chapter 2 is based on Ref. [16], and was coauthored by Jessie Muir and Dragan Huterer. It grew out of their general study of using large-scale structure surveys to reconstruct the Integrated-Sachs Wolfe effect map in Ref. [17], which identified calibration errors as a limiting factor. It was published in *Physical Review D*.

Chapter 3 is based on Ref. [18] and was coauthored by Dragan Huterer. It was published in *Monthly Notices of the Royal Astronomical Society*.

Chapter 4 is adapted from Ref. [19], which was published in *The Astrophysical Journal* and for which I was second author to Xiaolei Li. I was highly involved in both the analysis and writing of that work, and Saroj Adhikari, Dragan Huterer, Jessie Muir, and Hao-Yi Wu contributed as coauthors.

Toward the end of my PhD studies, I became involved with the Dark Energy Survey, where I have contributed to a number of analyses to varying degrees (E.g. [20–23]), including (by request) a significant effort to assess the robustness of the Year 3 galaxy clustering results using the methods described in Chapter 3. Because my contribution for most of these works was considerably less than the ones contained here, I have not included them in this dissertation.

Table of Contents

Acknowledgements **ii**

Preface **iv**

List of Tables **vii**

List of Figures **viii**

List of Appendices **xviii**

Abstract **xix**

Chapter 1: Introduction **1**

1.1 The Standard Model 2
1.2 Redshift and the Expanding Universe 6
1.3 History of the Universe: The Executive Summary 6
1.4 Inflation and Primordial Perturbations 8
1.5 Structure Growth 10
1.6 Large-Scale Structure Surveys 13
1.7 Cosmic Microwave Background 15
1.8 Connecting Theory to Data 16
1.9 Outline of Thesis 20

Chapter 2: Reconstructing the Integrated Sachs-Wolfe Map in the Presence of Systematic Errors **22**

2.1 Introduction 22
2.2 Methodology 25
2.3 Results 34
2.4 Impact of Calibration Errors on S/N Estimates 43

2.5 Conclusions 46

Chapter 3: Mitigating Contamination in LSS Surveys: A Comparison of Methods 50

3.1 Introduction 50
3.2 Contamination Model 54
3.3 Background: Existing Mitigation Methods 55
3.4 Placing into a Common Mathematical Framework 63
3.5 Applications 72
3.6 Evaluating Performance 79
3.7 Simulation Results 82
3.8 Summary of Methods 92
3.9 Conclusions 93

Chapter 4: Chasing the Inflationary Spectral Runnings in the Presence of Systematic Errors 97

4.1 Introduction 97
4.2 Methods 100
4.3 Results 107
4.4 Systematic Biases in Model Parameters 111
4.5 Constraining the Second Running: ΛCDM $+\alpha_s + \beta_s$ 115
4.6 Conclusions 116

Chapter 5: Closing Remarks 119

Appendices 124

Bibliography 145

List of Tables

1.1 Characteristic behavior of the density and scale factor in a flat Universe for various epochs where a single component dominated the energy budget. . . . 5

1.2 Cosmological parameters. Parameters above the line correspond to the six parameters that define the standard ΛCDM model. Below the line are additional parameters that are varied when testing extended models, but fixed to the reported values in ΛCDM. 18

2.1 Mean reconstruction quality coefficients $\bar{\rho}$ of ISW map reconstructions for various combinations of input maps and select levels of calibration error. The second column indicates $\bar{\rho}$ for the case of zero calibration error. The following columns show the reconstruction quality for two nonzero values of the calibration error variance; here $R_\ell(C_\ell^{\rm th})$ $[R_\ell(C_\ell^{\rm sky})]$ indicates the case where calibration errors are unaccounted [accounted] for in the estimator. Note, when $\sigma^2_{\rm cal} = 0$, $C_\ell^{\rm th} = C_\ell^{\rm sky}$. 41

4.1 Cosmological parameters, their fiducial values, and numerical derivative step sizes used for the Fisher matrix calculation. The last two parameters correspond to the Mead model for describing nonlinear effects. 100

List of Figures

1.1 Schematic history of the Universe leading up to now (purple star). Dashed lines indicate transitions between epochs dominated by a single matter-energy component, with labels and symbols to indicate key events. The approximate times and redshifts of these are given by the timeline at bottom, which also shows the ranges probed by LSS (red) and CMB (blue) experiments. The time dependence of the scale factor during each epoch is indicated at top. . . 7

1.2 68.3%, 95.4%, and 99.7% credible intervals on the matter and dark energy densities under a ΛCDM model with curvature allowed. Combining data from Type Ia supernovae, CMB, and LSS experiments results in much stronger parameter constraints (gray) than any probe individually, as they are sensitive to different parameter combinations. SNe Ia (blue) show a clear detection of Ω_Λ and an accelerating Universe, BAO constraints from LSS surveys (green) place tight constraints on Ω_m which sources structure growth, and measurements of the CMB (orange) tell us that the Universe is very close to flat, with $\Omega_{TOT} \approx 1$. Adapted from Suzuki et al. [24]. 17

2.1 *Left:* Planck CMB map (using the `SEVEM` algorithm); *Right*: Reconstructed ISW map using the Planck CMB map and the CMB lensing potential as tracers. Adapted from Figures 1 and 11 in Ref. [25]. 23

2.2 Effects of calibration errors on galaxy power. The solid curve shows the theoretical angular power spectrum for the NVSS survey [25, 26]. The colored dashed curves show the theoretical spectrum with two representative levels of calibration error. The shaded region is the 1σ uncertainty from the survey's sample variance, and the dotted curves indicate the ideal, all-sky cosmic variance. 25

2.3 Same as Fig. 2.2, but for the SDSS MphG catalog, following Refs. [25, 26]. In this case, the sample variance is driven by sky coverage (f_{sky}=0.22) as opposed to number density as for NVSS. 26

2.4 Schematic illustration of the general analysis methodology described in Sec. 2.2. Correlated "true" galaxy and ISW maps are simulated for each realization, given a set of cosmological parameters and LSS survey properties. A realization of a calibration error field modulates the true galaxy map to produce the observed galaxy field, which is then used to reconstruct the ISW map using the estimator described in Sec. 2.2.2, and compared to truth. 31

2.5 Quality of map reconstruction ρ vs. the calibration error variance $\sigma^2_{\rm cal}$ for our fiducial, Euclid-like survey. The colors of the lines indicate how tomographic information is handled, showing that splitting the survey into six redshift bins (red) improves the reconstruction compared to the single-bin case (blue). Solid curves indicate cases when the calibration error is included in the ISW estimator $[R_\ell(C_\ell^{\rm sky})]$, while the dashed curves show the reconstructions in which the effects of the calibration errors are not included ($[R_\ell(C_\ell^{\rm th})]$ (see Sec. 2.2.2 for details). Points (offset horizontally for clarity) show the mean ($\bar{\rho}$) of 2,000 realizations, with error bars indicating the 68% spread across realizations. The corresponding smooth curves are $\hat{\rho}$, the analytical estimate of $\bar{\rho}$ from Eq. (2.17). The inset illustrates the redshift distribution across bins overlaid with the ISW kernel in gray (reproduced from Ref. [17]). The vertical, shaded regions show the approximate current and projected levels of control over residual calibration errors. Calibration errors between redshift bins are modeled as uncorrelated. 35

2.6 Reconstruction quality when using binned Euclid-like survey and CMB intensity data separately and in combination. The purple curve and shaded band show the mean and 68% spread of ρ from simulations. As in Fig. 2.5, red curves are results when using the binned Euclid-like survey alone, whereas blue curves are the result of using both the fiducial survey and CMB intensity map. Like in Fig. 2.5, solid curves are for the case where calibration error power is correctly modeled in the estimator and dashed curves are for when they are not modeled at all. Neglecting the presence of calibration errors in a LSS map can actually degrade the quality of the ISW reconstruction compared to using the CMB temperature alone. 36

2.7 Comparison of ISW reconstruction quality using the LSS surveys and CMB temperature individually and in combination, for various levels of calibration error in the Euclid-like and SPHEREx-like surveys. Colors are the same as those of Fig. 2.6. Both of the LSS surveys are split into six redshift bins (see Sec. 2.3.1), with calibration errors uncorrelated between bins and surveys. The dashed curve shows the combined reconstruction if calibration errors are not included in the estimator. Using LSS surveys to improve the ISW map reconstruction from the CMB temperature only case requires calibration errors to be controlled to $\sigma^2_{\rm cal} \lesssim 10^{-4}$. 40

2.8 *Top:* Effect of cross-correlation between calibration errors in different bins of the fiducial Euclid-like survey, given by $C_\ell^{{\rm cal},ij} = r_{\rm cc}\sqrt{C_\ell^{{\rm cal},ii} C_\ell^{{\rm cal},jj}}$, for bins $i \neq j$. Solid curves have calibration errors accounted for in the estimator $[R_\ell(C_\ell^{\rm sky})]$. The dashed curve indicates the case where only the autopower contributions of the calibration errors are accounted for in the estimator $[R_\ell(C_\ell^{\rm XY,th})]$ and the dotted curve indicates the case where calibration errors are not accounted for at all $[R_\ell(C_\ell^{\rm th})]$. Cross-correlation of the errors results in mild degradation of the reconstruction for $\sigma^2_{\rm cal} \lesssim 10^{-4}$, but otherwise has little effect as long as the auto-correlation is correctly modeled in the estimator. *Bottom:* Dependence of ρ on the shape of C_ℓ^{cal}. Solid curves indicate C_ℓ^{cal} of the same form as Eq. (2.6) but with width $w^{\rm cal}$ varied. The dashed curve indicates the case where the error spectrum takes the form $C_\ell^{cal} \propto (\ell+1)^{-2}$. Reconstruction fares worse when calibration error power contributions are more concentrated at the largest angles, where the ISW kernel is largest. In all cases, the estimator uses the true observed LSS power spectrum ($C_\ell^{\rm sky}$). . 42

2.9 Contribution to squared signal-to-noise ratio per multipole for our fiducial Euclid-like survey with varying levels of calibration error. Total combined S/N for each level of calibration error $\sigma^2_{\rm cal}$ is given in the legend. 43

3.1 *Left:* The two-point correlation function of DES Y3 `redMaGiC` galaxies with $0.8 < z \leq 0.9$. The uncorrected sample (red) shows significantly higher galaxy clustering than after the template-based corrections are applied (green), demonstrating the importance of cleaning in modern-day LSS analyses. The black line shows the best-fit ΛCDM theory curve from the fiducial 3×2-pt analysis (which includes weak lensing data) and is in much better agreement after correction. Adapted from Ref. [23]. *Right:* Real world sky maps of several observational quantities used as templates for the correction scheme in the DES Y3 analysis. 52

3.2 Analysis procedure for a single map. A set of templates is generated (dashed box) along with a true overdensity map $\delta_{\rm true}$. A subset of the templates (orange box) contaminate the true overdensity map to generate the observed overdensity field $\delta_{\rm obs}$. We generate an estimated signal map $\hat{\delta}$ using one of the cleaning methods, and compare it to the truth, either at a map-level or power-spectrum-level. This is repeated for many realizations of the signal map and the performance of each cleaning method is assessed. 53

3.3 Illustration of the DES-Y1 cleaning method, showing the total observed pixel overdensity ($\delta_{\rm obs}$) as a function of a *template's* pixel overdensity, in ten evenly-spaced bins. Given the estimated covariance matrix (diagonals shown by blue error bars), the best-fit trend (blue line) can be calculated and used to reweight the observed map, producing a corrected map whose dependence on the template is removed (orange points, with corresponding standard errors on the pixel means). The process is then iterated for other templates until a satisfactory threshold is reached; see text for details. 57

3.4 Schematic illustration of the difference between the Template Subtraction and (pseudo-C_ℓ) Mode Projection methods. Template Subtraction allows templates to have different levels of contamination at each scale. This is analogous to performing Mode Projection, but first decomposing each template map into a series of derived templates, each corresponding to a different harmonic ℓ. See Sec. 3.4.1 for details. 65

3.5 The error in estimates of the overdensity δ in a toy Gaussian map when contaminated with a single template. Gray points indicate the pixel-based difference between the observed, uncleaned overdensity and the true overdensity when the contamination is multiplicative (additive contamination would lie directly along the dotted line). Orange points are the result when erroneously assuming the contamination is only additive. Blue points are the result when correctly treating the multiplicative component. 71
3.6 Distribution of pixel errors before cleaning (gray), after cleaning with Mode Projection but *before* multiplicative correction (orange), and *after* multiplicative correction (blue). The errors have been calculated as the RMSE of each pixel across 100 cleaned mocks in our fiducial configuration of a DES-like survey as described in Sec. 3.6, and have been normalized to the expected dispersion from the true overdensity field. 72
3.7 Best-fit L1 and L2 penalty coefficients in the regularization technique described in Sec. 3.5.2, as a function of the number of templates used for cleaning, $N_{\rm tpl}$ (new signal and template maps are generated at each value of $N_{\rm tpl}$). In all cases, 12 templates are contaminating the observed data (vertical dashed line). The importance of the L1 penalty, facilitating template selection, becomes increasingly important as more templates are included for cleaning. Lines and shaded region indicate the median and central 68% probability mass of 50 mocks at each $N_{\rm tpl}$ for the central bin of our fiducial DES-like survey. Here, $\rho_{\rm tpl} = 0.2$ within template groups, though plots for other $\rho_{\rm tpl}$ look similar. See Sec. 3.6 for details of implementation. 76
3.8 Root-mean-square error of pixel overdensity estimates, normalized to expected dispersion from the true overdensity due to cosmic variance, vs. pixel leverage for 100 signal realizations. The vertical axis shows the standard error across pixels in 1000 equal-sized bins (in this case 197 pixels per bin at `Nside`=128). The error in both observed and estimated overdensity scales as roughly $\propto H_{ii}^{1/2}$ for all methods (dashed line, to guide the eye). The dotted vertical line indicates a commonly used threshold of $3\times$ the mean leverage across pixels to identify pixels that may have an undue impact on regression fit parameters. The histogram in the top panel indicates the number of pixels at a given leverage. $N_{\rm tpl} = 27$, $N_{\rm sys} = 11$, $\sigma^2_{\rm sys} = 0.01$. 78

3.9 Error in overdensity estimates for different cleaning methods, binned in deciles of the true overdensity and with points offset for clarity. The top plot includes error bars indicating the standard deviation of pixel errors in each bin, while the bottom plot is a zoomed-in version to better display how the means deviate from zero. Overcorrection at the map-level is only significant for Template Subtraction, which under-estimates the magnitude of both peaks and voids, while other methods are very close to unbiased. See text for details. 84

3.10 *Left*: Error in map reconstruction for each method as a function of multipole ℓ in a DES-like survey, shown as the deficit in correlation at each multipole between the true and cleaned maps ($1 - C_\ell^{s\hat{s}}/C_\ell^{ss}$). A perfect reconstruction corresponds to 0, whereas pure noise corresponds to 1. For all methods (except perhaps Template Subtraction), the cleaned map is a good approximation of the true map for cross-correlation purposes, especially at scales $\ell \gtrsim 30$. *Right*: Error in power spectrum estimation, shown as the residual angular power relative to sample variance ($\tilde{C}_\ell^{\rm est} - \tilde{C}_\ell^{ss}$)/$\sigma_{C_\ell}$ in bins of $\Delta\ell = 10$. Solid lines indicate means of cleaning performed on 50 signal realizations of each bin and shaded regions indicate the central 68% probability mass of the 250 total realizations. The multiplicative correction applied to Mode Projection removes most of the bias of the method (green to orange). Here we use 27 templates, of which 11 are contaminating the data. 85

3.11 Bias in the angular power spectrum, $\Delta\chi^2_{C_\ell}$, as a function of the number of templates fit to the map. We consider Gaussian templates which have a correlation of $\rho_{tpl} = 0.2$ within each of the four template classes, as defined in Sec. 3.6.1. We generate two template realizations per class with which to contaminate each signal map ($N_{\rm sys} = 2$, denoted by the vertical dotted line). The templates used to perform the cleaning vary from one to 48 for each type of template spectrum, for a total of four to 196 templates, with new realizations generated for each $N_{\rm tpl}$ (this is the source of the noise in the 'Uncleaned' line). The Template Subtraction, Mode Projection, and Forward Selection methods are all mildly susceptible to overfitting — signaled by the increase in $\Delta\chi^2_{C_\ell}$ for $N_{\rm tpl} > 2$ — though only Template Subtraction to a degree where it overcomes the penalty for neglecting a contaminating template ($N_{\rm tpl} = 1$). For the *additive* Mode Projection method, $\Delta\chi^2_{C_\ell}$ is dominated by the bias from not addressing the multiplicative contribution to the power spectrum (see Fig. 3.10, right panel), while the other methods are dominated by increased variance from chance correlations. The bias from failing to correct for the multiplicative term dominates even when fitting for ~ 200 templates. The DES-Y1 and Elastic Net display a lesser dependence on $N_{\rm tpl}$, and so are more robust to overfitting. See Sec. 3.7.2 for details. 87

3.12 Error in the angular power spectrum, $\Delta\chi^2_{C_\ell}$, as a function of the level of cross-correlation imposed between the templates within the same class. We assumed contamination from two realizations from each of the four classes (i.e. $N_{\rm sys} = 8$). The left panel assumes cleaning with only one of the contaminating templates from each class, while in the right panel we clean for four templates from each class, including the contaminating ones. Note that in the case where template correlation $\rho_{\rm tpl} \to 1$, the two templates are identical and it is equivalent to cleaning only for one contaminating templates, an ideal scenario. In the right panel we see that while the DES-Y1 outperforms others when templates are completely orthogonal, it suffers as the level of correlation between templates increases. The Elastic Net method mitigates this problem. 89

4.1 Matter power spectrum at the Planck ΛCDM best-fit values (black) and with a non-zero second running of $\beta_s = 0.03$ (green). The largest and smallest scales are those most sensitive to a change in the spectral runnings. While the information content from large scales is limited by cosmic variance, small-scales in the nonlinear regime offer an intriguing means to further constrain the spectral runnings. Shading indicates scales that are increasingly in the nonlinear regime (above $k \gtrsim 0.1h\,\mathrm{Mpc}^{-1}$). Adapted from Ref. [27]. 99
4.2 Galaxy number density $n(z)$ of Euclid and DESI in each redshift bin. The features in the DESI $n(z)$ are due to the fact that the sample is a combination of several populations of sources. 106
4.3 Error (1σ here and everywhere) in the spectral running α_s as a function of $k_{\rm max}$, evaluated for our fiducial Euclid-like survey, assuming $\beta_s = 0$. The legend shows our assumption about modeling of the systematics, while T_{k_α,k_β} refers to the inclusion of the trispectrum to the data covariance. Note that here and in subsequent plots, the value of the running denoted as the "Inflationary prediction" (purple horizontal line) is only approximate. 108
4.4 Marginalized constraints on α_s when combining information from different surveys (DESI-like, Euclid-like, CMB-S4, and also in combination). Solid curves include the MFP description of the systematic errors in galaxy surveys (see Eq. (4.8), while the dashed curves do not. Results using the Mead parameterization are similar to the No Nuisance (No Nuis) case, and so we we omit them here for clarity. 109
4.5 Relative difference between the nonlinear predictions from two popular fits: that of Takahashi et al [6] and of Mead et al [28] scaled so as to correspond to about 1% maximum difference at small scales. The quantity shown, $\delta P/P = 0.2(P_{\rm taka} - P_{\rm mead})/P_{\rm mead}$, is the fiducial model for the small-scale systematics that we employ in subsequent plots to gauge the protection offered by our systematics parametrizations. 112
4.6 1-σ statistical errors (solid curves) and bias (dashed) in the first spectral running, as a function of $k_{\rm max}$. We adopt the Euclid+CMB-S4 combination of surveys. The legend on top denotes three alternate assumptions about the systematic error modeling: none (blue), Mead (red), and MFP (black). . . . 113
4.7 Similar to Fig. 4.6, except now β_s is allowed to vary. The left panel shows the 1-σ error and parameter bias in α_s as a function of $k_{\rm max}$, while the right panel shows the same for β_s. The curves have the same meaning as in Fig. 4.6. . . 115

B.1 Distribution of pixel overdensities across all 100 realizations of the lognormal (orange) and Gaussian (blue) maps of the galaxy overdensity in the lowest redshift bin of our fiducial DES-like survey. The Gaussian maps contain pixels with $s < -1$, which is nonphysical in cases like this where s corresponds to an overdensity. 132

B.2 Box plot showing the performance of each cleaning method when using Gaussian (blue, left) versus lognormal (orange, right) signal maps, as measured by $\Delta\chi^2_{C_\ell}$ of the power spectrum. Filled boxes show the 25-50-75% quartiles, with whiskers encompassing the rest of the distribution out to 1.5× the interquartile range. Points beyond this range are indicated by diamonds. Regardless of whether lognormal or Gaussian maps are used, the relative performance of the methods to one another is largely unchanged, and the Gaussian approximation is negligible compared to neglecting the multiplicative correction of Sec. 3.4.2. 133

B.3 Impact of prewhitening before cleaning with the multiplicative and additive versions of the Mode Projection method on 1000 realizations for our fiducial contamination model. The standard Mode Projection method assumes a flat power spectrum for the target signal, resulting in a suboptimal estimate of contamination. This can be improved through 'prewhitening' the data vector and templates using a prior power spectrum, which can be shown to be equivalent to a standard weighted regression procedure in harmonic space. There is clear but modest improvement from the standard case (blue) to the nearly-optimal, prewhitened case (orange), with the most improvement seen for realizations that have large error. This can be seen by the preferential reduction of extreme points at the high end of the box plots in the prewhitened case (note the log scale). 138

B.4 Validation tests of the Mode Projection map-cleaning procedure. *Left panel:* Comparison of Mode Projection performance on an additive-only contaminated map (as assumed by the Mode Projection method), using `NaMaster` (blue) and our own implementation (orange). The agreement between the two is very good. *Right panel:* Impact of not pre-centering cleaning templates in `NaMaster`. The blue curve indicates the standard use case, where contamination is additive and completely described by the templates, which have been individually centered at zero. If templates are instead centered at another value (here we add a constant $2\sigma_{\mathrm{tpl},i}$ offset to each template, where $\sigma_{\mathrm{tpl},i}$ is the standard deviation of values in template map i. Adding a monopole template completely mitigates the bias from non-centered templates. 139

B.5 Same as Fig. 3.11 but showing the RMSE in the estimated overdensity map for each method, rather than error in the power spectrum. Trends are very similar. See Sec. 3.7.2 for details. 139

B.6 Dependence of the power spectrum error ($\Delta\chi^2_{\mathrm{C}_\ell}$) on the level of contamination σ^2_{sys} (x-axis), and on the stopping criterion $\Delta\chi^2/\Delta\chi^2_0$ used for the DES-Y1 method (colors). Points are offset for clarity. For comparison, the variance across pixels from the true overdensity in each bin ranges from $\sigma^2_{\mathrm{sig}} \in [0.075, 0.122]$ for the 5 redshift bins of our fiducial survey, corresponding to factors of 7.5 — 1220× larger than σ^2_{sys} for the points shown. 140

C.1 Constraints on the spectral running α_s for a hypothetical survey with $n(z) \to 1000 \times [n(z)_{\mathrm{Euclid}}]$ alone (left, compare to Fig. 4.3) and with a CMB-S4 experiment (right, compare to Fig. 4.6). While the LSS constraints improve with the increased number density, the trispectrum still limits the information that can be gained from nonlinear scales of $k \gtrsim 0.6h\,\mathrm{Mpc}^{-1}$ (*left*, dashed vs. solid). If $P(k, z)$ is mismodeled, then only the MFP prescription (black) improves constraints over CMB-S4 before significantly biasing the results. . . 142

C.2 Other systematic shifts in $P(k)$ tested to verify that the results of Sec. 4.4 are robust to choice of $\delta P(k)$. Note that because Eq. (4.17) is linear in $\delta P(k)$ and we are interested in where $|\delta|/\sigma = 1$, the overall sign of $\delta P(k)$ is inconsequential. 143

C.3 Parameter bias from different $\delta P(k, z)$ for ΛCDM+α_s + β_s using Euclid + CMB-S4 for α_s (top) and β_s (bottom). The 1σ uncertainty is in black and columns correspond to different nonlinear prescriptions from Sec. 4.2. 143

List of Appendices

Appendix A: ISW Reconstruction Appendices **124**
A.1 Equivalence with Other Estimators 124
A.2 Estimating ρ with $R_\ell(\tilde{C}_\ell)$ 127

Appendix B: LSS Systematics Mitigation Appendices **131**
B.1 Lognormal vs. Gaussian Signal Maps 131
B.2 Effect of Prewhitening 133
B.3 Comparison with `NaMaster` 135
B.4 Accounting for the Monopole 136
B.5 Map Error when Varying Number of Templates 137
B.6 Impact of $\Delta\chi^2/\Delta\chi^2_0$ on DES-Y1 Analysis 137

Appendix C: Chasing the Spectral Runnings Appendices **141**
C.1 Increasing the Number Density $n(z)$ 141
C.2 Robustness of Results to Choice of $\delta P(k, \mu, z)$ 142

Abstract

Cosmology as a field has made rapid progress over the last 30 years, driven in large part by a massive increase in observational data. Large-scale structure (LSS) surveys have played a key role in this rise of 'Precision Cosmology', having grown from mapping the location of thousands of galaxies to the hundreds of millions. As statistical errors plummet, however, progress in the field increasingly hinges on our ability to model and control systematic errors to an exquisite degree. Here we investigate the impact of systematic errors in LSS surveys in three different cosmological contexts. These studies progress from quantifying the level of calibration necessary to accurately perform specific analyses to developing practical methods to correct for systematics in order to achieve such levels of calibration.

First we present a detailed study of the Integrated Sachs-Wolfe (ISW) effect and how errors in the large scale photometric calibration of LSS surveys impact estimation of the ISW signal. The ISW effect is an imprint of dark matter in the Cosmic Microwave Background and contains important information about dark energy, including possible signatures for modifications to General Relativity. We quantify the necessary levels of calibration to produce accurate reconstructions of the ISW map and power spectrum for next-generation surveys. We provide a roadmap for ISW reconstruction, including the optimization of survey configuration and an improved estimator to render the analysis more robust to calibration errors.

Next, we perform a detailed study of the leading methods for removing spatially-dependent systematic errors in galaxy surveys, such as those induced by interstellar dust, variable atmospheric conditions, and other effects that modulate the observed number of galaxies across the sky. We recast them into a common statistical framework, elucidating assumptions implicit within each method and characterize their performance on a suite of simulations. We propose extensions to current methods that are more robust, simpler to implement, and exhibit greater suppression of systematic errors. We further derive uncertainty estimates for the galaxy-level corrections, enabling the propagation of errors from the correction methods into the LSS galaxy catalogs and any subsequent analyses that use them.

The final portion of this thesis focuses on small scale systematic errors in LSS analyses,

such as arise from theoretical uncertainties in the non-linear growth of dark matter, baryonic effects, and other astrophysical phenomena. We characterize how errors in modeling such small scales impact our ability to accurately infer the primordial power spectrum of curvature fluctuations, which initially seeded structure in the early Universe. We show that if unaccounted for, current and predicted modeling uncertainties can strongly bias measurements of the "runnings" of the spectral index, key parameters for testing single-field slow-roll models of inflation, thought to be responsible for the rapid, early expansion of the Universe. We compare methods designed to mitigate such small-scale systematic errors and demonstrate that, even with optimistic improvements in small scale modeling, only exotic models of inflation will be testable via constraints on the runnings from near-future LSS surveys.

These three studies represent important steps for continued progress in the field and towards ensuring that analyses of large-scale structure are robust and accurate in the era of Precision Cosmology.

Chapter 1

Introduction

Almost every culture has a cosmology — a description of the natural order of Universe and how it came to be. Humans have always sought to understand our place in the cosmos, but over the last ∼100 years, theoretical and technological developments have transformed this pursuit into a rigorous science. We have developed a mathematical framework to describe the dynamics and evolution of the Universe on the largest scales under different physical models, and the technological tools that enable us to test these models against a growing body of observational data that extend all the way to the earliest moments of the Big Bang. In the last 30 years in particular, leaps in observational capabilities have ushered in the era of so-called 'Precision Cosmology' and a standard cosmological model with sub-percent-level constraints on the parameters thought to describe the composition and evolution of the Universe. Large-scale structure (LSS) surveys have a played a key role in the birth of this new paradigm, having grown from mapping the locations of thousands of galaxies around us to hundreds of millions, allowing us to precisely test our understanding of how structure grows and how the Universe evolves over time. This precision comes with formidable challenges, however; as *statistical errors* continue to plummet, there is increasing risk that unmodeled, non-cosmological effects can significantly bias results and masquerade as new physics. Progress in the field increasingly relies on our ability to detect, characterize, and mitigate the impacts of such *systematic errors* in our analyses. In this thesis, we investigate the impact of systematic errors in LSS surveys in three different cosmological contexts. The results of these studies will inform and improve our ability to probe the cosmos with the unprecedented volume of data from upcoming large-scale structure surveys.

In this introduction, we give a brief overview of key facets of modern cosmology, including our current understanding of the forces governing the dynamics of the Universe as a whole and the assumptions and observations on which that understanding is based. We provide an abbreviated history of key events in the Universe's evolution and some of the mathematical

framework used to connect observational data to our theoretical models, which will be useful background for the body of the thesis beginning in Chapter 2. Much of this information can be found in introductory cosmology texts, e.g. Refs. [29–33].

1.1 The Standard Model

The standard ΛCDM model of cosmology describes an accelerating Universe dominated by an energy inherent to the vacuum of space itself (parameterized by Λ) and cold dark matter ("CDM"), a non-relativistic form of matter that does not interact electromagnetically. These comprise approximately $\sim$70% and $\sim$25% of the total energy density in the current Universe, with only $\sim$5% making up the so-called "baryonic" matter that is the basis of everything we know (planets, stars, coffee, etc.)[1] We can infer this because of the theoretical groundwork laid by Einstein and other early cosmologists 100 years ago, enabling us to relate the shape and time evolution of the Universe to its energetic composition.

In a seminal 1917 paper, Einstein applied his recent general theory of Relativity to the cosmos at large, relating mass and energy to the dynamics of spacetime itself. Believing (without evidence) that the Universe should be static, he introduced a cosmological constant Λ to perfectly counteract the gravitational pull of a matter-filled Universe.[34]

Building on this work, Soviet physicist Alexander Friedmann found exact solutions for Einstein's field equations for a Universe that is homogeneous (translationally invariant) and isotropic (directionally invariant). These assumptions lead to a spacetime metric

$$ds^2 = -c^2dt^2 + a(t)^2\left[\frac{dr^2}{1-kr^2} + r^2d\theta^2 + r^2\sin^2(\theta)d\phi^2\right], \tag{1.1}$$

where k encodes the curvature of space today[2], c is the speed of light in a vacuum, and the expansion or contraction of the universe as a whole is permitted and captured by an overall *scale factor* $a(t)$. $a(t)$ then relates *proper* distances affected by Universal expansion to *comoving* distances where this expansion is factored out, and the two are set to be equivalent at present day (t_0) with $a(t_0) = 1$.

Eq. 1.1 is known as the Friedmann-Lemaître-Robertson-Walker (FLRW) metric, in honor of its independent derivation by Friedmann's contemporaries. Much of modern cosmology is built on the twin assumptions of large-scale homogeneity and isotropy of the Universe, such

[1] Note the term "baryonic" would *normally* refer only to matter that is made up of an odd number of quarks, but cosmologists have a long history of defying norms so pretty much any non-relativistic standard model particle counts.

[2] Just like the 2D surface of the earth is (positively) curved, with parallel lines eventually converging, 3D space can be curved as well, with positive, zero, and negative curvature corresponding to the sign of k.

that they are collectively termed the *cosmological principle*. Testing the validity of these assumptions is still an active area of research (e.g. Ref. [35]).

Friedmann derived the equations governing the time evolution of the scale factor given General Relativity, which are now referred to as the Friedmann equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho(t) + \frac{\Lambda c^2}{3} - \frac{kc^2}{a(t)^2} \tag{1.2}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \tag{1.3}$$

where G is a universal gravitational constant, Λ is Einstein's cosmological constant, and $\rho(t)$ and p are the mass density[3] and pressure of the Universe, both of which serve to decelerate Universal expansion.

A third Friedman equation known as the continuity equation can be derived from the first two and relates how the density changes at any given point in time as a function of density and pressure:

$$\dot{\rho} = -3\left(\frac{\dot{a}}{a}\right)\left(\rho(t) + \frac{p(t)}{c^2}\right) \tag{1.4}$$

This relation holds not just for the Universe as a whole, but also for individual and distinct components that contribute to the total density, such as matter, radiation or dark energy. These different components effectively act as distinct, perfect fluids and have different equations of state w relating their pressure to energy density ($p = w\rho c^2$). Note that while Eq. 1.2 is only impacted by the density of each component, the acceleration given by Eq. 1.3 also depends on pressure and thus the equations of state.

To understand how different components affect the expansion rate, it is useful to consider simplified cases with only one component. Using the 0 subscript to denote a quantity evaluated at present day, we can write the time evolution of the density of a single-component Universe with constant equation of state as

$$\rho(t) = \rho_0 a^{-3(1+w)}. \tag{1.5}$$

Non-relativistic matter has essentially no pressure and thus equation of state $w_{\text{matter}} \approx 0$, resulting in $\rho(t) = \rho_{M,0}/a(t)^3$, the expected relation of density scaling as the inverse of volume. Radiation is somewhat less familiar — in addition to the inverse-volume scaling of the density of photons, the expansion of the Universe also stretches wavelengths of light,

[3] We will use the terms energy density and mass density somewhat interchangeably, recognizing that they are related via a simple scaling of c^2.

inducing a so-called *redshift*, which results in the energy density of radiation dropping as a^{-4}, with $w_{rad} = 1/3$.

While Eqs. 1.2 and 1.3 treat the cosmological constant as a separate entity, they can be reformulated to treat it as another energy component with density $\rho_\Lambda = \Lambda c^2/8\pi G$. This has no time dependence, implying $w_\Lambda = -1$ (c.f. Eq. 1.5). This is the sense in which the cosmological constant is interpreted as a type of "dark energy," and in this case where the energy density does not change with a, it can be interpreted as an energy inherent to the vacuum of space itself.[4] So far, the ΛCDM model provides a very good fit to existing data, but it is possible that dark energy does not correspond to a vacuum energy and is instead some other phenomenon. A more general parameterization allows the equation of state to vary (typically denoted simply w), or to even be time dependent such as via the CPL parameterization [37, 38] $w(a) = w_0 + w_a(1-a)$. Allowing either w or w_0 and w_a to vary are two commonly-tested alternative models to ΛCDM, in the hopes of learning more about the nature of dark energy.

We now define several parameters in order to rewrite Eq. 1.2 in a form that contains the cosmological parameters as they are typically reported. The Hubble parameter characterizes the relative expansion rate of the scale factor,

$$H(a) \equiv \frac{\dot{a}}{a}, \tag{1.6}$$

while the *critical density* of the Universe today is that required to ensure the Universe is flat (i.e. $k = 0$):

$$\rho_{crit,0} = \frac{3H_0^2}{8\pi G}. \tag{1.7}$$

The energy density is usually reported as

$$\Omega_i = \rho_{i,0}/\rho_{crit,0}, \tag{1.8}$$

where i corresponds to matter, radiation, or dark energy. A flat Universe (as observations suggest ours nearly is) then has $\Omega_{TOT} \equiv \Omega_M + \Omega_r + \Omega_{DE} = 1$.

Eq. 1.2 can then be written as

$$H^2(a) = H_0^2\left(\Omega_M a^{-3} + \Omega_r a^{-4} + \Omega_{DE} a^{-3(1+w)} - \Omega_K a^{-2}\right), \tag{1.9}$$

where we have included the more general parameterization of dark energy with a free equation

[4] It is worth noting that quantum field theory predicts a vacuum energy, but one that is larger than dark energy by some 120 orders of magnitude. Physicists have actually puzzled over why we don't observe such a strong vacuum energy since the 1980s, before dark energy was even discovered [36].

of state, and it can be seen that in this Friedman equation, the effect of curvature is as an effective energy component whose density dilutes as a^{-2}.

This parameterization neatly separates the Hubble parameter today (H_0, aka the Hubble constant) from its time evolution. It also explicitly shows how the different energy densities evolve with scale factor. A consequence of the different rates of dilution is that a multi-component Universe will go through different epochs, where the behavior of $H(a)$ is largely governed by whichever term in Eq. 1.9 is dominant, and *when* these epochs occur is governed by the energy densities today. Eq. 1.9 forms the basis of using measurements of the expansion history to constrain the cosmological parameters.

Using Eqs. 1.6 and 1.9, it is straightforward to solve for the time dependence of the scale factor in a Universe that is dominated by each component (assuming a flat Universe, $\Omega_k = 0$). Table 1.1 shows this along with other characteristics of single-component epochs.

Dominant Component	w	$\rho(a)$	$a(t)$	Redshifts Dominant
Radiation	1/3	a^{-4}	$t^{1/2}$	$z \in [?, 3500]$
Matter	0	a^{-3}	$t^{2/3}$	$z \in [3500, 0.33]$
Cosmological Constant (Λ)	-1	a^0	e^{Ht}	$z \in [0.33, ?]$

Table 1.1: Characteristic behavior of the density and scale factor in a flat Universe for various epochs where a single component dominated the energy budget.

The scale factor a is neatly related to a quantity that is directly observable: the cosmological redshift z. As space expands, it stretches the wavelengths of light traveling through it, shifting them to lower frequency. Thanks to quantum mechanics, the precise and unique patterns of allowed electron energy transitions within atoms allow for absorption and emission spectra to function as extremely good identifiers of far-off elements even when their absolute frequencies are shifted. By comparing the observed spectra to corresponding spectra at rest, the redshift can be very precisely determined: $1 + z = \lambda_o/\lambda_e$, where λ_o and λ_e correspond to the observed and rest-frame emitted wavelengths, respectively. The cosmological redshift of an object[5] is related to the scale factor of the Universe at the time the light was emitted by $1 + z = a(t_0)/a(t_e) = 1/a(t_e)$. The observability of redshift (and thus the scale factor) makes it much more useful than time for characterizing "when" events happened; the associated time t is model dependent. For the very early Universe, it is also common to use thermal temperature or energy, related to the scale factor as $T = E/k_B \propto 1/a$, as these are more relevant quantities for characterizing which particles are in equilibrium with the hot plasma

[5]While the peculiar velocity of an object also causes a Doppler shift of roughly $z_{pec} \approx v_{los}/c$ (ignoring relativistic corrections), this is quickly dominated by the cosmological redshift for $z \gtrsim 0.01$. Nevertheless, the effect is still important and even contains cosmological information, which is captured in measurements of so-called *Redshift Space Distortions*.

of the early Universe.

1.2 Redshift and the Expanding Universe

The concept of redshift has served a pivotal role in the development of cosmology. In an expanding Universe, all objects will appear in general to recede from an observer with velocity $v = H_0 d$, where d is the proper (i.e. instantaneous) distance to the object and H_0 is a constant that characterizes the expansion rate. For small redshifts ($z \ll 1$), this is well-approximated by $z \approx H_0 d/c$, such that the expansion rate H_0 can be determined if one can measure both the redshift and distance to nearby objects in the "Hubble flow" of cosmological expansion. It was exactly this technique that gave the first evidence for an expanding Universe.

While Georges Lemaître first proposed what is now called "Hubble's Law" relating distance to recession velocity for an expanding Universe [39], it is named for Edwin Hubble who provided the first observational evidence that our Universe was undergoing such expansion. Hubble used Cepheid variable stars as standard candles[6] to estimate the distance to so-called nubulae. He found that such nebulae were far outside the Milky Way, and were in fact entirely distinct galaxies [41] Using redshifts for the nebulae that had been meticulously determined by Vesto Slipher in the years prior [42], he found that in addition to being very far away, all such nebulae were *receding*, suggesting that the Universe was in fact expanding. Hubble's discovery, coupled with the theoretical framework laid out by Einstein, Friedmann, Lemaître and others, led to the serious consideration of a very early Universe that was much more dense, and began with what has come to be known as the Big Bang.[7]

1.3 History of the Universe: The Executive Summary

The Big Bang model has firmly established itself since it was first proposed in the 1920s and, coupled with the theory of *Inflation*, it forms the basis of the ΛCDM model. The Big Bang

[6] A *standard candle* is any object whose absolute luminosity can be determined, such that the difference from its *apparent* (observed) luminosity can be used to estimate its distance. In 1908, Henrietta Leavitt discovered that Cepheid variables pulsate with a period that is related to their absolute luminosity, enabling Cepheids to be used as standard candles. This same technique for measuring H_0 is used today, though now Cepheid variables make up just a single rung in a cosmic "distance ladder," used to calibrate distances to even deeper probes such as Type Ia supernovae (e.g. Ref. [40]).

[7] An alternative hypothesis that had traction was the so-called *Steady State* model, in which the Cosmological Principle held for *all time*, and thus the Universe had no beginning or end, and new matter was continuously created to maintain density in an expanding Universe (see e.g. [43, 44]). The discovery of the Cosmic Microwave Background in 1964 largely put the Steady State model to rest.

deftly explains not only the expansion of the Universe, but also the cosmic abundances of the lightest elements and the existence of the black body Cosmic Microwave Background radiation (CMB) that we see in all directions. ΛCDM fills in the details of what happened after the Big Bang, explaining the variations in the CMB and evolution of large-scale structure, while Inflation explains the initial conditions.

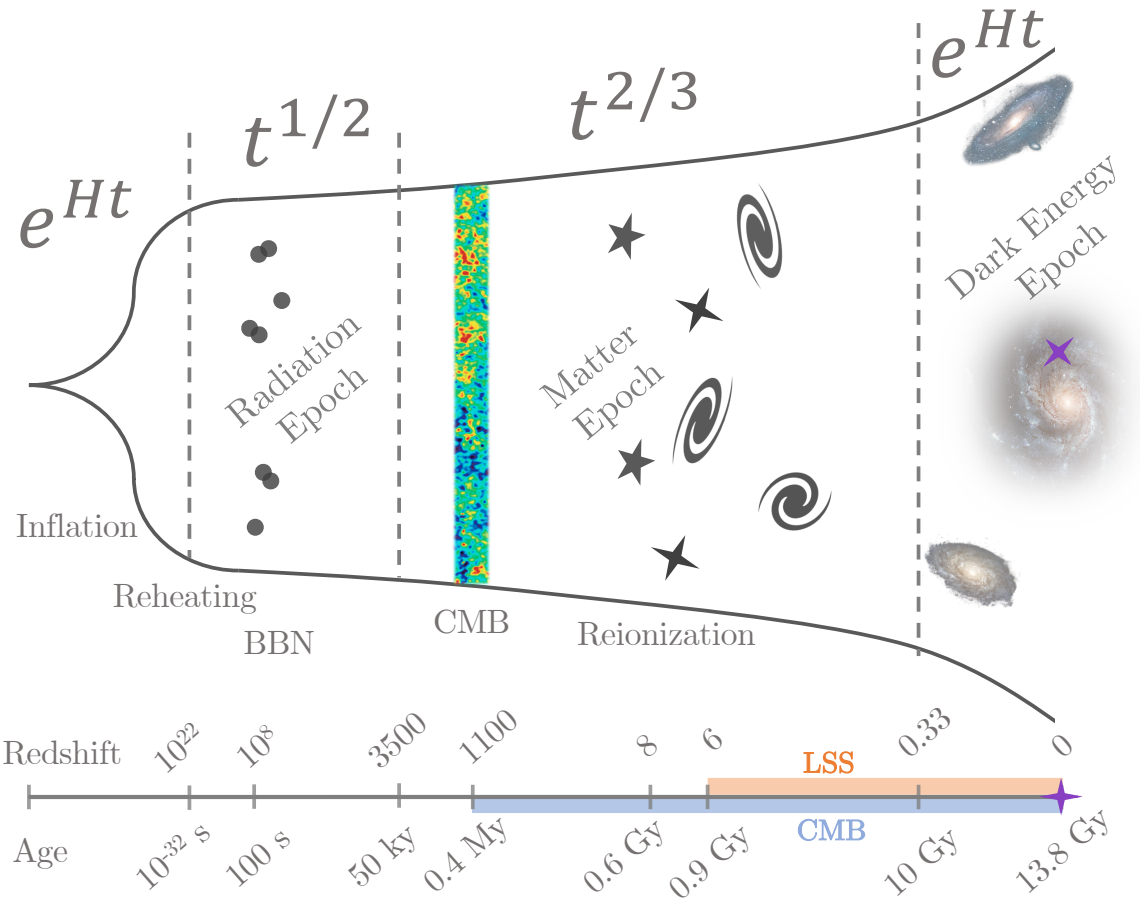

Figure 1.1: Schematic history of the Universe leading up to now (purple star). Dashed lines indicate transitions between epochs dominated by a single matter-energy component, with labels and symbols to indicate key events. The approximate times and redshifts of these are given by the timeline at bottom, which also shows the ranges probed by LSS (red) and CMB (blue) experiments. The time dependence of the scale factor during each epoch is indicated at top.

Fig. 1.1 shows a timeline of some of the key events in the Universe's 14 billion year history. The very earliest moments up to $t \sim 10^{-12}$ seconds are still speculative, but it is believed that the Universe underwent a period of exponential expansion known as *inflation*, during which the Universe increased in size by > 25 orders of magnitude. When inflation ceased, the field responsible decayed and filled the Universe with hot radiation in a process known as *reheating*, starting what is typically known as the (hot) Big Bang. As the Universe expanded, it cooled,

causing particles to fall out of equilibrium with the primordial plasma as their decreasing interaction rates fell below the expansion rate. Within the first few minutes, atomic nuclei fused together in a process known as *Big Bang Nucleosynthesis* (BBN) — detailed calculation of the reaction rates during this process of expansion and cooling produce remarkably good predictions of the measured abundances of primordial hydrogen, helium, deuterium and lithium (see [45] for a review.).

The energy density of matter decayed more slowly than that of radiation, marking a transition around $z \sim 3500$, after which the energy budget became dominated by non-relativistic matter. Overdense regions of dark matter began to grow appreciably under gravity and the rate of expansion increased. At $z \sim 1100$ ($t \sim 380,000$ yr), photons cooled enough to decouple from baryons, allowing neutral hydrogen to stably form. Photons free-streamed until the present in the form of the Cosmic Microwave Background, providing a detailed snapshot of the Universe at that time. Baryons were then free to fall into the gravity wells of dark matter and began to form the first stars and galaxies around $z \sim 20$ ($t \sim 200$ My). Eventually, they produced enough energy to re-ionize the hydrogen-dominated Universe by $z \sim 8$ and continued to aggregate into larger clusters of galaxies, tracing extended filaments of large-scale structure throughout the Universe. The dark energy epoch began around $z \sim 0.33$, initiating another period of accelerated expansion with the scale factor going as $a(t) \propto e^{Ht}$ and slowing the growth of structure. We now discuss some of these events in greater detail and introduce some of the mathematical formalism for characterizing their impact on cosmological observables, which will be used throughout the rest of this work.

1.4 Inflation and Primordial Perturbations

It is currently thought that the initial density perturbations that eventually grew into the large scale structure we see today were seeded by quantum fluctuations of a primordial quantum field at around $t \sim 10^{-34}$ seconds. The presence of this field would cause exponential and superluminal expansion, which separated regions that were previously in causal contact, smoothing out inhomogeneities, reducing the observed curvature, and diluting any exotic features that were generated before the time of inflation (such as magnetic monopoles). All of these solve certain observational problems that were unexplained before the theory of inflation was proposed, and they place lower limits on how much inflation must have occurred.

The quantum fluctuations of the supposed "inflaton" field got enlarged to macroscopic scales, imprinting spatial inhomogeneities of energy density which, once inflation ended, decayed into the physical fields and particles we see today. This results in the inflationary

prediction of adiabatic density fluctuations, wherein the different components (e.g. matter and radiation) have coherent density fluctuations about their respective means. Generically, we can parameterize fluctuations from the mean for some field i using the *overdensity*:

$$\delta_i(\boldsymbol{r}, t) \equiv \frac{\rho_i(\boldsymbol{r}, t) - \bar{\rho}_i(t)}{\bar{\rho}_i(t)}, \tag{1.10}$$

or its Fourier transform:

$$\delta_i(\boldsymbol{k}, t) = \int \frac{d^3\boldsymbol{r}}{(2\pi)^{3/2}} \delta_i(\boldsymbol{r}, t) e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{1.11}$$

The power spectrum characterizes the amplitude of fluctuations at different wavenumbers $\boldsymbol{k}$ as

$$\langle \delta(\boldsymbol{k}, t) \delta^*(\boldsymbol{k}', t) \rangle = (2\pi)^3 P(k, t) \delta_D^3(\boldsymbol{k} - \boldsymbol{k}'), \tag{1.12}$$

where $\langle ... \rangle$ is the ensemble average over realizations of the field and $\delta_D^3(\boldsymbol{k} - \boldsymbol{k}')$ is the Dirac delta function (a result of the statistical isotropy and homogeneity of δ). We will make extensive use of overdensities and power spectra throughout this work.

Most simple inflation models produce primordial curvature fluctuations that are nearly Gaussian and have a power spectrum[8] that is almost scale-independent. It is common to parameterize the power spectrum of the scalar curvature perturbations as a power law that is Taylor expanded about a pivot scale k_*:

$$\frac{k^3}{2\pi^2} P_s(k) = A_s \left(\frac{k}{k_*} \right)^{(n_s - 1) + \frac{1}{2}\alpha_s \ln(k/k_*) + \frac{1}{6}\beta_s (\ln(k/k_*))^2 + ...}, \tag{1.13}$$

where A_s is the *scalar amplitude* of the fluctuations, n_s is the *spectral index*, and $\alpha_s = (dn_s/d\ln k)|_{k^*}$ and $\beta_s = (d^2 n_s/d(\ln k)^2)|_{k^*}$ are the *spectral runnings*. The parameters n_s, α_s and β_s depend on the potential of the inflaton field, $V(\phi)$, such that their measurement gives crucial insight into the form of the field that seeded the primordial density fluctuations.

In the simplest inflationary models where inflation is generated by a single field that is slowly rolling down its potential $V(\phi)$ (rather ingeniously called *single-field slow-roll* inflation), it can be shown that inflation generally predicts a scalar index slightly less than one:

$$n_s = 1 - 6\epsilon + 2\eta + \mathcal{O}(\epsilon^2, \eta^2). \tag{1.14}$$

ϵ and η are the *slow-roll parameters* characterizing the slope and curvature of $V(\phi)$ re-

[8] A Gaussian distribution is fully characterized by its mean and variance and since $\bar{\delta} = 0$ by definition, if δ is a Gaussian field it is fully characterized by its power spectrum $P(k, t)$. Even if a field is not fully Gaussian, as the second moment of the field the power spectrum often carries a substantial amount of information and is thus a key summary statistic for characterizing the statistics of random fields, as will be evident throughout this work.

spectively, with $\epsilon \ll 1$ and $\eta \ll 1$ required for inflation. The current constraint of $n_s = 0.9649 \pm 0.0042$ [46] is very much aligned with the inflationary paradigm. Chapter 4 investigates the prospects of near-future measurements to detect α_s and β_s at the levels predicted by inflationary models.

A key feature of inflation is that it separates regions that were previously in causal contact through the superluminal expansion of space. The comoving Hubble radius (c/aH) defines the scale at which two particles are in causal contact; inflation is characterized by a *shrinking* Hubble radius, and it is straightforward to show that this occurs whenever there is accelerating expansion of space:

$$\frac{d}{dt}\left(\frac{c}{aH}\right) = -\frac{c}{(aH)^2}\ddot{a} \tag{1.15}$$

thus

$$\ddot{a} > 0 \iff \frac{d}{dt}\left(\frac{c}{aH}\right) < 0, \tag{1.16}$$

which from Eq. 1.3 will occur whenever $\rho + 3p/c^2 < 0$, or $w < -1/3$. Incidentally, with the dark energy epoch we recently entered another phase of accelerating expansion so the comoving horizon is shrinking once again!

Curvature perturbations are generated almost equally across all scales (per Eq. 1.13) and during inflation many of these will expand to the point where their wavelengths are larger than the Hubble radius, such that their peaks and troughs are out of causal contact. These perturbations no longer oscillate or change but instead maintain a constant amplitude. The result is a set of frozen long-wavelength inhomogeneities outside the Hubble radius, which seed coherent oscillations once they "re-enter the horizon" of the Hubble radius after inflation ends.

1.5 Structure Growth

Density perturbations with wavelengths smaller than the Hubble radius are governed by the opposing forces of gravity and radiation pressure in an expanding space, with gravity acting to grow overdense regions and pressure and expansion working to dilute them. Assuming General Relativity holds, then for a non-relativistic fluid such as matter (though potentially embedded in a relativistic background), we can solve for the dynamics where $\delta \ll 1$ using standard fluid equations such as the Euler momentum equation (governing motion) and the continuity equation (ensuring conservation of mass) coupled with the Poisson equation

(relating the gravitational potential to its source mass) to get [29]

$$\ddot{\delta_k} + \underbrace{2H\dot{\delta_k}}_{\text{Hubble drag}} + \left[\underbrace{\frac{k^2c_s^2}{a^2}}_{\text{Pressure}} - \underbrace{4\pi G\rho_M(t)}_{\text{Gravity}}\right]\delta_k = 0. \tag{1.17}$$

δ_k indicates that we are working in Fourier space and $c_s = (\partial p/\partial\rho)^{1/2} = \sqrt{w}c$ is the sound speed at which pressure fluctuations travel. Considering the case where the Hubble drag from expansion is negligible (i.e. $H = 0$), the solutions are straightforward: when the bracketed term is positive, δ_k oscillates, whereas if the bracketed term is negative then there exist both exponentially growing and decaying modes. These correspond to the cases where pressure or gravity dominate the dynamical evolution, respectively, with the separation between modes that experience oscillation vs. growth (or decay) characterized by the Jeans length:

$$\lambda_J \equiv \frac{2\pi a}{k_J} = c_s\sqrt{\frac{\pi}{G\rho_M(t)}} \tag{1.18}$$

The Hubble drag acts as a damping term in each of these regimes.

We can solve this for perturbations of dark matter during different epochs, taking into account that the time-dependence of H changes (c.f. Table 1.1), and taking advantage of the fact that $c_s \approx 0$ for dark matter such that the pressure term is negligible. We find that dark matter perturbations grew[9] as

$$\delta_k(t) \propto \begin{cases} \ln a(t), & \text{Radiation Dominated} \\ a(t), & \text{Matter Dominated} \\ 1. & \Lambda \text{ Dominated} \end{cases} \tag{1.19}$$

In contrast to dark matter modes, baryons were ionized and highly coupled to photons in a photon-baryon plasma at this time, such that baryonic modes oscillated in what is known as *Baryonic Acoustic Oscillations* (BAO) until photon decoupling, at which point radiation pressure dropped to zero. This halted the oscillation of baryons, imprinting a clear characteristic scale on the distribution of baryon overdensities corresponding to the maximum distance the acoustic waves could travel before recombination. This scale can be observed in the distribution of galaxies today and serves as an extremely useful *standard ruler* for measuring the expansion history.

After recombination, baryons fell into the gravitational wells already created by the prior

[9] There are in general two solutions but we are interested in the growing modes, not the static or decaying ones.

growth of dark matter perturbations and continued to grow per Eq. 1.19.[10] The distribution of galaxies (made of baryons) therefore traces the distribution of dark matter, and the mapping between them is parameterized by the so-called *galaxy bias* $b_g(k,z)$:

$$\delta_g = b_g(k,z)\delta_{dm}, \tag{1.20}$$

where b_g depends on the specifics of the galaxy sample, but for large scales can be approximated as independent of k.[11]

The 3D power spectrum of galaxies[12] can then be written in terms of the dark matter power spectrum as

$$P_g(k,z) = b^2(k,z)P_m(k,z), \tag{1.21}$$

and the the dark matter power spectrum related to the primordial power spectrum as

$$P_m(k,z) = \langle \delta_M \delta_M^* \rangle = \frac{4}{25}\left(\frac{kc}{aH}\right)^4 T^2(k,z)P_s(k)T_{nl}(k,z). \tag{1.22}$$

The measured power spectrum of a galaxy sample can thus be used to constrain the physics of the very early Universe. $T(k,z)$ is the transfer function that encodes details of how different modes evolve through the different Universal epochs (as idealized in Eq. 1.19). It is strongly sensitive to the transition between radiation and matter dominated epochs, and thus constrains Ω_M (Ω_r is well measured by the temperature and density of CMB photons), while also being sensitive to Ω_{DE} which slows the growth of perturbations at later times.[13]

$T_{nl}(k,z)$ parameterizes the effects of *nonlinear* collapse. As overdensities grow to $\delta \gtrsim 1$, they decouple from the Universal expansion and begin to grow nonlinearly, violating the linear assumptions that were used to derive Eq. 1.19. This occurs on small scales (high k

[10] Indeed this head-start before recombination from cold dark matter is necessary to achieve the level of structure growth we observe today.

[11] An exception is if there is *primordial non-Gaussianity* in the initial fluctuations, which would result in the appearance of a scale-dependent bias on large scales. This is an active area of research.

[12] The observed galaxy power spectrum is actually anisotropic when using redshift to characterize the radial distance, as the peculiar velocities of the galaxies affect the radial distance measure, but not the transverse distance measure. This results in so-called *Redshift Space Distortions* (RSD), with two main effects: the galaxy sample appears stretched along the line of sight (the *Fingers-of-God* effect) [47], and both over- and underdensities become slightly amplified along the line of sight (Kaiser) [48]. The former is due to orbital velocities of galaxies in gravitationally bound structures (and thus more apparent on small scales) whereas the latter is a result of the tendency for galaxies to fall toward overdense regions. The effects of RSD are significantly reduced for photometric surveys where galaxy redshift uncertainties dominate over the RSD signal. In Chapter 4 we forecast constraints with spectroscopic surveys and so include the effects of RSD on P_g.

[13] $T(k,z)$ is sometimes separated into a k-dependent piece and a separate z-dependent piece known as the *growth function*, in order to more explicitly separate the effects of the matter-radiation transition from the effects of dark energy.

modes) first, and extends to larger scales over time. $T_{nl}(k, z)$ is usually computed using large-scale N-body simulations to inform emulators (i.e. interpolators) or fitting functions such as Halofit [6] and HMCode [28]. Improving models of nonlinear evolution is a highly active area of research, and uncertainties in modeling the nonlinear power spectrum are a key motivator of the investigations in Chapter 4.

1.6 Large-Scale Structure Surveys

LSS surveys put direct constraints on how the distribution of galaxies and matter changes with time, and so have proven invaluable for measuring structure growth and generally testing the theoretical models about the processes that govern the structure, composition, and evolution of the universe. There are two broad classes of LSS survey: spectroscopic and photometric, corresponding to the precision with which redshifts can be determined. The high resolution spectra from spectroscopic surveys like BOSS and DESI require much more time to observe each galaxy, since there are many more (and finer) wavelength "bins" in which photons must be collected, and this limits the number of galaxies that can be targeted for observation. Photometric surveys like SDSS, DES, and LSST instead view large areas of the sky with a few (typically five or six) filters with wide bandwidths, and thus the redshift estimates have much larger errors but the sky area and number of observed galaxies is far greater. Important synergies exist between the two, with overlapping galaxy samples enabling more precise calibration of the photometric redshift estimates.

The precise 3D information from spectroscopic surveys allows $P_g(k, z)$ to be measured very accurately. In contrast, galaxies in photometric samples are typically grouped into redshift bins, with sample redshift distributions used to characterize the radial information and an angular power spectrum computed for that 2D redshift slice. The 2D projected overdensity *map* for redshift bin i can be computed from the 3D overdensity $\delta(\hat{\mathbf{n}}, z)$ as

$$\delta_i(\hat{\mathbf{n}}) = \int \frac{dn_i}{dz} \delta(\hat{\mathbf{n}}, z) b(z) dz \tag{1.23}$$

where $\hat{\mathbf{n}}$ denotes a direction on the sphere and dn_i/dz is the distribution of galaxy redshifts in the bin.

Just as we worked in Fourier space to compute the power spectrum $P(k, z)$ of the 3D density field, we can expand a projected overdensity field into a weighted sum of spherical

harmonic basis functions $Y_{\ell m}(\hat{\mathbf{n}})$:

$$\delta_i(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a^i_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{1.24}$$

where

$$a_{\ell m} = \int \delta(\hat{\mathbf{n}}) Y^*_{\ell m}(\hat{\mathbf{n}}) d\Omega, \tag{1.25}$$

ℓ characterizes the typical wavelength of fluctuations in the basis function $Y_{\ell m}$ (with $\ell \sim \pi/\theta$), and m characterizes the phase. The $a_{\ell m}$ coefficients can then be used as a harmonic space representation of the real-space maps, and this comes with certain useful properties such as a 2D analogue to Eq. 1.12:

$$\langle a_{\ell m} a^*_{\ell' m'} \rangle = C_\ell \delta_{\ell \ell'} \delta_{m m'}, \tag{1.26}$$

where C_ℓ is the *angular* power spectrum, and $\delta_{\ell\ell'}$ and $\delta_{mm'}$ are Kronecker deltas. In other words the covariance of the coefficients is diagonal, thanks to statistical isotropy,[14] and if the 2D field is Gaussian then the spectrum of C_ℓ's contain all of the statistical information about it.

The power spectrum of an observed overdensity field i can be estimated by averaging over the $2\ell + 1$ coefficients for each harmonic ℓ:

$$\hat{C}^i_\ell = \sum_{m=-\ell}^{\ell} \frac{|a^i_{\ell m}|^2}{2\ell + 1}, \tag{1.27}$$

where $\hat{C}_\ell$ indicates an estimate using our single Universe as opposed to the true theoretical quantity that generated it. This can be compared to the theoretical prediction given by

$$C^i_\ell = \frac{2}{\pi} \int dz \int dz' \int dk \left(\frac{dn_i}{dz}\right) \left(\frac{dn_i}{dz'}\right) k^2 P_{gg}(k, z, z') j_\ell(kr(z)) j_\ell(kr(z')) \tag{1.28}$$

where $r(z)$ is the comoving distance at z and $j_\ell(x)$ is a spherical Bessel function of the first kind. It is also common to use the two-point correlation function as an alternative statistic to the observed power spectrum. Both of these "two-point statistics" contain identical

[14] This only strictly holds when the entire sky is observed — when portions of the sky are masked as is usually the case, neighboring harmonic modes become coupled. However this effect can be computed and accounted for, and simply binning in ℓ can largely mitigate the impact.

information, and are related as

$$w(\theta) = \langle \delta(\hat{\mathbf{n}})\delta(\hat{\mathbf{n}}')\rangle = \sum_{\ell=1}^{\infty} \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos\theta) \tag{1.29}$$

where $P_\ell(\cos\theta)$ are the Legendre polynomials.

We make extensive use of angular power spectra in our work with photometric surveys in Chapters 2 and 3.

Other probes from LSS surveys include the aforementioned BAO signal (also captured in the two-point measurements, albeit suboptimally) and *cosmic shear* measurements. The latter uses the shapes and orientations of galaxies as a means to measure the foreground matter distribution via weak gravitational lensing. It's very cool, but not the focus of this work (see e.g. Refs. [49, 50]).

1.7 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is the relic radiation that was released after protons and electrons combined to form neutral hydrogen, allowing photons to free-stream. The radiation is effectively a picture of the Universe at $z \approx 1100$. It is remarkably uniform in all directions, corresponding to a near-perfect black body spectrum with temperature $T = 2.73$K, with typical deviations of $\delta T/T \sim 10^{-5}$ across the sky. An enormous amount of information is encoded in these anisotropies, and yet further information is gleaned from polarization data.

As effectively a 2D field, the angular power spectrum is the typical summary statistic for characterizing the temperature (or polarization) anisotropies of the CMB and comparing them with theoretical predictions. The *acoustic peaks* observed in the spectrum are a clear imprint of coherent oscillations of the photon-baryon fluid before recombination, which themselves were seeded by the primordial fluctuations. Thus the CMB constrains both the shape of the primordial spectrum $P_s(k)$ and the cosmological parameters impacting the oscillatory dynamics. The observed spectra agree remarkably well with the predictions of inflation, which provides the mechanism by which these oscillations were generated and their phases synchronized.

The amplitude of the spectrum constrains A_s, while the tilt is dependent on n_s, α_s and β_s. Baryonic matter is coupled to the fluid, and so affects the peaks differently than does dark matter, allowing relative peak heights to constrain each. Furthermore, the location of the first peak corresponds to the furthest distance sound could have traveled between inflation and recombination. Comparing this transverse distance measure with the angle at

which it is observed and the distance to the CMB allows us to put very tight constraints on the geometry of the Universe, with the most recent observations from the Planck spacecraft suggesting that it is flat to within $\sim$1% [51].

Modes at $\ell \lesssim 200$ were still outside of the Hubble radius at recombination, such that they theoretically provide a pristine measurement of the primordial perturbations projected onto 2D. These large-scale modes get modified however at later times by the presence of dark energy in what is known as the *Integrated Sachs-Wolfe* effect. The main focus of Chapter 2 is how to use LSS surveys to accurately separate the primordial from this late time contribution.

The CMB has been called the cosmic Rosetta Stone because of its ability to test and inform our cosmological understanding. Combining it with other probes such as LSS surveys and Type Ia supernovae allows for dramatically improved constraints. Not only does it serve as an early-universe complement to late-time observations, providing a large lever arm with which to constrain our understanding of structure growth and expansion, but the probes are sensitive to different combinations of parameters, allowing for degeneracies to be broken when combined (Fig. 1.2).

1.8 Connecting Theory to Data

For reference, a standard parameterization of the six-parameter ΛCDM model is given in the top section of Table 1.2, with common extended parameters given in the bottom section. These are typically fixed to the indicated value in ΛCDM, but can be varied in order to test alternative models.

Using observations from LSS, the CMB and other cosmological probes, we can infer these parameters and thus test and improve our current models of the Universe. This happens in two ways:

- sharpening constraints on the parameters of the standard cosmological model, and
- identifying inconsistencies between our observations and the standard model predictions (and potentially identifying an alternative model that rectifies the inconsistency).

If it is the result of new physics, then the latter can herald a paradigm shift where the contemporary model is superseded. This has happened on numerous occasions. The ΛCDM paradigm arose from the discovery of dark energy in 1998, when two independent teams [52, 53] separately measured the Universe's expansion history with Type Ia supernovae and discovered that the Universe is in fact accelerating. The discovery garnered a Nobel Prize.

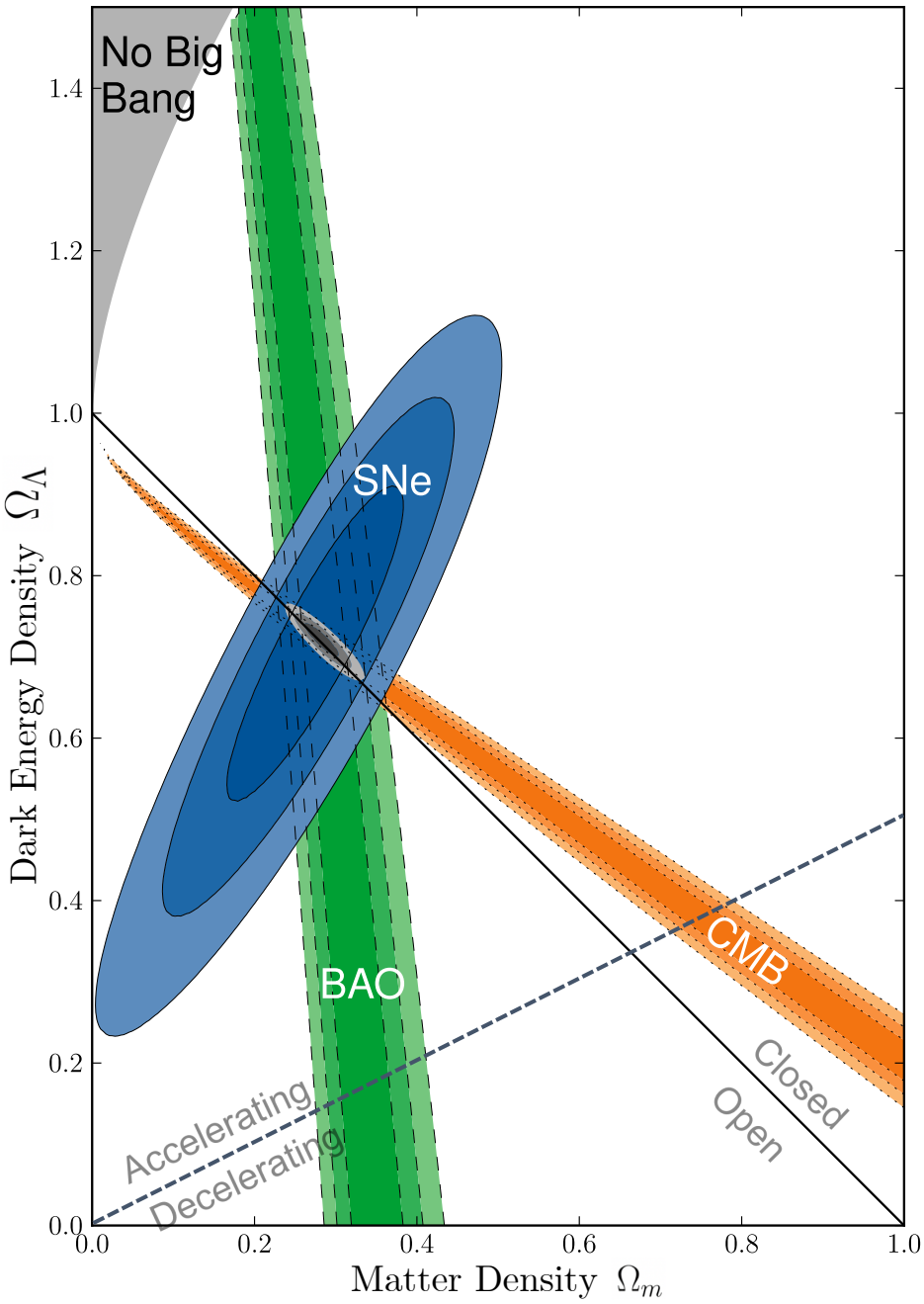

Figure 1.2: 68.3%, 95.4%, and 99.7% credible intervals on the matter and dark energy densities under a ΛCDM model with curvature allowed. Combining data from Type Ia supernovae, CMB, and LSS experiments results in much stronger parameter constraints (gray) than any probe individually, as they are sensitive to different parameter combinations. SNe Ia (blue) show a clear detection of Ω_Λ and an accelerating Universe, BAO constraints from LSS surveys (green) place tight constraints on Ω_m which sources structure growth, and measurements of the CMB (orange) tell us that the Universe is very close to flat, with $\Omega_{TOT} \approx 1$. Adapted from Suzuki et al. [24].

Ω_b	Baryon density
Ω_M	Total matter density
H_0	Hubble constant (km/s/Mpc)
n_s	Scalar spectral index
A_s	Scalar power spectrum amplitude
τ	Optical depth of Reionization
$\Omega_{DE} = 1 - \Omega_r - \Omega_M$	Dark energy density
$\Omega_k = 0$	Curvature energy density
$N_{\text{eff}} = 3.046$	Effective number of relativistic degrees of freedom
$w = -1$	Dark energy equation of state
$w_a = 0$	Slope of dark energy equation of state
$\alpha_s = 0$	Running of the scalar spectral index
$\beta_s = 0$	Second running of the scalar spectral index

Table 1.2: Cosmological parameters. Parameters above the line correspond to the six parameters that define the standard ΛCDM model. Below the line are additional parameters that are varied when testing extended models, but fixed to the reported values in ΛCDM.

However, discrepancies can also arise due to systematic errors in the analysis which can lead researchers astray. The false discovery of B-mode polarization in the CMB by the BICEP2 collaboration in 2014 — also touted as Nobel prize material — offers a cautionary tale. After making international headlines, it was discovered that their 'smoking gun' of inflation was in fact a spurious signal due to the failure to adequately account for foreground dust in the Milky Way.

As the great Carl Sagan noted, *extraordinary claims require extraordinary evidence* [54]. Our observations are mediated by a wide variety of non-cosmological effects related to astrophysics, experimental apparatus, and even researcher bias; it is crucial to eliminate sources of systematic error and ensure that cosmological analyses are robust to a range of unknowns. Incorporating these effects into our models with appropriate parameterization allows us to bring them firmly out of the realm of the "unknown unknown" that could result in systematic error in the analysis and into the realm of the *parameterized* unknown. These can then be included in the model as *nuisance parameters*, which can either be fixed using external data, or inferred jointly along with the cosmological parameters of interest.[15] This requires careful study of which effects may have a significant impact on the cosmological analysis in question.

We now describe the process by which we obtain constraints on the cosmological parame-

[15]Of course what may be systematic effects treated with nuisance parameters in one analysis may be of interest in their own right in another analysis (e.g. the impact of Active Galactic Nuclei on small-scale clustering). One person's systematic is another person's signal.

ters, with a focus on LSS surveys. To constrain the parameters of a model, one must typically be able to compute the *likelihood*, $P(\boldsymbol{d}|\boldsymbol{\theta})$, which gives the probability of the observed data ($\boldsymbol{d}$) as a function of the model parameters ($\boldsymbol{\theta}$). For a frequentist analysis (e.g. Maximum Likelihood approaches) this is all one needs, but most cosmological analyses take a Bayesian approach, in which case one must also specify *prior* probabilities for the model parameters, $P(\boldsymbol{\theta})$. The *posterior* probabilities of the model parameters can then be computed as

$$P(\boldsymbol{\theta}|\boldsymbol{d}) = \frac{P(\boldsymbol{d}|\boldsymbol{\theta})P(\boldsymbol{\theta})}{P(\boldsymbol{d})}, \tag{1.30}$$

where the *evidence* (or *marginal likelihood*) $P(\boldsymbol{d})$ can often be ignored for inferring parameter constraints for a given model, but becomes important when comparing the ability of different models to explain the same data.

It is common to assume that the likelihood is Gaussian in the data[16], with the form

$$\ln P(\boldsymbol{d}|\boldsymbol{\theta}) \propto (\boldsymbol{d} - \boldsymbol{d}_{thry}(\boldsymbol{\theta}))^T \mathbf{C}^{-1} (\boldsymbol{d} - \boldsymbol{d}_{thry}(\boldsymbol{\theta})), \tag{1.31}$$

where $\mathbf{C}$ is the covariance matrix of the data (estimated analytically or numerically, see e.g. Ref. [55]) and $\boldsymbol{d}_{thry}(\boldsymbol{\theta})$ is the model-predicted data vector at $\boldsymbol{\theta}$.

The summary statistics used for $\boldsymbol{d}$ vary depending on the analysis and the cosmological probe but for the standard cosmological analyses using LSS and CMB surveys the approach is typically to construct a data vector made of multiple two-point statistics. For instance, the analysis in Chapter 4 uses the auto and cross-power spectra of CMB temperature and polarization measurements as well as 3D galaxy power spectrum measurements as elements in the data vector. As noted, the work in Chapters 2 and 3 work primarily with angular power spectra of LSS surveys of the type used for photometric surveys like DES.

The observed data vector is computed (e.g. via Eq. 1.27) and compared to the theoretical prediction $\boldsymbol{d}_{thry}(\boldsymbol{\theta})$ over a large range of parameter space $\boldsymbol{\theta}$. Open source Boltzmann codes like `CLASS` [5] and `CAMB` [4] readily compute the evolution of the background and linear perturbations given a set of cosmological parameters to produce power spectra for dark matter and the CMB. The effect of nonlinear growth is incorporated using tools like `Halofit` [6] or `HMCode` [28], which can then be mapped to the observed galaxy power spectra given survey-dependent quantities such as dn/dz and $b(k, z)$ via Eq. 1.28.

Each of these steps includes additional assumptions and uncertainties, which can be parameterized and included as additional degrees of freedom in $\boldsymbol{\theta}$. For instance, in the fiducial analyses of Dark Energy Survey (DES), the mean and variance of dn/dz is allowed

[16] Note that the likelihood is a function of the parameters θ; it is not a probability distribution.

to vary in each redshift bin (with a Gaussian prior). One of the primary benefits of `HMCode` over `Halofit` is some freedom to capture the uncertain effect of baryons at small scales; DES instead uses `Halofit` and removes clustering data at the small scales where the uncertainty is non-negligible.

While adding nuisance parameters captures additional model uncertainty in the final posteriors of the cosmological parameters, the dimensionality of the model parameter space can quickly grow unwieldy. The fiducial DES 3x2pt analysis varies 26 parameters, far more than the $\sim$6 cosmological parameters that are most of interest (and indeed, only Ω_M and σ_8 are actually well measured by DES alone) [21]. With $\mathcal{O}(1)$ seconds to compute $\boldsymbol{d}_{thry}$ at a single point, an exhaustive grid-search of the parameter space is clearly infeasible and so Markov Chain Monte Carlo (MCMC) or nested sampling methods are the default approach for obtaining estimates of the posterior. Even so, producing a single set of posteriors typically takes days even when massively parallelized and run on high-performance computing clusters.[17] It is therefore crucial to understand which potential systematics must be addressed because of their ability to significantly impact the analysis. This will become especially important with the increased computational challenges of and improved statistics of Stage IV LSS surveys like the Vera Rubin Observatory's Legacy Survey of Space and Time (LSST).

1.9 Outline of Thesis

The studies in this thesis contribute to the body of literature investigating the impacts and mitigation of systematic errors in various cosmological analyses. We focus on broad classes of LSS systematic errors, which capture the effects of a wide range of real-world processes as they impact the observed galaxy fields and the theoretical matter power spectrum at small scales.

Chapter 2 presents a study of the Integrated Sachs-Wolfe (ISW) effect and how errors in the large-scale photometric calibration of LSS surveys impact estimation of the ISW signal. The ISW effect is an imprint of dark matter in the Cosmic Microwave Background and contains important information about dark energy, including possible signatures for modifications to General Relativity. We use a generic formalism to capture spatially-varying residual systematic errors in the observed galaxy fields and quantify the necessary levels of calibration to produce accurate reconstructions of the ISW map and power spectrum for next-generation LSS surveys. We provide a roadmap for ISW reconstruction, including the

[17] It is for this reason that we make use of the Fisher matrix formalism in Chapter 4, approximating the posterior surface as Gaussian in the parameters and enabling us to rapidly estimate changes in the posterior for thousands of different analysis configurations.

optimization of survey configuration and an improved estimator to render the analysis more robust to calibration errors.

In Chapter 3, we perform a detailed study of the leading methods for calibrating galaxy clustering measurements. These assess and remove spatially-dependent systematic errors such as those induced by interstellar dust, variable atmospheric conditions, and other effects that modulate the observed number of galaxies across the sky. We recast them into a common statistical framework, elucidating assumptions implicit within each method and characterize their performance on a suite of simulations. We propose extensions to current methods that are more robust, simpler to implement, and exhibit greater suppression of systematic errors. We further derive uncertainty estimates for the galaxy-level corrections, enabling the propagation of errors from the correction methods into the LSS galaxy catalogs and any subsequent analyses that use them.

In Chapter 4 we focus on small scale systematic errors in LSS analyses, such as arise from theoretical uncertainties in the non-linear growth of dark matter, baryonic effects, and other astrophysical phenomena. We characterize how errors in modeling such small scales impact our ability to accurately measure the primordial power spectrum of curvature fluctuations, which initially seeded structure in the early Universe. We show that if unaccounted for, current and predicted modeling uncertainties can strongly bias measurements of the "runnings" of the spectral index, key parameters for testing single-field slow-roll models of inflation, thought to be responsible for the rapid, early expansion of the Universe. We compare methods designed to mitigate such small-scale systematic errors and demonstrate that, even with optimistic improvements in small scale modeling, only exotic models of inflation will be testable via constraints on the runnings from near-future LSS surveys.

Chapter 2

Reconstructing the Integrated Sachs-Wolfe Map in the Presence of Systematic Errors

2.1 Introduction

Cosmic microwave background (CMB) photons undergo a frequency shift as they travel to us from the last scattering surface. On top of the redshift due to the expansion of the Universe, an additional contribution to the temperature anisotropy is introduced whenever the universe is not matter dominated—for example, right after recombination when radiation contributes non-negligibly, or at late times when dark energy becomes important. This so-called integrated Sachs-Wolfe (ISW) effect is given by [56, 57]

$$\left.\frac{\Delta T}{\bar{T}}\right|_{ISW}(\hat{\mathbf{n}}) = \frac{2}{c^2}\int_{t_*}^{t_0} dt\,\frac{\partial\Phi(\mathbf{r},t)}{\partial t}, \tag{2.1}$$

where t_0 is the present time, $t_\star$ is that of recombination, c is the speed of light, $\mathbf{r}$ is the position in comoving coordinates, and Φ is the gravitational potential. The late-time ISW signal (referred to hereafter simply as 'ISW') has been statistically detected via measurements of the cross-correlation of CMB temperature maps with galaxy maps [25, 26, 58–71] and, more recently, with maps of CMB lensing convergence [25, 26]. These detections serve as an important consistency test of the standard model of cosmology, and can help constrain the properties of dark energy.

The ISW can provide additional information beyond its power spectrum if its *map* can be reconstructed with sufficient signal-to-noise. Since the total large-angle CMB temperature

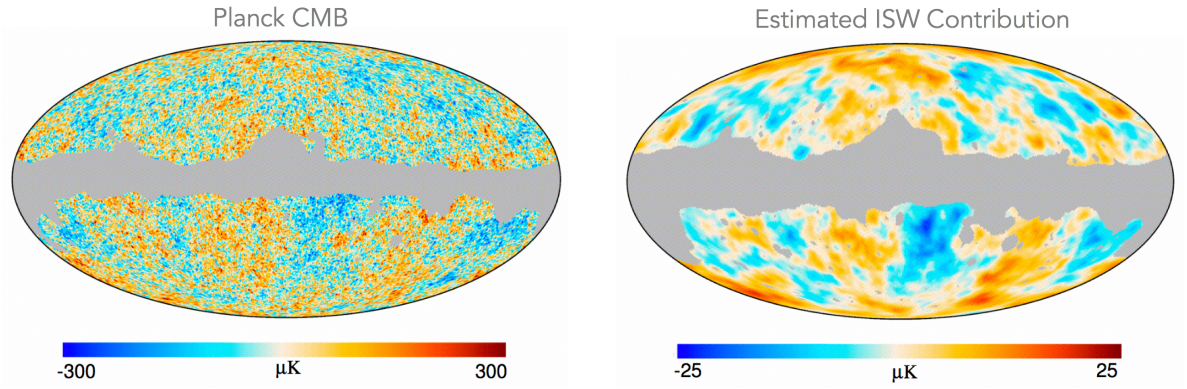

Figure 2.1: *Left:* Planck CMB map (using the SEVEM algorithm); *Right*: Reconstructed ISW map using the Planck CMB map and the CMB lensing potential as tracers. Adapted from Figures 1 and 11 in Ref. [25].

anisotropy is the sum of early- (hereafter 'primordial') and late-time contributions,

$$\left.\frac{\Delta T}{\bar{T}}\right|(\hat{\mathbf{n}}) = \left.\frac{\Delta T}{\bar{T}}\right|_{\rm prim}(\hat{\mathbf{n}}) + \left.\frac{\Delta T}{\bar{T}}\right|_{\rm ISW}(\hat{\mathbf{n}}), \tag{2.2}$$

reconstructing the ISW map would allow us to isolate the primordial-only anisotropy. This can be accomplished using the expected spatial correlation of the ISW signal with the CMB and with tracers of LSS. Fig. 2.1 shows one example, where the Planck collaboration used maps of the total CMB and the CMB lensing potential to estimate the ISW contribution [25].

This separation of the CMB into early- and late-time contributions can also be useful for a variety of cosmological tests. For example, one could study the temporal origin of the large-angle CMB anomalies reported in, e.g., Ref [72]. One could also subtract the realization-specific contaminating ISW contribution to estimation of primordial non-Gaussianity [73], something that is currently done using theoretical templates for the ISW-lensing bispectrum [74]. Motivated by these considerations, reconstruction of the ISW map has been the focus of a number of recent analyses [17, 25, 26, 75–82].

In this chapter we study how ISW map reconstruction is affected by a class of observational and astrophysical systematic errors which we will refer to broadly as photometric calibration errors or, for conciseness, calibration errors. These systematics afflict all galaxy surveys at large angular scales, contributing to the significant excess of power at large scales found in many recent surveys, including the Sloan Digital Sky Survey (SDSS) [71, 83–89], MegaZ [90], WISE-AGN and WISE-GAL [25], and NVSS [69, 71, 91, 92]. Calibration errors are thus already established as one of the most significant systematics impacting large-angle measurements of galaxy surveys, a fact that has broad implications, such as for measuring

scale-dependent bias as a signal of primordial non-Gaussianity. As the statistical power of galaxy surveys continues to grow, the control and understanding of systematics like calibration errors is becoming even more important.

There is a variety of ways in which modern photometric surveys assess and mitigate contamination from systematics, many of which rely on cross-correlating galaxy maps with known systematics templates. This can be used to identify contaminated regions, which are then masked or excluded from the analysis (as in Ref. [93]). A more nuanced approach is to use these templates to subtract or marginalize over systematics-induced spatial variations in the calculation of, for example, the two-point clustering signal [89, 94–99]. In Chapter 3 we take a detailed look at (and extend) some of these methods.

Ref. [89] took this approach when studying the overall detection significance of the ISW effect in SDSS data. The authors found results similar to Ref. [71], the authors of which instead accounted for excess power by adding a low-redshift spike in the source distributions. Most of these correlation corrections are perturbative, however, and will miss any contamination that is not captured by a perfectly linear relation. Additionally, only recently were some of the leading methods extended to remove systematics from the maps themselves, as opposed to simply cleaning the two-point statistics. Suchyta et al. [100] propose an alternative approach, wherein measurement biases are characterized by injecting fake objects into Dark Energy Survey images. This neatly avoids the reliance on having small levels of contamination in the input maps, but it still cannot account for certain systematics, such as dust or flux calibration. Whatever the approach taken, some level of residual calibration error will remain.

Muir and Huterer [17] (MH16) showed that at levels of calibration control consistent with current and near-future surveys, residual calibration errors are by far the dominant systematic for ISW signal reconstruction. This motivates us to study their impact in more detail. Namely, we would like to study whether the presence of residual calibration errors can be mitigated by combining information from multiple input maps or through better modeling of the contributions of systematics to observed galaxy power. We also wish to investigate to what extent residual calibration errors similarly impact the signal-to-noise ratio of galaxy-CMB cross-correlation and, in turn, the significance of ISW detection. With this aim, we use ensembles of simulated maps to characterize the performance of ISW reconstruction based on surveys like Euclid and SPHEREx, two proposed wide-angle surveys of which the properties are expected to be good for ISW detection and reconstruction. We also consider the benefits of including Planck-like simulations of CMB intensity in the reconstruction effort.

We begin in Sec. 2.2 by describing our model for calibration errors, how we reconstruct the ISW map and evaluate its quality, and which input data sets we use. In Sec. 2.3, we

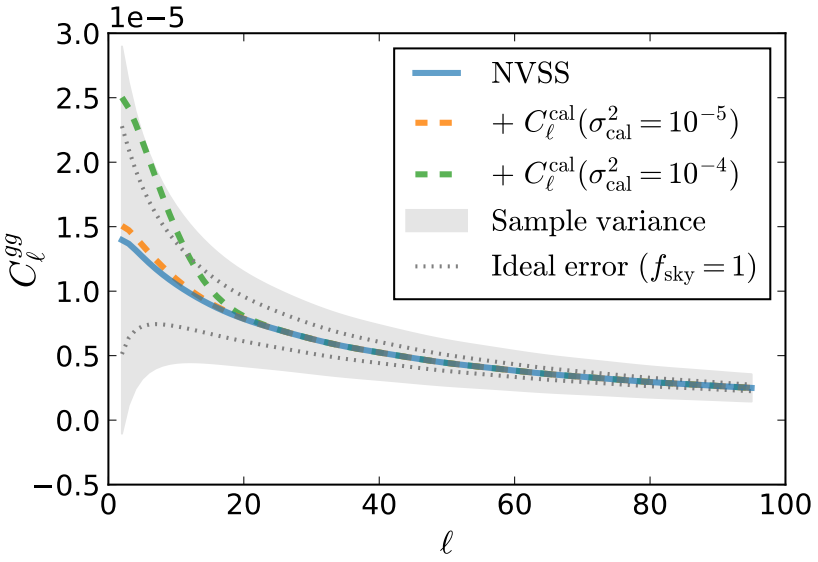

Figure 2.2: Effects of calibration errors on galaxy power. The solid curve shows the theoretical angular power spectrum for the NVSS survey [25, 26]. The colored dashed curves show the theoretical spectrum with two representative levels of calibration error. The shaded region is the 1σ uncertainty from the survey's sample variance, and the dotted curves indicate the ideal, all-sky cosmic variance.

compare the performance of ISW reconstruction when using one versus multiple surveys and investigate the impact different assumptions have on the results. In Sec. 2.4, we relate map reconstruction to the total signal-to-noise ratio of ISW detection, and we conclude in Sec. 2.5.

2.2 Methodology

2.2.1 Modeling Calibration Errors

Photometric calibration is a challenge faced by all photometric galaxy surveys. It refers to the adjustments required to establish a consistent spatial and temporal measurement of flux of the target objects in different observation bands. A number of different systematics must be accounted for in calibration, including but not limited to detector sensitivity variation on the focal plane, variation in observing conditions, the presence of foreground stars (as galaxies near them are less likely to be detected), and extinction by interstellar dust. Calibration errors are introduced if these systematics are incompletely or inaccurately accounted for.

Our focus is on how calibration errors affect galaxy number counts. To illustrate this, imagine that a perfectly uniform screen (of e.g. dust) blocks some light from all galaxies. This pushes the faintest galaxies below the survey's flux limit, and leads to observation of

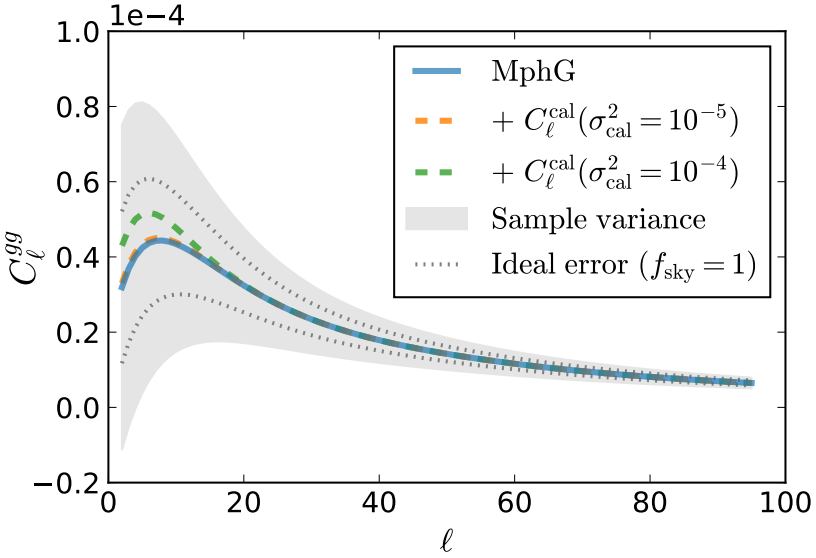

Figure 2.3: Same as Fig. 2.2, but for the SDSS MphG catalog, following Refs. [25, 26]. In this case, the sample variance is driven by sky coverage (f_{sky}=0.22) as opposed to number density as for NVSS.

fewer galaxies in all directions. A pure monopole change such as this increases shot noise but does not affect the angular clustering signal of galaxies. In contrast, in a more realistic scenario where the opacity of this "screen" depends on direction, it affects the observed galaxy clustering signal by adding spurious power and by coupling different multipoles of the measured power spectrum [101, 102]. The presence of unaccounted-for calibration errors can thus introduce biases in cosmological parameter estimates from large-scale structure surveys. These are particularly severe for the ISW effect and other measurements based on signals that, like calibration errors, enter primarily at large angular scales.

In keeping with the picture of calibration errors as a direction-dependent screen, we model them as a modulation of the true galaxy number counts $N(\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is the direction on the sky. The observed, modulated counts are [101]

$$N_{\rm obs}(\hat{\mathbf{n}}) = [1 + c(\hat{\mathbf{n}})]\, N(\hat{\mathbf{n}}), \tag{2.3}$$

where the field $c(\hat{\mathbf{n}})$ describes the screening effect of calibration errors. While we will generally refer to this kind of modulation as the result of "calibration errors," as Eq. (2.3) makes clear, this formalism can describe any residual effect that modulates a survey's selection function.

Though the expression in Eq. (2.3) will necessarily couple different multipoles, at low ℓ, the impact of calibration errors on the observed galaxy power spectrum is well approximated

by

$$C_\ell^{\rm obs} \approx C_\ell^{\rm gal} + C_\ell^{\rm cal}, \tag{2.4}$$

neglecting multiplicative terms.

Following MH16, we model the calibration error field $c(\hat{\mathbf{n}})$ as a Gaussian random field with power spectrum $C_\ell^{\rm cal}$ and quantify the level of residual calibration errors using its variance,

$$\sigma^2_{\rm cal} \equiv {\rm Var}({\rm c}(\hat{\mathbf{n}})) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{4\pi} {\rm C}_\ell^{\rm cal}. \tag{2.5}$$

While the conversion between $\sigma_{\rm cal}$ and the rms magnitude error depends on the faint-end slope of the luminosity function of tracers in the survey, they can be related roughly as $(\delta m)^{\rm rms} \simeq \sigma_{\rm cal}$ [101]. Thus a survey with residual calibration errors of $\sigma^2_{\rm cal} = 10^{-6}$ has been calibrated to roughly a milimagnitude.

Motivated by power spectrum estimates for maps of dust extinction corrections and magnitude limit variations in existing surveys (see Fig. 5 and 6 in Ref. [101]), we choose the fiducial calibration error power spectrum to be

$$C_\ell^{\rm cal} = \alpha^{\rm cal} \exp\left[-(\ell/w^{\rm cal})^2\right], \tag{2.6}$$

with $w^{\rm cal} = 10$. The normalization constant $\alpha^{\rm cal}$ is varied to achieve the desired $\sigma^2_{\rm cal}$. Figures 2.2 and 2.3 show the impact of calibration errors of this form on the angular power spectrum of the NVSS and SDSS MphG galaxy surveys, which have been used to reconstruct the ISW map in previous studies [25, 26, 80]

For our fiducial model, we assume that calibration error maps for different redshift bins and surveys are uncorrelated with one another. We briefly examine the impact of relaxing such an assumption in Sec. 2.3.4.

2.2.2 ISW Estimator

Similarly to MH16, we work with the optimal estimator derived by Manzotti et al [80]. It takes as input n maps, which can include any tracers that carry information about the ISW signal, namely LSS, CMB, or lensing convergence maps. Letting $g^i_{\ell m}$ represent the spherical components of the ith input map, where $i \in \{1, \ldots, n\}$, the optimal estimator for the spherical component of the ISW signal is

$$\hat{a}^{\rm ISW}_{\ell m} = \sum_{i=1}^{n} R^i_\ell \, g^i_{\ell m}. \tag{2.7}$$

The operator

$$R^i_\ell \equiv -N_\ell [D_\ell^{-1}]_{\mathrm{ISW}-i} \tag{2.8}$$

is a reconstruction filter derived from the covariance matrix,

$$D_\ell = \begin{pmatrix} C_\ell^{\mathrm{ISW,ISW}} & C_\ell^{\mathrm{ISW},1} & \cdots & C_\ell^{\mathrm{ISW},n} \\ C_\ell^{1,\mathrm{ISW}} & C_\ell^{1,1} & \cdots & C_\ell^{1,n} \\ \vdots & \vdots & \ddots & \vdots \\ C_\ell^{n,\mathrm{ISW}} & C_\ell^{n,1} & \cdots & C_\ell^{n,n} \end{pmatrix}. \tag{2.9}$$

In this expression, superscript numbers label the input maps and $N_\ell \equiv 1/[D_\ell^{-1}]_{\mathrm{ISW}-\mathrm{ISW}}$ estimates the variance of the reconstruction at multipole ℓ. When a single input map A, is used, this expression reduces to a simple Wiener filter,

$$\hat{a}^{\mathrm{ISW}}_{\ell m} \xrightarrow{\text{single map}} \frac{C_\ell^{\mathrm{ISW}-\mathrm{A}}}{C_\ell^{\mathrm{A}-\mathrm{A}}} g^{\mathrm{A}}_{\ell m}. \tag{2.10}$$

We demonstrate in Appendix A.1 that Eq. (2.7) is equivalent to the estimator of Manzotti et al [80], where the CMB temperature map is treated separately from LSS maps, and show that it reduces to the Linear Covariance Based (LCB) filter first proposed by Barreiro et al [75].

In constructing this ISW estimator, one must make a choice about how to obtain the necessary angular power and cross-power spectra in the covariance matrix. The C_ℓ's can either be extracted from observations (as in Refs. [77, 79]) or computed analytically for an assumed cosmology (as in Refs. [17, 25, 26, 80, 82]). Analytic calculation is straightforward but introduces a model dependence which can potentially bias results if, for example, calibration error contributions are not modeled correctly [17]. Measuring C_ℓ from observations produces a model-independent estimator and so can help in the case where the theory spectra are inaccurate, but at the expense of limited precision due to sample variance, especially at large scales, scales with low power, or for map combinations that have little correlation.[1] Hybrid methods can also be used, as in Ref. [78], which accounts for observed excesses in the autopower of NVSS data by using a smoothed fit to data to get the galaxy map's autopower, but analytically computes its cross-correlation with the ISW signal.

We therefore consider two limiting cases of constructing the estimator in order to investigate how calibration errors impact the ISW reconstruction:

1. a 'worst' case estimator filter, $R_\ell(C_\ell^{\mathrm{th}})$, where we use the fiducial theory C_ℓ's in the

[1] Using the observed spectra also violates an assumption in the maximum likelihood derivation of the estimator, in which the covariance is assumed to be known (i.e. independent of the measured signal).

estimator, in which calibration errors' power contributions are not modeled at all, and

2. a 'best' case estimator filter, $R_\ell(C_\ell^{\rm sky})$, in which calibration error power contributions are modeled perfectly (i.e. the covariance matrix is known). This case may be approximated by, e.g. a smoothed fit of the observed LSS power.

The theoretical spectra are related simply through the expression

$$C_\ell^{\rm sky} = C_\ell^{\rm th} + C_\ell^{\rm cal}. \tag{2.11}$$

where $C_\ell^{\rm cal}$ is the power spectrum of the calibration error field described in Sec. 2.2.1. We consider these cases in Secs. 2.3.1 and 2.3.2 respectively.

2.2.3 Quality Statistic

To quantify the accuracy of a given reconstruction, we use the correlation coefficient between the temperature maps of the true $[T^{\rm ISW}(\hat{\mathbf{n}})]$ and reconstructed $[T^{\rm rec}(\hat{\mathbf{n}})]$ ISW signal,

$$\rho = \frac{\frac{1}{N_{\rm pix}}\sum_k^{N_{\rm pix}}(T_k^{\rm ISW} - \bar{T}^{\rm ISW})(T_k^{\rm rec} - \bar{T}^{\rm rec})}{\sigma_{\rm ISW}\sigma_{\rm rec}}, \tag{2.12}$$

where $\bar{T^X}$ and σ_X^2 are the mean and variance of map $T^X(\hat{\mathbf{n}})$, respectively.[2] We do not include pixel weights in our calculation of ρ, as is done to account for masking effects in Ref. [82]. This is because we work with only full-sky maps, as will be discussed in the next section.

The correlation coefficient can be rewritten in terms of the cross-power between the true ISW map realization and the input tracers,

$$\rho = \frac{\frac{1}{4\pi}\sum_{\ell,i}(2\ell+1)R_\ell^i\tilde{C}_\ell^{\,{\rm ISW}-i}}{\sigma_{\rm ISW}\sigma_{\rm rec}}, \tag{2.13}$$

where the tilde denotes pseduo-C_ℓ measured from a given map realization, and we have used Eq. (2.7) to write

$$\tilde{C}_\ell^{\rm ISW-rec} = \sum_i \frac{1}{2\ell+1}\sum_m [a_{\ell m}^{\rm ISW}]^* R_\ell^i g_{\ell m}^i \tag{2.14}$$

$$= \sum_i R_\ell^i \tilde{C}_\ell^{{\rm ISW}-i}. \tag{2.15}$$

[2] We also considered s, which measures the rms error between true and reconstructed ISW maps as a complementary quality statistic, but found that for the cases studied here, the information it provided was largely redundant to that given by ρ.

Because the measured correlation coefficient depends on the specific realization, we assess reconstruction accuracy for a given set of input map properties as follows. We simulate a large number of realizations of correlated maps, then apply the ISW estimator to obtain associated reconstructed ISW maps, and by comparing these with the true ISW maps we obtain a sample distribution for ρ. Its mean value $\bar{\rho}$, which in the limit of an infinitely large ensemble will approach an expectation value $\langle\rho\rangle$, provides a statistical measure of how accurately the estimator can reproduce the true ISW signal. Studying how $\bar{\rho}$ changes in response to variations in survey properties and modeling choices therefore allows us to understand which factors are most important for obtaining an accurate ISW reconstruction.

We can avoid the computational cost of generating many simulation ensembles by noting that we can obtain a good estimate for the expectation value of ρ if we make the approximation

$$\langle\rho\rangle = \left\langle \frac{\frac{1}{4\pi}\sum_{\ell,i}(2\ell+1)R_\ell^i \tilde{C}_\ell^{\,\mathrm{ISW}-i}}{\sigma_{\mathrm{ISW}}\sigma_{\mathrm{rec}}} \right\rangle \tag{2.16}$$

$$\approx \frac{\frac{1}{4\pi}\sum_{\ell,i}(2\ell+1)R_\ell^i C_\ell^{\mathrm{ISW}-i}}{\hat{\sigma}_{\mathrm{ISW}}\hat{\sigma}_{\mathrm{rec}}}, \tag{2.17}$$

that is, we replace the pseudo-C_ℓ's with their expectation value across realizations, $\tilde{C}_\ell \to C_\ell$. We will refer to the quantity in Eq. (2.17) as $\hat{\rho}$, defining

$$\hat{\sigma}_{\mathrm{ISW}} = \sqrt{\frac{1}{4\pi}\sum_\ell (2\ell+1)\, C_\ell^{\mathrm{ISW}}} \tag{2.18}$$

$$\hat{\sigma}_{\mathrm{rec}} = \sqrt{\frac{1}{4\pi}\sum_{\ell,i,j} (2\ell+1)\, R_\ell^i R_\ell^j C_\ell^{ij}}\,, \tag{2.19}$$

to approximate the rms fluctuations in the true and reconstructed ISW maps. Here the indices i and j label the input tracer maps and the sum over ℓ runs over the multipoles $\ell \in [2, 95]$, a range chosen to conservatively to include all scales where the ISW signal is important.

We have tested the approximation $\hat{\rho} \approx \langle\rho\rangle$ in Eq. (2.17) extensively and found it works well when the estimator filter R_ℓ is built from analytically computed spectra but can break down if R_ℓ is composed of $\tilde{C}_\ell$'s extracted from map realizations. This behavior is related to the way in which using measured C_ℓ's makes ρ depend on $\tilde{C}_\ell$, such that $\bar{\rho} = \rho(\langle\tilde{C}_\ell\rangle)$ is no longer a good approximation of $\langle\rho(\tilde{C}_\ell)\rangle$. Appendix A.2 discusses this in more detail.

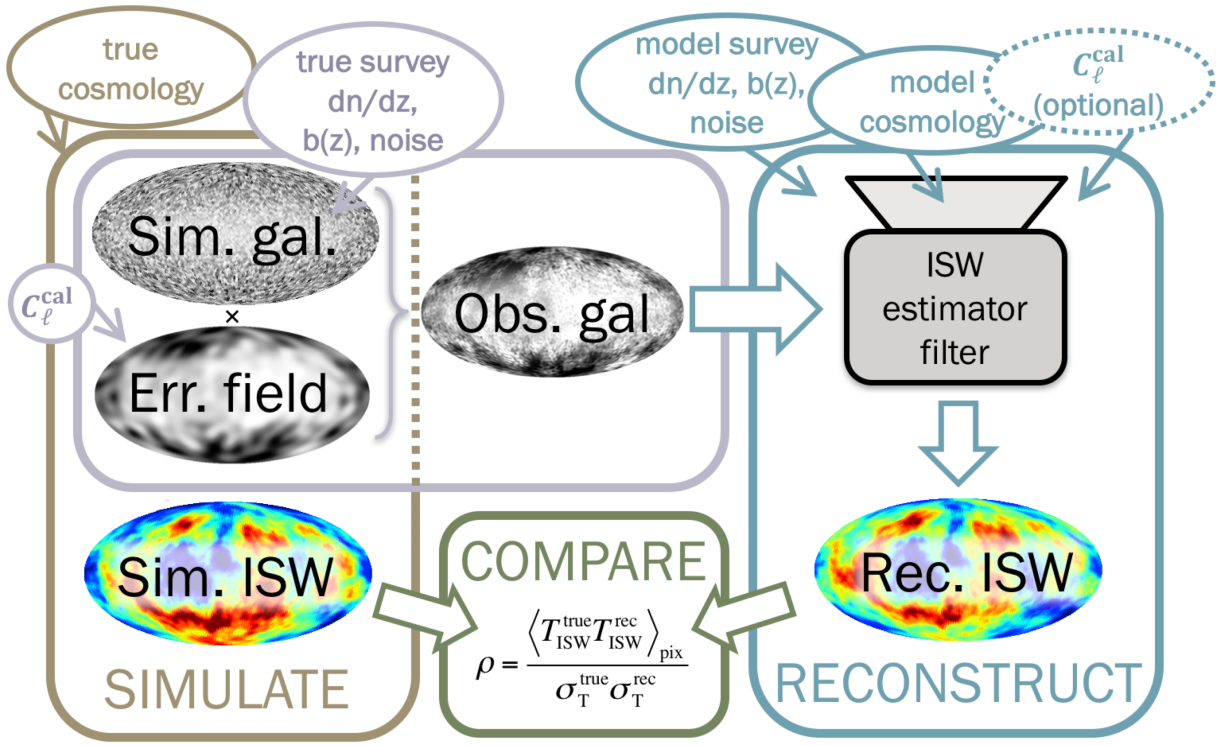

Figure 2.4: Schematic illustration of the general analysis methodology described in Sec. 2.2. Correlated "true" galaxy and ISW maps are simulated for each realization, given a set of cosmological parameters and LSS survey properties. A realization of a calibration error field modulates the true galaxy map to produce the observed galaxy field, which is then used to reconstruct the ISW map using the estimator described in Sec. 2.2.2, and compared to truth.

2.2.4 Simulated Surveys

Fig. 2.4 schematically illustrates our overall methodology, including the generation of correlated ISW, CMB and LSS maps (described here), and the ISW reconstruction and quality assessment using the tools described in the previous sections.

By working with simulated maps, we are able to study in detail how calibration error levels and modeling choices affect ISW signal reconstruction.

Since we are concerned only with large scales, we model the ISW signal, total CMB temperature anisotropy, and galaxy number density fluctuations as correlated Gaussian fields. We use `HEALPY` [12] to generate map realizations based on input auto- and cross-power spectra which we compute analytically following the standard expressions given e.g. in Ref. [17]. We use the Limber approximation for $\ell \geq 20$, having verified that this affects ρ at the level of 0.1% or less for the surveys and range of $\sigma^2_{\rm cal}$ considered here. We compute C_ℓ for multipoles with $\ell \leq 95$, as this range contains almost all of the ISW signal [103]. Accordingly, our simulations are sets of `HEALPIX` maps of resolution NSIDE $= (\ell_{\rm max} + 1)/3 = 32$. We refer

the reader to Ref. [17] for a more detailed description of the reconstruction pipeline.

Because our goal is to study the impact of calibration errors and not survey geometry, we assume full-sky coverage in all of our analyses. Ref. [82] found that in overlapping regions of partial sky LSS surveys, ISW reconstruction quality degrades only slightly compared to the full-sky case. Therefore, the performance of a given estimator using full-sky maps should be indicative of its performance using maps with only partial sky coverage.

Our fiducial cosmological model is ΛCDM, with the best-fit cosmological parameter values from Planck 2015, $\{\Omega_c h^2, \Omega_b h^2, \Omega_\nu h^2, h, n_s\} = \{0.1188, 0.0223, 0, 0.6774, 0.9667\}$. Unless otherwise stated, ISW reconstructions are performed on 2000 map realizations for each analysis and include multipole information down to $\ell_{\rm min} = 2$.

Within this framework, four pieces of information are required to model a LSS survey: the distribution of its sources along the line of sight $n(z)$, a prescription for how they are binned in redshift, their linear bias $b(z)$, and their projected number density per steradian $\bar{n}$. Below we describe how our choices for these characteristics are based on the properties of promising future probes of the ISW effect.

Euclid-like LSS Survey

Our fiducial survey is modeled on Euclid, a future LSS survey with large sky coverage and a deep redshift distribution [104], which is expected to be an excellent probe of the ISW effect [103, 105]. We assume the redshift distribution used by Martinet et al [106],

$$\frac{dn}{dz} = \frac{3}{2z_0^3} z^2 \exp\left[-(z/z_0)^{1.5}\right], \tag{2.20}$$

which has a maximum at $z_{\rm peak} \simeq 1.21 z_0$. We choose $z_0 = 0.7$ and $\bar{n} = 3.5 \times 10^8$, with a photo-$z$ redshift uncertainty of $\sigma(z) = 0.05(1+z)$ which smoothes the edges of redshift bins. For simplicity, we assume a constant galaxy bias of $b(z) = 1$. Our results are qualitatively insensitive to this choice as long as the bias is reasonably well approximated for the input maps. This is because the bias term cancels in the estimate of the ISW signal, so that fractional differences between true and modeled bias have little impact on ρ.

Muir and Huterer [17] investigated the impact of numerous systematic errors stemming from mismodeling the survey properties and underlying cosmology in the ISW estimator, finding residual calibration errors to be the dominant systematic. We therefore set cosmological and survey parameters in our models to be equivalent to the simulated truth and focus on the impact of residual calibration errors. We refer the reader to Ref. [17] for further details on both fitting for bias and the impact that mismodeling can have on reconstruction.

In Sec. 2.3.1 we investigate the improvement in ISW map reconstruction when the fiducial

Euclid-like survey is split into six redshift bins with edges at $z \in \{0.01, 0.4, 0.8, 1.2, 1.6, 2, 3.5\}$ (see inset of Fig. 2.5), as compared to the unbinned case. We subsequently use the six-binned Euclid survey as our fiducial case.

SPHEREx-like LSS Survey

We model a second survey on the SPHEREx All-Sky Spectral Survey (SPHEREx), a proposed survey that has been optimized to study LSS in the low-redshift universe. One of its goals is to place stringent limits on primordial non-Gaussianity [107], which will require rigorous control of calibration errors. Given this, SPHEREx will provide excellent input map candidates for ISW map reconstruction. Its shallower reach makes it complementary to the deeper mapping of the LSS provided by Euclid.

SPHEREx will identify galaxies with varying levels of redshift uncertainty, ranging from $\sigma_z < 0.003(1+z)$ up to $\sigma_z > 0.1(1+z)$. Grouping these into catalogs with different levels of precision provides collections of galaxies useful for different science goals. The $\sigma_z < 0.1(1+z)$ catalog with a projected ~ 300 million galaxies was identified in Ref. [107] as the best subsample for detecting the primordial non-Gaussianity parameter $f_{\rm NL}^{\rm loc}$, whose signal is highest at the large scales. Our investigations confirm this catalog of SPHEREx to be the best for ISW detection as well. We therefore fit its projected redshift distribution given in Ref. [107] to the functional form for dn/dz given in Eq. (2.20). We select $z_0 = 0.46$, which results in a peak dn/dz of $z_{\rm peak} \simeq 0.56$. We have confirmed that our results are not strongly sensitive to changes in this redshift distribution, in agreement with the findings of Ref. [17].

We use a projected number density of $\bar{n} = 6.6 \times 10^7$ and consider the case where the survey is split into six redshift bins. We choose their edges by scaling the Euclid-like survey's binned redshift distribution to the SPHEREx median redshift, resulting in redshift bin edges at $z \in \{0.01, 0.26, 0.53, 0.79, 1.05, 1.31, 2.30\}$. This still provides sufficient sampling of the field in each bin to ensure that shot noise is subdominant to the galaxy signal power.

Planck-like CMB Survey

CMB data have frequently been used in conjunction with LSS data for ISW map reconstruction. Recent examples include Ref. [25], which used NVSS radio data, the Planck lensing convergence map, and Planck temperature data. That analysis was subsequently extended to include more LSS tracers in Ref. [26]. However, in both of these cases, residual systematics limit the usefulness of lensing data to scales of $\ell \geq 10$ and $\ell \geq 8$, respectively. Ref. [82] investigated the usefulness of CMB data for ISW reconstruction using a simulation pipeline

similar to ours, finding that both CMB temperature and polarization data only modestly improve reconstruction quality but carry a greater benefit when the LSS tracers themselves contain less information (due to e.g. noise or other properties of the survey).

It is then natural to ask whether CMB data can help mitigate the impact of calibration errors in LSS maps. We therefore consider CMB temperature as an additional input map. To compute the total CMB temperature power spectrum, C_ℓ^{TT}, we compute the primordial-only contributions using a modified version of `CAMB` [4] and add them to our calculations for $C_\ell^{\rm ISW}$. As the CMB power spectrum is determined within the limits of cosmic variance at low ℓ and the ISW signal is already dominated by the primary (that is, non-ISW) CMB anisotropies, we do not include calibration errors in the generation of CMB temperature maps. Though CMB polarization and lensing could provide additional information, residual systematics remain at large scales for each (see Refs. [108] and [109], respectively), so for simplicity we do not include them in this analysis.

2.3 Results

To characterize the impact of calibration errors in LSS surveys on the ISW map reconstruction, and the potential to mitigate these impacts, we look at multiple combinations of input maps with different properties. Specifically, we consider the impact of binning in redshift, of adding CMB intensity data, and of including additional LSS information from another survey. For each of these studies, we examine two limiting cases for the estimator. The best case scenario, which we will reference as $R_\ell(C_\ell^{\rm sky})$, is when one perfectly models all contributions to the galaxy power, including residual calibration errors. The worst case, referenced by $R_\ell(C_\ell^{\rm th})$, is when the estimator is built out of theoretical spectra with no power from calibration errors. The power spectra in these two cases are related by Eq. (2.11).

We use the analytical $\hat{\rho}$ to estimate the mean reconstruction quality across a wide range of $\sigma^2_{\rm cal}$, while performing reconstruction on simulated maps for selected values, to both verify the accuracy of $\hat{\rho}$ and to generate error bars for the spread of ρ across simulations.

2.3.1 One Survey: Binning in Redshift

We first consider the Euclid-like survey alone and investigate the impact of binning in redshift on the quality of reconstructions in the presence of calibration errors. We model calibration errors in the binned case by adding the calibration field's power to the autopowerpower spectrum of each bin i per Eq. (2.4): $C_\ell^{i,i} \to C_\ell^{i,i} + C_\ell^{\rm cal}$. We do not add any power to the

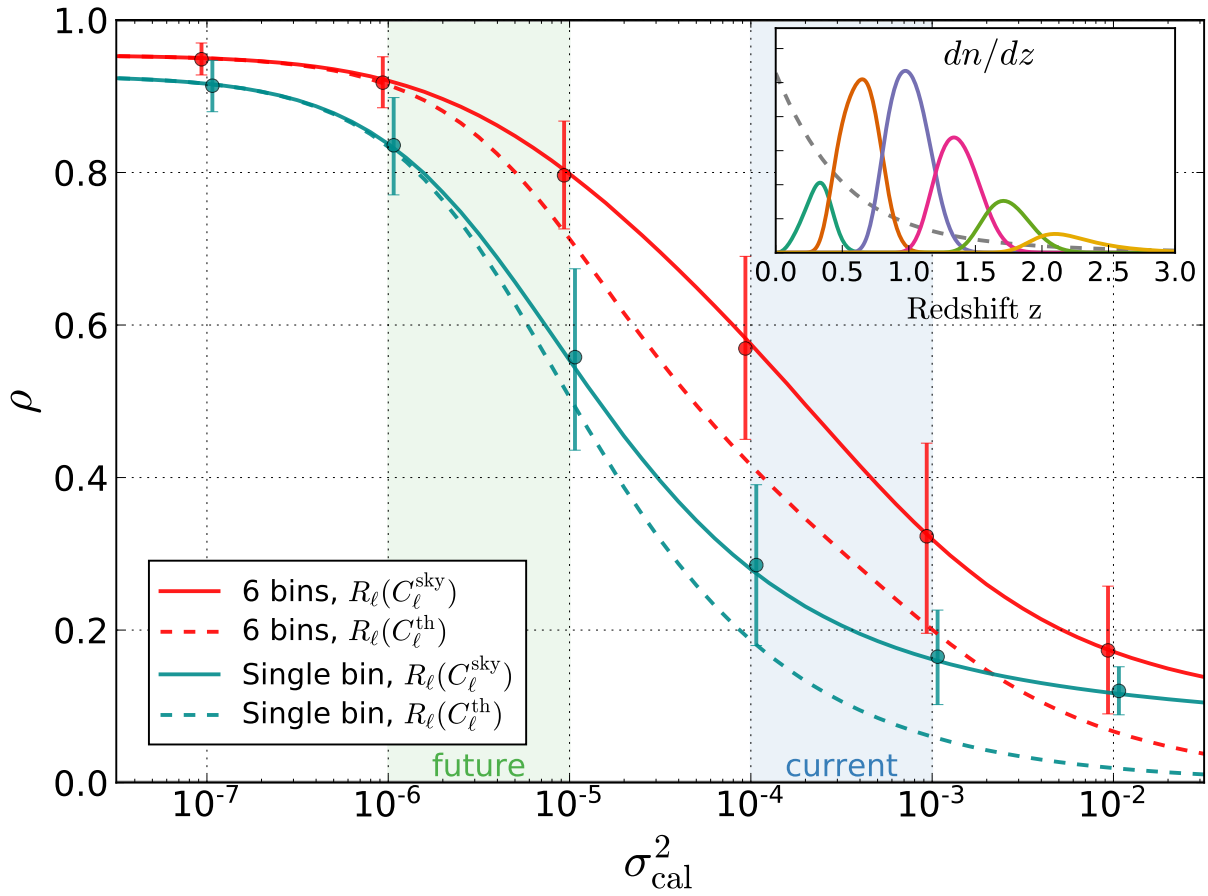

Figure 2.5: Quality of map reconstruction ρ vs. the calibration error variance $\sigma^2_{\rm cal}$ for our fiducial, Euclid-like survey. The colors of the lines indicate how tomographic information is handled, showing that splitting the survey into six redshift bins (red) improves the reconstruction compared to the single-bin case (blue). Solid curves indicate cases when the calibration error is included in the ISW estimator $[R_\ell(C_\ell^{\rm sky})]$, while the dashed curves show the reconstructions in which the effects of the calibration errors are not included ($[R_\ell(C_\ell^{\rm th})]$ (see Sec. 2.2.2 for details). Points (offset horizontally for clarity) show the mean ($\bar{\rho}$) of 2,000 realizations, with error bars indicating the 68% spread across realizations. The corresponding smooth curves are $\hat{\rho}$, the analytical estimate of $\bar{\rho}$ from Eq. (2.17). The inset illustrates the redshift distribution across bins overlaid with the ISW kernel in gray (reproduced from Ref. [17]). The vertical, shaded regions show the approximate current and projected levels of control over residual calibration errors. Calibration errors between redshift bins are modeled as uncorrelated.

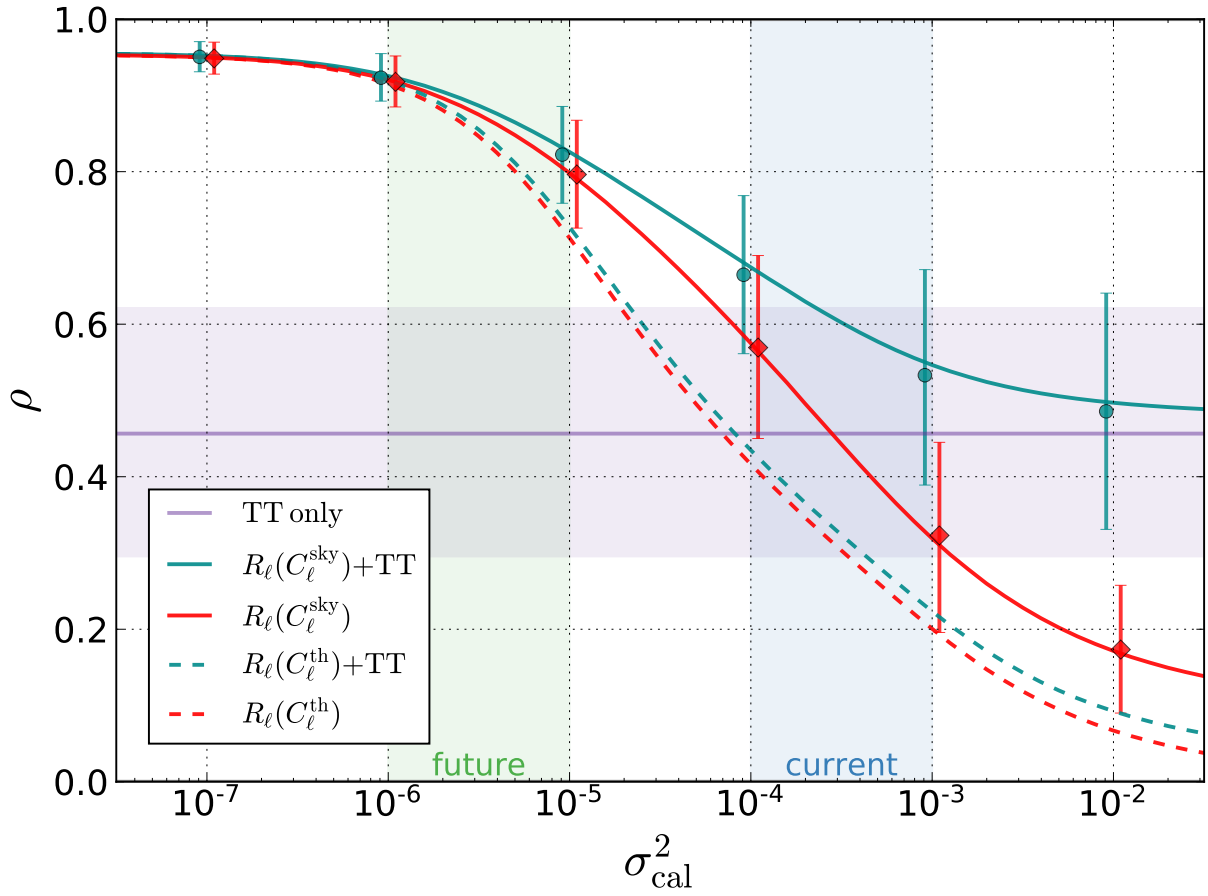

Figure 2.6: Reconstruction quality when using binned Euclid-like survey and CMB intensity data separately and in combination. The purple curve and shaded band show the mean and 68% spread of ρ from simulations. As in Fig. 2.5, red curves are results when using the binned Euclid-like survey alone, whereas blue curves are the result of using both the fiducial survey and CMB intensity map. Like in Fig. 2.5, solid curves are for the case where calibration error power is correctly modeled in the estimator and dashed curves are for when they are not modeled at all. Neglecting the presence of calibration errors in a LSS map can actually degrade the quality of the ISW reconstruction compared to using the CMB temperature alone.

cross-spectra, though we test the impact of contamination in the cross-spectra in Sec. 2.3.4.[3]

The results of this study are shown in Fig. 2.5. For reference, we use a vertical shaded band to mark the level of calibration corresponding to current surveys, defined roughly as the range bounded by the residual SDSS DR8 limiting magnitude variations [110] and the SDSS über-calibration [111]. ('Future' levels are defined roughly as those between that required to limit bias on cosmological parameters to below their projected uncertainties and an intermediate level before bias becomes unacceptable; see Refs. [101] and [17] for details.)

As shown in Fig. 2.5, splitting the survey into six redshift bins results in significant improvement in reconstruction at all levels of calibration error. This improvement is comparable to reducing $\sigma^2_{\rm cal}$ of the single-bin case by a factor of 10 at 'current' levels.

Tomographic information is useful because it allows galaxy counts to be weighted more optimally, taking advantage of the fact that the ISW effect becomes stronger at lower redshift as dark energy becomes more dominant and structure growth slows. For instance, considering the expected ISW reconstruction power from each bin when using optimal weights (i.e. the squared contribution of each term in Eq. 2.7, using $R_\ell(C_\ell^{\rm sky})$), we find that with no calibration error, the first redshift bin contributes 87% as much power as the second bin, with subsequent bins contributing 58%, 31%, 15% and 10% as much power, respectively. There is additional benefit to binning when calibration errors are considered. Since the low-redshift bins have a higher clustering signal than the high-redshift bins, they are less impacted by the same level of calibration error. Thus, the optimal weighting changes depending on the level of calibration error. When calibration errors are increased to $\sigma^2_{\rm cal} = 10^{-4}$, for example, the first bin contributes the most power to the reconstruction, with bins $2-6$ only contributing 39%, 12%, 4%, 2%, and 1% as much power. As we will show later, this error-level-dependent weighting will mean adding information from a shallower survey such as SPHEREx makes reconstruction more robust against calibration errors.

The importance of accounting for calibration errors in the estimator is apparent in the difference between the dashed and solid curves, where doing so improves $\bar{\rho}$ for $\sigma^2_{\rm cal} \gtrsim 10^{-6}$, with $\Delta\bar{\rho} \approx 0.1-0.2$ at current levels of calibration. This improvement is roughly comparable to the improvement seen from binning in redshift.

Though for clarity we do not include this case in the Figure, we additionally studied the effect of using the observed, unsmoothed galaxy-galaxy power in the estimator (that is, $\tilde{C}_\ell$, power spectra extracted from map realizations rather than computed analytically). We find that in this case $\bar{\rho}$ converges to the same value as the $R_\ell(C_\ell^{\rm sky})$ case when calibration

[3] In reality, the power contribution from calibration errors will also vary somewhat across bins, depending on the redshift dependence of the faint-end slope of the luminosity function for the tracer population. We have assumed here for simplicity that the power contribution is independent of redshift.

errors are very large, but is greatly reduced from $\bar{\rho}$ found using either $R_\ell(C_\ell^{\rm sky})$ or $R_\ell(C_\ell^{\rm th})$ when calibration errors are small ($\sigma^2_{\rm cal} \ll 10^{-5}$). For example, for a single input map in the limit of no calibration errors, quality reduces from $\bar{\rho} = 0.93$ to 0.83 when we switch to using observed $\tilde{C}_\ell$'s. If we also use the observed (unsmoothed) *cross*-correlation between the LSS map and the CMB for the galaxy-ISW term in the estimator, reconstruction quality is further degraded to $\bar{\rho} = 0.74$ in the absence of calibration errors. This is because primary CMB anisotropies are large compared to ISW contributions, causing the measured galaxy-CMB correlation to receive relatively large noise contributions from chance correlations between LSS maps and the primordial CMB.

Given the significant improvement in reconstruction that binning provides, from here forward we adopt the configuration with six tomographic bins as our fiducial Euclid-like survey.

2.3.2 Effect of Adding Planck TT Data

We now consider adding information from the Planck-like CMB temperature map described in Sec. 2.2.4. When used as the only input map, the reconstruction is considerably worse than that found using the ideal Euclid-like survey (Table 2.1). We include it in our study, however, because any realistic study attempting to reconstruct the ISW signal will likely include CMB temperature data. Additionally, the reconstruction quality attainable with CMB temperature data alone provides a useful baseline against which to compare the performance of estimators based on LSS maps.

With CMB temperature data alone, we find an average reconstruction quality of $\bar{\rho} = 0.46$, in good agreement with Ref. [82]. To put this into proper context, however, it is important to note that there is a large scatter around that mean; while the average reconstruction quality is indicative of performance, any single realization, such as that of our own Universe, can vary substantially in fidelity. The purple band in Fig. 2.6 shows the extent of this scatter for ISW reconstruction based on just the CMB map.

When CMB temperature information is combined with that from LSS maps, it significantly improves reconstruction quality, but only if the true galaxy power spectrum $C_\ell^{\rm sky}$ (including calibration error contributions) is used in the estimator, as can been seen by the behavior of the solid curves on the right-hand side of Fig. 2.6. The blue $\rho(\sigma^2_{\rm cal})$ curve describing the CMB+LSS reconstruction tracks the maximum of the curves corresponding to reconstructions using the CMB and LSS input maps separately, shown by the purple and red curves, respectively. This occurs because the estimator down-weights the LSS survey the more it is affected by calibration errors, converging to the TT-only reconstruction quality in

the limit of large calibration errors. If one does not model calibration error power contributions, however, then any improvement from combining multiple input maps is marginal at best and can in fact result in a *worse* reconstruction than just using the CMB data alone. This demonstrates the importance of ensuring that the LSS C_ℓ's used in the ISW estimator are a good fit to the observed spectra.

2.3.3 Effect of an Additional LSS Survey: SPHEREx-like

We now consider the addition of our fiducial six-bin SPHEREx-like survey described in Sec. 2.2.4, assuming for simplicity that it has the same level of calibration errors as the Euclid-like survey. Results are shown in Fig. 2.7.

In the limit of no calibration errors, the SPHEREx-like survey offers little additional information. In fact, adding both SPHEREx and CMB TT results in negligible improvement over the Euclid-like only case ($\Delta\bar{\rho} < 0.003$ compared to a spread of $\sigma_{\rm Euc+Spx+TT} = 0.019$).

However, by comparing the black and blue curves we see that including the SPHEREx-like survey does make the reconstruction somewhat more robust against calibration errors. The reason for this is similar to why binning in redshift is helpful: recall that, in the case of binning, having narrow, low-redshift bins means having some bins with higher galaxy autopower than the unbinned case, which then have less susceptibility to a given level of calibration error. Similarly, SPHEREx has a shallower redshift distribution, and thus an intrinsically higher clustering signal, so that it can actually provide a better reconstruction than the Euclid-like survey at moderate levels of calibration error. We would expect to see similarly increased robustness to calibration errors for any tracer with a larger clustering signal, including tracers with a larger bias.

Finally, just as for Euclid, we find that if calibration errors are not accounted for in the estimator, then adding LSS data can actually result in a worse reconstruction than that from using CMB temperature data alone.

2.3.4 Effect of Varying Calibration Error Properties

We now test how sensitive the results in the previous sections are to our assumptions about calibration errors, showing the results in Fig. 2.8.

First, the left panel shows what happens when we vary the level of cross-correlation between the calibration errors of different LSS maps. It is conceivable that residual calibration errors can be correlated across different bins of a single survey, or even across different surveys, especially if the error has an astronomical origin. To model such correlation, we set the level of cross-correlation between the calibration errors of maps i and j using a parameter

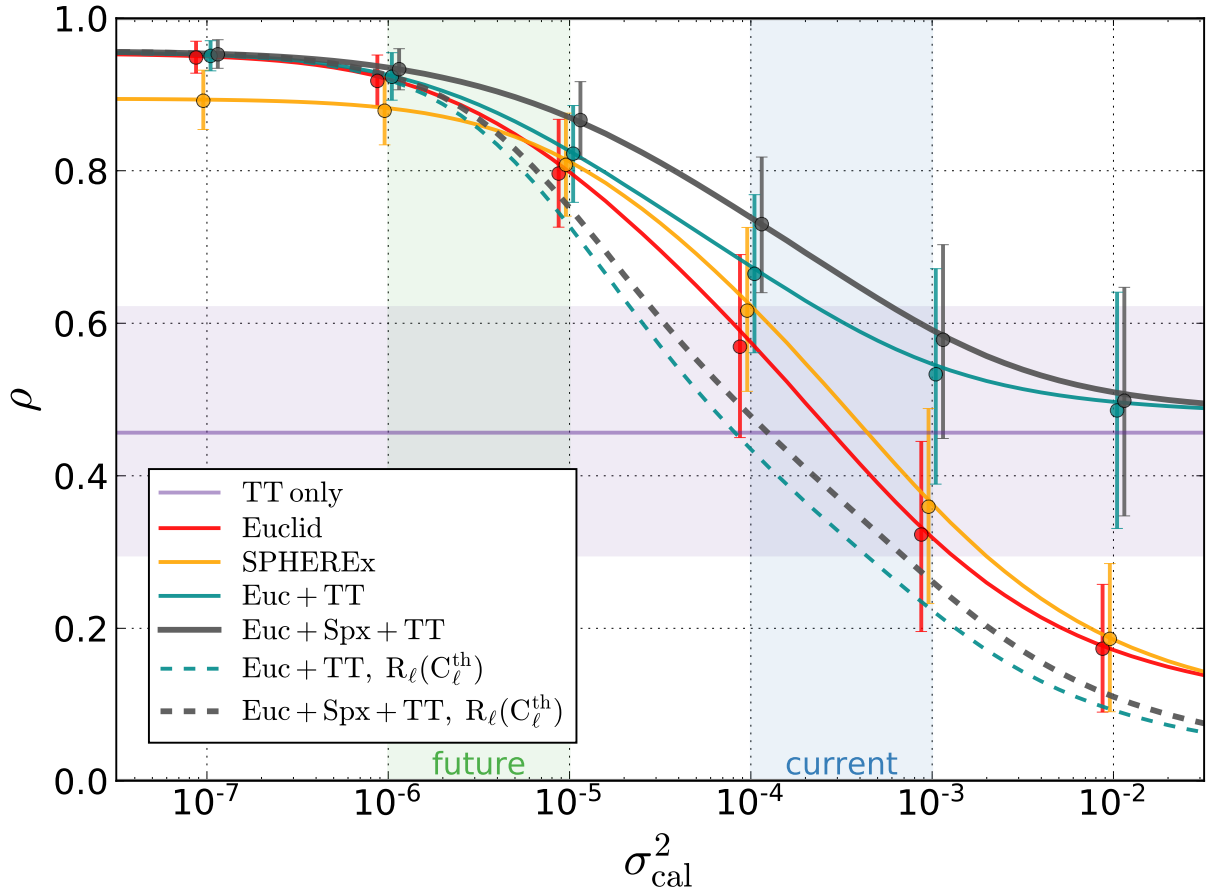

Figure 2.7: Comparison of ISW reconstruction quality using the LSS surveys and CMB temperature individually and in combination, for various levels of calibration error in the Euclid-like and SPHEREx-like surveys. Colors are the same as those of Fig. 2.6. Both of the LSS surveys are split into six redshift bins (see Sec. 2.3.1), with calibration errors uncorrelated between bins and surveys. The dashed curve shows the combined reconstruction if calibration errors are not included in the estimator. Using LSS surveys to improve the ISW map reconstruction from the CMB temperature only case requires calibration errors to be controlled to $\sigma^2_{\rm cal} \lesssim 10^{-4}$.

	R_ℓ	$R_\ell(C_\ell^{\rm th})$		$R_\ell(C_\ell^{\rm sky})$	
$\sigma^2_{\rm cal}$	0	10^{-6}	10^{-4}	10^{-6}	10^{-4}
TT	0.46	-	-	-	-
Euclid (1 bin)	0.92	0.83	0.19	0.84	0.29
Euclid (6 bin)	0.95	0.91	0.41	0.92	0.57
SPHEREx (6 bin)	0.89	0.88	0.52	0.88	0.62
Euc + Spx + TT	0.96	0.92	0.47	0.93	0.73

Table 2.1: Mean reconstruction quality coefficients $\bar{\rho}$ of ISW map reconstructions for various combinations of input maps and select levels of calibration error. The second column indicates $\bar{\rho}$ for the case of zero calibration error. The following columns show the reconstruction quality for two nonzero values of the calibration error variance; here $R_\ell(C_\ell^{\rm th})$ [$R_\ell(C_\ell^{\rm sky})$] indicates the case where calibration errors are unaccounted [accounted] for in the estimator. Note, when $\sigma^2_{\rm cal} = 0$, $C_\ell^{\rm th} = C_\ell^{\rm sky}$.

$r_{\rm cc}$, where

$$C_\ell^{{\rm cal},ij} = r_{\rm cc}\sqrt{C_\ell^{{\rm cal},ii} C_\ell^{{\rm cal},jj}}, \text{ for } i \neq j \tag{2.21}$$

As we only consider cases where calibration errors in all maps are characterized by the same $C_\ell^{\rm cal}$, this reduces to

$$C_\ell^{{\rm cal},ij} = r_{\rm cc} C_\ell^{\rm cal}, \text{ for } i \neq j. \tag{2.22}$$

We consider the six-bin fiducial Euclid-like survey and find that this kind of correlated error results in mild degradation of the reconstruction for $\sigma^2_{\rm cal} \lesssim 10^{-4}$, but otherwise it has little effect as long as calibration errors are correctly modeled in the estimator [that is, $R_\ell(C_\ell^{\rm sky})$ is used].

If calibration errors are *not* accounted for [$R_\ell(C_\ell^{\rm th})$ is used], reconstruction suffers considerably, as shown by the dotted curve. We also use a dashed curve [labeled $R_\ell(C_\ell^{\rm XY,th})$] for the case where the estimator filter correctly accounts for the autopower contributions of calibration errors but neglects the cross-power contributions. As seen by comparing the solid, dashed, and dotted orange curves in Fig. 2.8, reconstruction quality is far more sensitive to accurate modeling of the calibration error contribution to the autopower than to the cross-power. Thus, fitting the observed autopower for each map but using theoretical cross-powers, as is done in Ref. [78], should harm the reconstruction relatively little, depending on the fitting scheme; we find $\Delta\hat{\rho} \approx -0.03$ at $\sigma^2_{\rm cal} = 10^{-4}$ for $r_{\rm cc} = 0.2$, far less than the typical variation over realizations shown in Fig. 2.5.

Additionally, we study the impact of changing the shape of the calibration error power

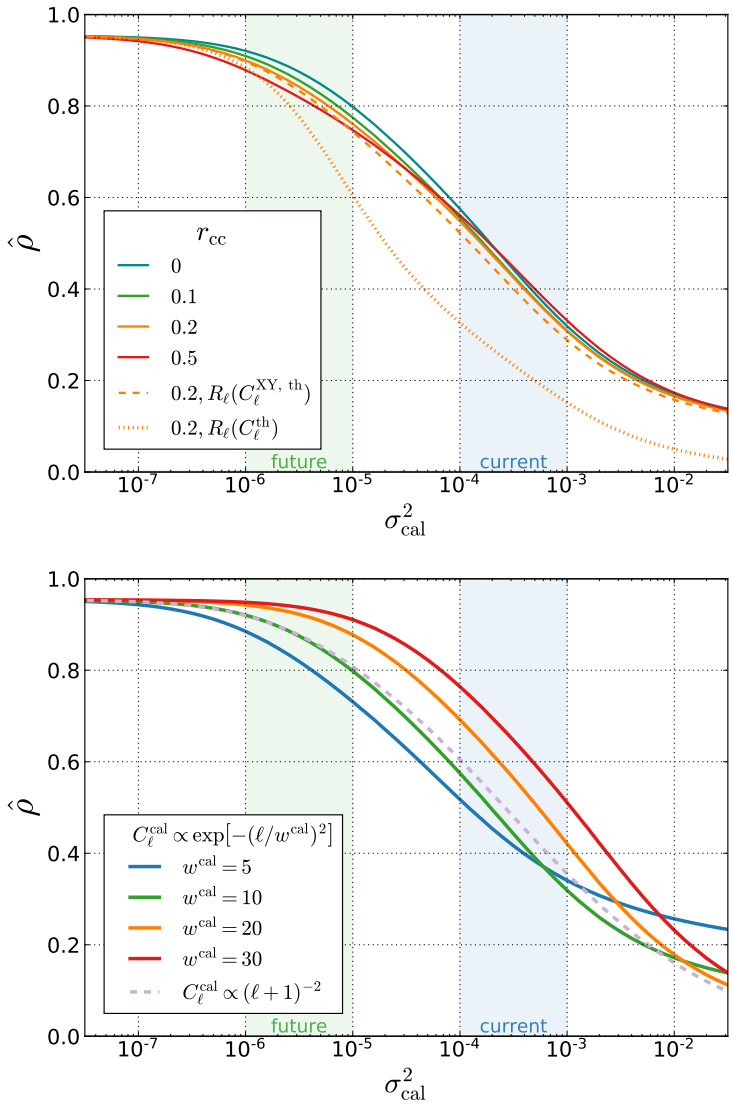

Figure 2.8: *Top:* Effect of cross-correlation between calibration errors in different bins of the fiducial Euclid-like survey, given by $C_\ell^{{\rm cal},ij} = r_{\rm cc}\sqrt{C_\ell^{{\rm cal},ii} C_\ell^{{\rm cal},jj}}$, for bins $i \neq j$. Solid curves have calibration errors accounted for in the estimator $[R_\ell(C_\ell^{\rm sky})]$. The dashed curve indicates the case where only the autopower contributions of the calibration errors are accounted for in the estimator $[R_\ell(C_\ell^{\rm XY,th})]$ and the dotted curve indicates the case where calibration errors are not accounted for at all $[R_\ell(C_\ell^{\rm th})]$. Cross-correlation of the errors results in mild degradation of the reconstruction for $\sigma^2_{\rm cal} \lesssim 10^{-4}$, but otherwise has little effect as long as the auto-correlation is correctly modeled in the estimator. *Bottom:* Dependence of ρ on the shape of C_ℓ^{cal}. Solid curves indicate C_ℓ^{cal} of the same form as Eq. (2.6) but with width $w^{\rm cal}$ varied. The dashed curve indicates the case where the error spectrum takes the form $C_\ell^{cal} \propto (\ell+1)^{-2}$. Reconstruction fares worse when calibration error power contributions are more concentrated at the largest angles, where the ISW kernel is largest. In all cases, the estimator uses the true observed LSS power spectrum ($C_\ell^{\rm sky}$).

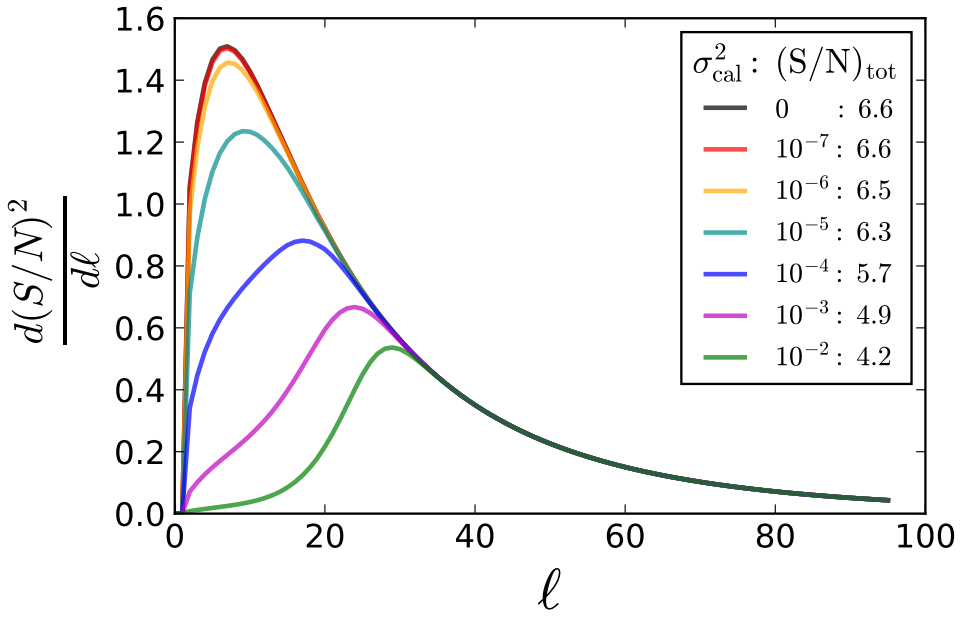

Figure 2.9: Contribution to squared signal-to-noise ratio per multipole for our fiducial Euclid-like survey with varying levels of calibration error. Total combined S/N for each level of calibration error $\sigma^2_{\rm cal}$ is given in the legend.

spectrum $C_\ell^{\rm cal}$, showing the results in the right panel of Fig. 2.8. We first vary the width parameter $w^{\rm cal}$ of the calibration error power spectrum $C_\ell^{\rm cal}$ given in Eq. (2.6). Results for different values of $w^{\rm cal}$ are qualitatively similar, though for fixed $\sigma^2_{\rm cal}$, the reconstruction is less sensitive to calibration errors when $w^{\rm cal}$ is larger. The reason for this is that ρ is most sensitive to contamination at the lowest multipoles, as will be discussed in Sec. 2.4. Using a power law $C_\ell^{\rm cal} \propto (\ell+1)^{-2}$ gives results similar to our fiducial Gaussian form with $w^{\rm cal} = 10$.

2.4 Impact of Calibration Errors on S/N Estimates

Given the extent to which calibration errors degrade the quality of ISW signal reconstruction, it makes sense to ask whether they also impact the signal-to-noise (S/N) of ISW detection. Detection of the ISW signal via the cross-correlation between the CMB and LSS maps has been the subject of considerable study, as it serves as an important consistency test for the presence of dark energy. The significance of detection varies considerably depending on the LSS tracer sample and the statistical methods used [112], as well as how systematics in the LSS data are treated [71, 89, 92]. Ref. [89] used systematics templates to correct the observed power spectra for SDSS galaxies, finding a S/N loss of ~ 0.5 if such corrections are neglected. Ref. [71] introduced a low-redshift spike in the source distributions in order

to reproduce the observed excess autopower in NVSS and SDSS catalogs and estimate that such systematics result in an uncertainty of $\Delta \mathrm{S/N} \pm 0.4$. The most recent results come from the Planck Collaboration, which found $\sim 4\sigma$ evidence for the ISW effect, with most of the signal coming from cross-correlation of the CMB temperature with the NVSS radio catalog and CMB lensing [26].

The maximum achievable signal-to-noise can be obtained by considering an ideal survey that perfectly traces the ISW (i.e. $C_\ell^{gg} = C_\ell^{Tg} = C_\ell^{\mathrm{ISW}}$), resulting in a maximum $\mathrm{S/N} \sim 6-10$ for ΛCDM cosmology [68, 69, 87, 103, 105, 112, 113].

Our goal is to study how calibration errors impact the significance of ISW detection. There are multiple ways one can quantify detection of the ISW effect, including correlation detection between LSS and the CMB, template matching to an assumed model, or model comparison. Each of these methods relies on different assumptions and tests different statistical questions (see Ref. [112] for a detailed review). Here we adopt the simple correlation detection statistic which quantifies the expected deviation from a null hypothesis of no correlation between LSS (g) and CMB temperature (T). In this formalism the S/N for ISW detection is

$$\left(\frac{S}{N}\right)^2 \simeq \sum_\ell (\mathbf{C}_\ell{}^{Tg})^* (\mathbf{C}_\ell^{\mathrm{cov}})^{-1} \mathbf{C}_\ell{}^{Tg}, \tag{2.23}$$

where we have assumed the multipoles contribute independently to the S/N. Here $\mathbf{C}_\ell{}^{Tg}$ is a vector of the ISW-LSS cross-spectra, and the covariance matrix elements corresponding to LSS maps i and j can be written as

$$C_\ell^{\mathrm{cov},ij} = \left\langle \Delta C_\ell^{Ti} \Delta C_\ell^{Tj} \right\rangle \tag{2.24}$$

$$\simeq \frac{C_\ell^{Ti} C_\ell^{Tj} + C_\ell^{TT} (C_\ell^{ij} + C_\ell^{\mathrm{cal},ij} + \delta_{ij} \frac{1}{\bar{n}_{ij}})}{f_{sky}(2\ell+1)}, \tag{2.25}$$

where the last term in the numerator is due to shot noise and δ_{ij} is the Kronecker delta.[4] Equations (2.23) and (2.25) demonstrate that all cosmological tests using LSS-CMB cross-correlation are limited in their constraining power due to sample variance and the relatively large amplitude of the primordial CMB fluctuations. They also make it clear that calibration errors will reduce the significance of ISW detection.

We assume calibration errors to be uncorrelated between maps, so $C_\ell^{\mathrm{cal},ij} \rightarrow \delta_{ij} C_\ell^{\mathrm{cal},ij}$.

[4] Strictly speaking, this will result in a slight underestimate of the significance, as technically the null hypothesis covariance, with $C_\ell^{Ti} = C_\ell^{Tj} = 0$ in Eq. (2.25), should be used. However, as the galaxy-ISW cross-power terms are small compared to the galaxy autopower, we follow the practice in most of the literature of keeping them in the S/N calculation.

For a single LSS map, Eq. (2.23) reduces to the form

$$\left(\frac{S}{N}\right)^2 \simeq f_{sky} \sum_\ell \frac{(2\ell+1)(C_\ell^{Tg})^2}{(C_\ell^{Tg})^2 + C_\ell^{TT}(C_\ell^{gg} + C_\ell^{\rm cal} + 1/\bar{n}_g)}. \tag{2.26}$$

If there are no calibration errors, we find S/N= 6.6 for our Euclid-like survey, which is near the maximum[5] for this cosmology, S/N= 6.7. As $\sigma^2_{\rm cal}$ increases from 0 to current levels, the total S/N reduces to $4.9-5.7$, a drop of only $\sim 15\%-30\%$. This can be seen in the S/N values listed for various $\sigma^2_{\rm cal}$ in the legend of Fig. 2.9. In contrast, for the same level of error, average reconstruction quality $\hat{\rho}$ is reduced by $40\%-60\%$. Clearly, ISW signal reconstruction is substantially more affected by calibration errors than is ISW detection significance.

The greater robustness of the total S/N to calibration errors is due to the fact that it has support at higher multipoles. This is most easily illustrated in the single-map case, where the contribution per multipole to the total signal-to-noise is

$$\left(\frac{S}{N}\right)^2_\ell \equiv \frac{d\,(S/N)^2}{d\ell} = (2\ell+1)\frac{(C_\ell^{Tg})^2}{C_\ell^{TT}C_\ell^{gg} + (C_\ell^{Tg})^2}. \tag{2.27}$$

Figure 2.9 shows how the contribution per multipole responds to different levels of calibration error.

As $\sigma^2_{\rm cal}$ increases, the signal-to-noise decreases at lower multipoles, but contributions at higher multipoles remain unchanged. These higher-multipole contributions are thus still available to contribute to the overall S/N.

Map reconstruction is more sensitive to the largest scales. For the single-map case, this can be illustrated analytically as follows. Using the single-map estimator from Eq. (2.10),

[5]This limit can in principle be increased, e.g., through the method of Ref. [114] in which the observed LSS map is used to reduce the local variance and which in our case brings the maximum possible S/N to 7.2, or through the inclusion of polarization data as in Ref. [115].

we can write the estimated reconstruction quality statistic as

$$\begin{aligned}
\hat{\rho} &= \frac{\frac{1}{4\pi}\sum_\ell (2\ell+1)\left(\frac{C_\ell^{Tg}}{C_\ell^{gg}}\right) C_\ell^{Tg}}{\sigma_{\rm ISW}\sqrt{\frac{1}{4\pi}\sum_\ell (2\ell+1)\left(\frac{C_\ell^{Tg}}{C_\ell^{gg}}\right)^2 C_\ell^{gg}}} \\
&= \frac{1}{\sigma_{\rm ISW}}\sqrt{\frac{1}{4\pi}\sum_\ell (2\ell+1)\frac{(C_\ell^{Tg})^2}{C_\ell^{gg}}} \\
&= \frac{1}{\sigma_{\rm ISW}}\sqrt{\frac{1}{4\pi}\sum_\ell \left(\frac{S}{N}\right)_\ell^2 C_\ell^{TT}\left(1+\frac{(C_\ell^{Tg})^2}{C_\ell^{gg}C_\ell^{TT}}\right)} \\
&\approx \frac{1}{\sigma_{\rm ISW}}\sqrt{\frac{1}{4\pi}\sum_\ell \left(\frac{S}{N}\right)_\ell^2 C_\ell^{TT}}. \qquad (2.28)
\end{aligned}$$

Here, $(S/N)_\ell^2$ is the quantity given by Eq. (2.27) which, when summed over ℓ, gives $(\mathrm{S/N})^2$. Thus, we see from Eq. (2.28) that $\hat{\rho}$ is proportional to a total (S/N) of which the terms are weighted by C_ℓ^{TT}. Since C_ℓ^{TT} drops sharply as $\sim \ell^{-2}$, the quality of map reconstruction $\hat{\rho}$ is more impacted by large-angle calibration errors than the overall S/N is. This is also a primary cause for the degradation in reconstruction quality seen when $\sigma^2_{\rm cal}$ was concentrated at lower multipoles in Sec. 2.3.4.

2.5 Conclusions

Reconstruction of the integrated Sachs-Wolfe signal would allow, for the first time, a clean separation of the CMB temperature anisotropies into contributions from 300,000 years after the big bang and those from some ∼10 billion years later. This, in turn, would allow for a more informed assessment of the origin of the "large-angle CMB anomalies" and a more complete elimination of ISW contaminants to CMB-based measurements of primordial non-Gaussianity. Accurate ISW reconstruction requires wide-angle large-scale structure maps from which the gravitational potential evolution can be inferred, but in practice, these maps are plagued by photometric calibration errors due to a host of atmospheric, instrumental, and selection-induced systematics. In our previous work, we found that the realistic levels of residual calibration error severely degrade the accuracy of the reconstructed ISW map.

In this chapter, we investigated how the effects of residual photometric calibration errors on the ISW map reconstruction can be mitigated by using tomographic information and by combining data from multiple surveys. To quantify the amount of residual calibration errors,

we use their variance $\sigma^2_{\rm cal}$, the square root of which is roughly equal to the rms *magnitude* fluctuations induced by these systematics.

We find that for a Euclid-like survey with a single redshift bin, to achieve a reconstruction comparable in quality to that derived from the CMB temperature map alone (with an average correlation between the true and reconstructed ISW maps of only $\bar{\rho} \approx 0.46$), one must limit the variance of the calibration error field to $\sigma^2_{\rm cal} \lesssim 10^{-5}$.

This can be improved significantly if we exploit the tomographic information available by binning the LSS data in redshift (Fig. 2.5). We also show that if the model spectra in the estimator differ substantially from those used to generate the input maps e.g. by using theoretical power spectra that do not account for the observed excess autopower in the LSS survey from calibration errors, then reconstruction quality is substantially degraded. It is therefore crucial to verify that the theoretical spectra in the estimator are a good fit to those observed or to otherwise use smoothed fits.

We next consider how using multiple input maps, probing different tracers and redshift ranges, improves ISW signal reconstruction. We find that as long as the excess power contributed by calibration errors is adequately modeled in the estimator, the resulting reconstruction is always better than that from either of the input maps individually. If the excess power from calibration errors is not modeled, however, adding a map can actually degrade reconstruction (Fig. 2.6). The CMB temperature map adds information to the reconstruction at all levels of calibration error, but is especially valuable if the LSS maps are subject to calibration errors with $\sigma^2_{\rm cal} \gtrsim 10^{-5}$. Using a six-bin SPHEREx-like survey provides qualitatively similar results to the Euclid-like one, but because it is shallower, the reconstruction is less accurate in the limit of no calibration errors ($\rho_{\rm Spx} = 0.89 \pm 0.04$ vs. $\rho_{\rm Euc} = 0.95 \pm 0.02$, where errors indicate the 68% spread across realizations). However SPHEREx's shallower depth and therefore intrinsically higher clustering signal becomes an asset in the presence of calibration errors, making the survey more robust against calibration errors and leading to a better reconstruction for $\sigma^2_{\rm cal} \gtrsim 6 \times 10^{-6}$. (Similarly, we would expect the increased clustering of tracers with larger bias to help mitigate the effects of calibration error as well.) Therefore, a combination of a shallower and a deeper LSS survey provides complementarity useful for separation of calibration errors from the ISW signal and necessary for a good ISW map reconstruction in the presence of such errors.

Using all three simulated surveys as input—Euclid, SPHEREx, and CMB temperature—significantly improves reconstruction compared to using a single survey with current levels of residual calibration errors, or CMB temperature data alone. We find that if the calibration errors can be controlled to the level of $\sigma^2_{\rm cal} \lesssim 10^{-6}$ ($\sigma^2_{\rm cal} \lesssim 10^{-5}$), then the combination of Euclid, SPHEREx, and CMB temperature maps can produce the ISW map reconstruction

to an excellent accuracy of $\rho = 0.93 \pm 0.03$ ($\rho = 0.87 \pm 0.05$). This is roughly the same level of calibration control required for future LSS surveys to avoid biasing measurements of cosmological parameters like the non-Gaussianity parameter f_{NL} and the dark energy equation of state. Thus, high-quality ISW reconstructions will, in a sense, "come for free" with the developments needed for cosmology constraints with next-generation surveys.

We additionally test the robustness of our results against changes to the properties of the calibration error field, looking at cross-correlations between calibration errors in different maps as well as the shape of their spectrum. We found that cross-correlation between the calibration errors of different tracer maps degrades the reconstruction most for $10^{-6} \lesssim \sigma^2_{\rm cal} \lesssim 10^{-4}$, but that this effect is relatively minor, provided the auto-correlation contribution is accounted for in the estimator (Fig. 2.8, left).

We also find that, compared to map reconstruction accuracy, the overall significance of ISW detection is less strongly affected by calibration errors. This is because its signal is distributed more broadly in multipole space than that of the map reconstruction quality statistic. To clarify this, we analytically relate $\hat{\rho}$ to the commonly used ISW detection S/N statistic in the case of a single input LSS map and show that $\hat{\rho}$ is weighted by an additional factor of C_ℓ^{TT}, demonstrating map reconstruction's greater sensitivity to the largest scales [Eq. (2.28)].

As an extension to this work, one could study how the inclusion of CMB lensing and polarization data can improve ISW map reconstruction, provided the systematics present in those data sets could be sufficiently accurately modeled. The results of [82] indicate that the use of lensing as input can contribute significantly to reconstruction quality, but they also show that current noise levels limit its effectiveness. Notably, the residual lensing systematics at $\ell \leq 8$ present a challenge, as this is where the ISW signal is strongest, and we expect these systematics to affect reconstruction with CMB lensing and polarization in a manner broadly similar to unaccounted for calibration errors in LSS maps at those scales.

Further work could also be performed to concretely explore how to best approximate the 'best case' reconstruction scenario, wherein calibration errors are fully accounted for, by using real input data. Here we have only characterized the limiting cases where the residual calibration error contribution to the LSS power is fully known or fully unknown, and have not addressed intermediate cases where they are partially accounted for.

Finally, we have only worked in the full-sky case whereas real data will necessarily have only partial sky coverage. Others have already shown that incomplete sky coverage only very minorly degrades reconstruction quality for areas covered by the input data sets [82], and we do not expect this to change in the presence of calibration errors. Nevertheless, a full analysis that attempts to make predictions for real surveys should take the actual sky

coverage and survey-specific systematics into account.

Even with these considerations, it is clear that accounting for the presence of residual calibration errors in LSS surveys is a critical step for any reconstruction of the ISW map, as their presence and treatment impact both the survey characteristics and set of input maps that produce the optimal reconstruction.

Chapter 3

Mitigating Contamination in LSS Surveys: A Comparison of Methods

3.1 Introduction

Over the past 40 years, cosmological surveys have produced increasingly detailed maps of the large-scale structure (LSS) in the Universe [116–124]. These observations have proven crucial for testing our understanding of gravity and cosmological structure formation, and helped to constrain cosmological parameters to the percent level [e.g. 124–126]. Recent observations from DES have for the first time imposed strong constraints on dark energy using an LSS survey alone, independently of the cosmic microwave background [127]. Upcoming ground-based missions like DESI [128], and the Rubin Observatory's LSST [129], along with space-based missions like SPHEREx [130], Euclid [131], and RST (formerly WFIRST) [132] will truly herald the age of precision cosmology, mapping up to ∼20 billion galaxies across the sky and bringing unprecedented precision to measurements of the dark energy equation of state and modified gravity. Such statistical precision makes the control of systematic errors in these datasets of paramount importance to avoid biasing cosmological analyses.

Cosmological information is extracted from LSS observations in multiple ways. The most common approach is to calculate the two-point correlation function or its Fourier counterpart, the power spectrum, to characterize the spatial distribution of galaxies (galaxy clustering) or their shapes (weak lensing). To date, these have been used in cosmological analyses to great success [133–152].

The two-point function contains all available information when the field it characterizes is Gaussian, but nonlinear gravitational collapse induces non-Gaussianity at late times and small scales. Therefore there is considerable cosmological information that is inaccessible to

the two-point function. This has led to growing interest in using complementary statistical representations of LSS observations, such as higher order N-point functions [153, 154, 154, 155, 155–162], statistics of peaks [163–165] and voids [166–172], density-split statistics [173], the log power spectrum [174], marked power spectra [175–178], Minkowski functionals [179–184], wavelet transforms [185, 186] and more. These methods rely on the accurate mapping of the underlying cosmological fields from which they are derived, so there is increasing need for tools to mitigate systematic contamination at the levels of both the map and the two-point functions.

Here, we consider the very general class of systematics that describes an arbitrary spatial modulation of the observed field. Such generic sources of error are one of the most serious contaminants in our quest to probe cosmology with future surveys. For definiteness we focus on the case of galaxy clustering, where the systematic error corresponds to a modulation of the galaxy selection function in redshift or across the observing footprint, similar to that discussed in Chapter 2. **However, the methods we test and develop in this chapter are general enough to apply to any real or complex field for which there exist maps of potential contaminants** (e.g. shear or Sunyaev-Zeldovich-effect fields).

Spatially-varying systematics in LSS maps may be caused by an large variety of physical effects. These include observing conditions and dust extinction (both of which effectively create a position-dependent "screen," obscuring background galaxies), bright objects and star-galaxy separation (which can eclipse, change the shape, or be confused for galaxies close to them on the sky), and variations in sensitivity of the detector (which include potentially time- and position-dependent variations in the focal plane), or imaging pipeline. In all of these cases, failure to fully account for variability in the selection function will result in residual artifacts — *calibration errors* — in the final data product and potentially bias results [16, 101, 102]. The presence of calibration errors is evidenced by a number of surveys [83, 85–88, 187–190] which have shown a significant excess of power at large scales where calibration errors are thought to be most prevalent. Recent observations (e.g. from the Dark Energy Survey [191]) demonstrate however that such contamination is by no means limited to large scales alone. In addition to adding power, calibration errors induce a multiplicative effect, coupling different scales and thus affecting all scales in the survey, including those smaller than the typical size of the calibration systematic itself [101, 102]. Much recent work [83, 88, 94–96, 98, 100, 110, 190, 192–203] has focused on mitigating these systematics in order to probe the underlying cosmology.

The simplest strategy to ameliorate the effects of calibration errors is to simply mask scales or data points suspected having large levels of contamination. More sophisticated strategies include using maps of suspected contaminants — so-called 'templates' — to cor-

rect the observations. An alternative and complementary approach is to forward-model many possible realizations of the cosmic initial conditions [e.g. 204–211]. One then evolves these initial conditions in time (while adding realizations of nonlinearities, bias, and observational/instrument systematics), and performs joint inference of cosmology, the initial conditions and late-time "true" fields, given observations. Another forward-modeling approach involves the injection of false images into observations in order to sample the selection function [100, 202]. While such forward approaches are powerful and very general, they also require extensive computational resources and are complicated to implement. In contrast, using templates to clean contaminated observations and directly infer the underlying fields is straightforward to implement and can be readily incorporated into ongoing or completed analyses. They have been the dominant approach in the community thus far, and so these are the methods we focus on here. Fig. 3.1 shows real-world results of the template-based correction scheme used in the just-released Year 3 analysis from the Dark Energy Survey. The left panel shows how the raw observed two-point correlation function of the `redMaGiC` catalog is artificially high (red), and is reduced significantly after template-based corrections are applied (green) (see Ref. [23] for details). The right panel shows sky maps with the characteristic DES footprint of some of the survey and sky properties used as templates in the correction scheme.

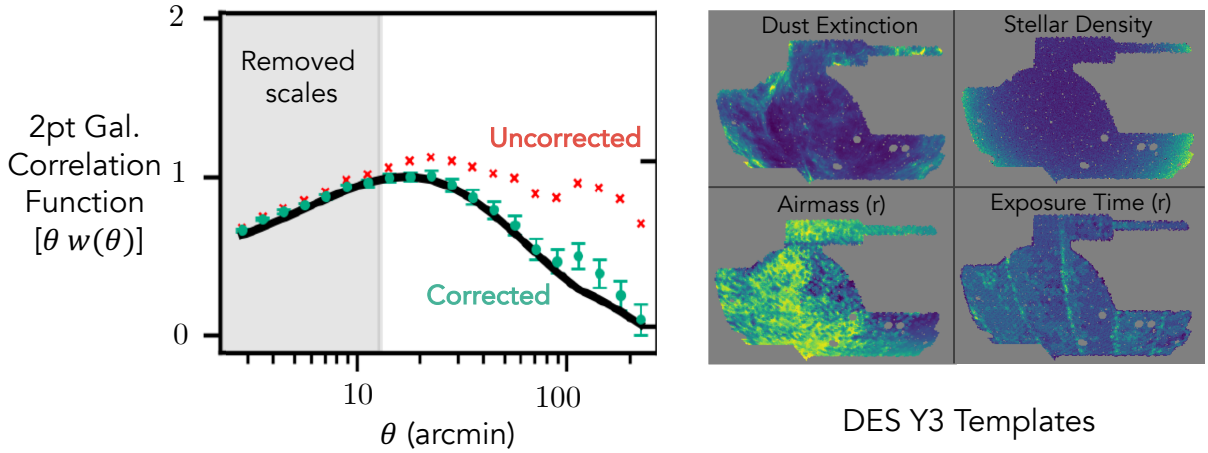

Figure 3.1: *Left:* The two-point correlation function of DES Y3 `redMaGiC` galaxies with $0.8 < z \leq 0.9$. The uncorrected sample (red) shows significantly higher galaxy clustering than after the template-based corrections are applied (green), demonstrating the importance of cleaning in modern-day LSS analyses. The black line shows the best-fit ΛCDM theory curve from the fiducial 3×2-pt analysis (which includes weak lensing data) and is in much better agreement after correction. Adapted from Ref. [23]. *Right:* Real world sky maps of several observational quantities used as templates for the correction scheme in the DES Y3 analysis.

In this work we revisit and extend state-of-the-art LSS systematics-cleaning strategies.

We interpret them through a regression framework to highlight commonalities and differences of the methods, as well as some tacit assumptions. In doing so, we show that the common pseudo-C_ℓ Mode Projection method is equivalent to linear regression. We use this framework to propose straightforward extensions that leverage the extensive body of literature and tools that have been developed for regression analyses. We rigorously test the performance of several existing methods, plus new ones that we propose, on a common set of simulated observations from current and future surveys.

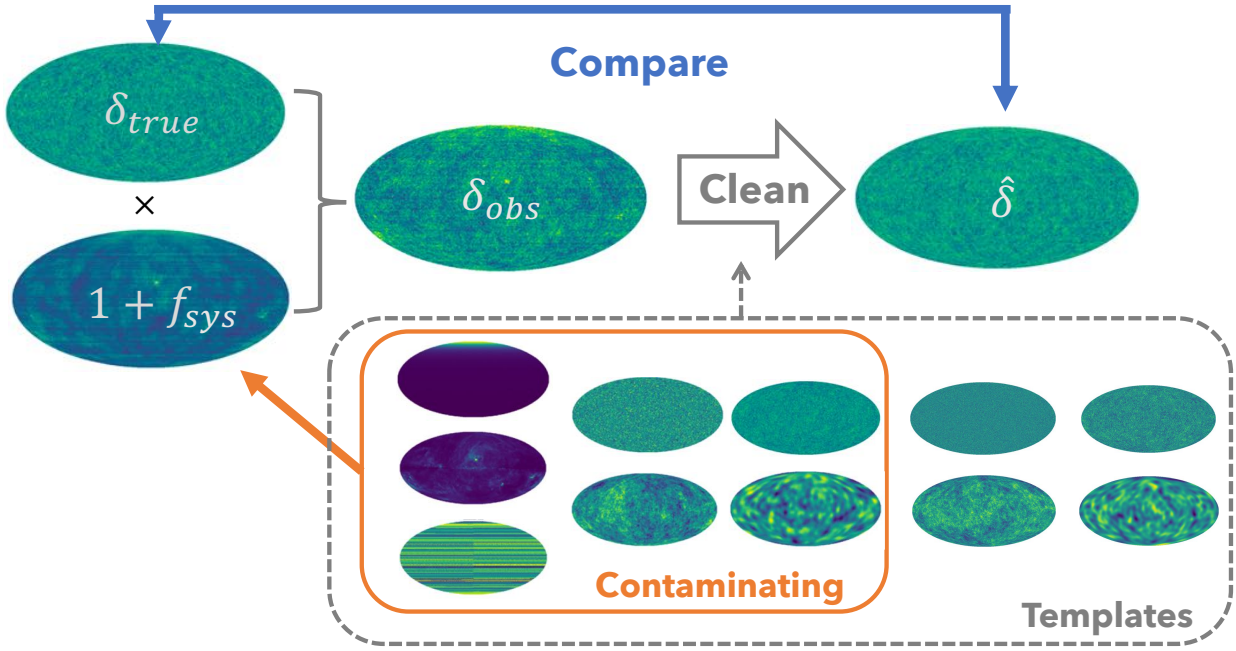

Figure 3.2: Analysis procedure for a single map. A set of templates is generated (dashed box) along with a true overdensity map $\delta_{\rm true}$. A subset of the templates (orange box) contaminate the true overdensity map to generate the observed overdensity field $\delta_{\rm obs}$. We generate an estimated signal map $\hat{\delta}$ using one of the cleaning methods, and compare it to the truth, either at a map-level or power-spectrum-level. This is repeated for many realizations of the signal map and the performance of each cleaning method is assessed.

We study performance using an ensemble of simulated galaxy overdensity maps, such that we can assess both the accuracy and precision of each method. We provide a library of templates and a contaminated overdensity map as input to each cleaning method, which then produces an estimate of the true overdensity map and power spectrum that we assess for accuracy. We repeat the process over a large number of sky realizations and for various configurations of templates to asses the precision and robustness of each method. A schematic outline of this process is shown in Fig. 3.2.

The chapter is organized as follows. In Sec. 3.2 we describe in detail our general model for contamination, which encompasses a wide range of systematics due to foregrounds or

instrument calibration errors. In Sec. 3.3 we describe several existing methods for systematics mitigation; in Sec. 3.4 we reinterpret the methods through a common framework to facilitate comparison, and in Sec. 3.5 we use this to map several aspects onto well-known techniques in statistics and propose two new mitigation methods. In Sec. 3.6 we describe the fiducial synthetic surveys on which we test the efficacy of the methods that we study. Sec. 3.7 shows the results of these performance comparisons, while Sec. 3.9 has our conclusions. Several Appendices show important but more technical and detailed aspects of the investigation.

3.2 Contamination Model

We first introduce the model for contamination of the observed LSS fields. It is very general, encompassing most known sources of real-world contamination. We can model the observed number density map as a combination of the true galaxy number density map ($N_{\rm true}(\hat{\mathbf{n}})$) modulated by a direction-dependent screen ($1 + f_{\rm sys}(\hat{\mathbf{n}})$), plus an additive contamination term $N_{\rm add}(\hat{\mathbf{n}})$:

$$N_{\rm obs}(\hat{\mathbf{n}}) = (1 + f_{\rm sys}(\hat{\mathbf{n}}))N_{\rm true}(\hat{\mathbf{n}}) + N_{\rm add}(\hat{\mathbf{n}}). \tag{3.1}$$

We will primarily address multiplicative contamination as this characterizes most known LSS contaminants; one exception is a contaminating population of objects such as stars, which we discuss briefly in Sec. 3.4.2. Therefore, we take $N_{\rm add}(\hat{\mathbf{n}}) \to 0$ for simplicity and focus on the first term $f_{\rm sys}(\hat{\mathbf{n}})$, which fully characterizes the systematic modulation of the true field such that pixels with $f_{\rm sys}(\hat{\mathbf{n}}) = 0$ are free of contamination. Using $N = \bar{N}(1 + \delta)$ and defining the ratio of true to observed mean number density as $\gamma = \bar{N}_{\rm true}/\bar{N}_{\rm obs}$, the observed overdensity can be written as

$$\delta_{\rm obs}(\hat{\mathbf{n}}) = \gamma(\delta(\hat{\mathbf{n}}) + 1)(f_{\rm sys}(\hat{\mathbf{n}}) + 1) - 1. \tag{3.2}$$

Here γ enforces the constraint that $\langle\delta_{\rm obs}\rangle_{\rm pix} = 0$ across the survey footprint, even though this is not necessarily true for the *true* overdensity field δ. This is due to the fact that we can only access the observed mean number density $\bar{N}_{\rm obs}$, which differs from the true mean both because of systematic contamination and because of sample variance from a limited survey footprint (see Sec. 3.4.2 for details).

This model for contamination is similar to the one used in [16, 17, 101, 102] to assess the impacts of *residual* calibration errors that remain in the data after cleaning. Here we focus on the methods used to perform such cleaning, and so use the screen model to describe contamination more generally.

We extend the screening formalism by considering that the total systematic modulation

is comprised of $N_{\rm sys}$ individual systematics, each of which acts as its own screen. Thus we have

$$1 + f_{\rm sys} = \prod_{i=1}^{N_{\rm sys}} (1 + f_i) \tag{3.3}$$
$$\simeq 1 + \sum_{i=1}^{N_{\rm sys}} f_i + \sum_{j \neq k}^{N_{\rm sys}} f_j f_k + [\text{higher order terms}]$$

where we have suppressed $\hat{\mathbf{n}}$ in the notation for convenience both here and in what follows. Note that even if a systematic individually contributes to $f_{\rm sys}$ linearly, there exist interaction terms with other systematics up to order $N_{\rm sys}$. Here and in general, $f_i \equiv f_i(\hat{\mathbf{n}})$ is a column vector with each element corresponding to a pixel, unless otherwise noted.

3.3 Background: Existing Mitigation Methods

The principal goal of this study is to compare various proposed systematics mitigation methods. The methods that we test are all designed to use maps that trace potential contamination in order to mitigate the impact of systematics, i.e. they assume that the systematic $f_i(\hat{\mathbf{n}})$ is a function of some tracer $t_i(\hat{\mathbf{n}})$. We refer to these tracer maps as *templates*, and examples include maps of stellar density, extinction, or summary statistics of observing conditions (e.g. mean g-band seeing) in each region of the sky throughout the duration of the survey (see Leistedt et al [191] for a detailed description of the process for creating templates from multi-epoch observational data for the Dark Energy Survey). Sources of error for which we have no templates (e.g. shot noise) are implicitly subsumed into the overdensity field.

We will investigate how effectiveness depends on analysis choices and suggest improvements where possible. We start with three principal methods that have been applied in the literature: the Dark Energy Survey Year 1 method (henceforth DES-Y1), the Template Subtraction method (TS), and the Mode Projection method (MP). While at face value the algorithms associated with these methods seem quite different, we demonstrate that they can be translated into a common mathematical framework of linear regression. Doing so allows us to distill commonalities and differences between the methods, as well to identify simplifications and extensions to them. We include three additional methods based on these insights.

For all the methods, we will work with maps that are divided into pixels in `HEALPix`[1] [12]

[1] `http://healpix.sourceforge.net`

format, which summarize the mean galaxy overdensity or template values within each pixel (see Sec. 3.6 for details). Furthermore, while we work in the context of cleaning galaxy overdensity fields, the methods are applicable more generally to corrections of any field for which we have templates of potential contamination, and so we denote the true signal more generally as s and the observed field as $d_{\rm obs}$. In our application, these correspond to the true and observed galaxy overdensity fields, $\delta_{\rm true}$ and $\delta_{\rm obs}$. In the sections that follow, we use $\hat{x}$ to denote an estimate of x, and $\tilde{C}_\ell^{xx}$ to indicate a realization-specific measurement of the power spectrum, as compared to its theoretical mean C_ℓ^{xx}.

3.3.1 Dark Energy Survey Y1 Method

The method used to derive galaxy weights for the Year-1 DES release is one of the more sophisticated mitigation methods applied to date. It is described in detail in Elvin-Poole et al [212] (EP17), but we review its main features here. Hereafter referred to as 'DES-Y1,' builds on the method first proposed as the 'Weights' method in Ross et al [94], wherein 1-dimensional relationships between observed galaxy densities and systematic templates (there called 'survey property maps') are removed by iteratively applying multiplicative factors ('weights') to galaxies. Fig. 3.3 shows one example of how the observed overdensity varies with a template. Multiplicative weights are applied to galaxies to de-trend the data, shifting the blue line to lie atop the dashed line. This method is explicitly a regression method, with versions employing linear fits [213–215], splines [89] or higher-order polynomials [216] as fitting functions for the 1D relationships.

Here we describe the version we adopt, which closely follows the implementation in EP17 used on the DES-Y1 data. For each template t_i, we group pixels into 10 evenly-spaced bins based on their template values, and independent of location on the sky (e.g. all pixels with a mean i-band seeing value within 10% of the max would be grouped). We then find the mean galaxy overdensity over the pixels in each bin[2]. A 10×10 covariance matrix of these bin means is estimated by performing the same bin-averaging process on a set of 400 uncontaminated mock maps, generated with a fiducial power spectrum for the overdensity field (we assume the true overdensity power spectrum to generate these mocks).

Next, we use `scipy.optimize` and the estimated covariance to find the parameters

[2]In EP17, extreme regions are removed by eye: each template is inspected and bins that exhibit an average fluctuation in number density of $> 20\%$ are masked, as are regions where visual inspection suggests a deviation from non-monotonic behavior (see their Fig. 3). We neglect this step, as it is difficult to automate robustly and in our tests we found that it did not alter our results.

$\{m_i, b_i\}$ of the best-fit line[3] of the binned overdensity to each binned template i:

$$\frac{\langle N_{\rm obs}\rangle_j}{\bar{N}_{\rm obs}} = m_i \langle t_i\rangle_j + b_i \tag{3.4}$$

where $\langle\cdot\rangle_j$ indicates the average pixel value in bin j of the given template. See Fig. 3.3 (blue points and trend) for an illustration.

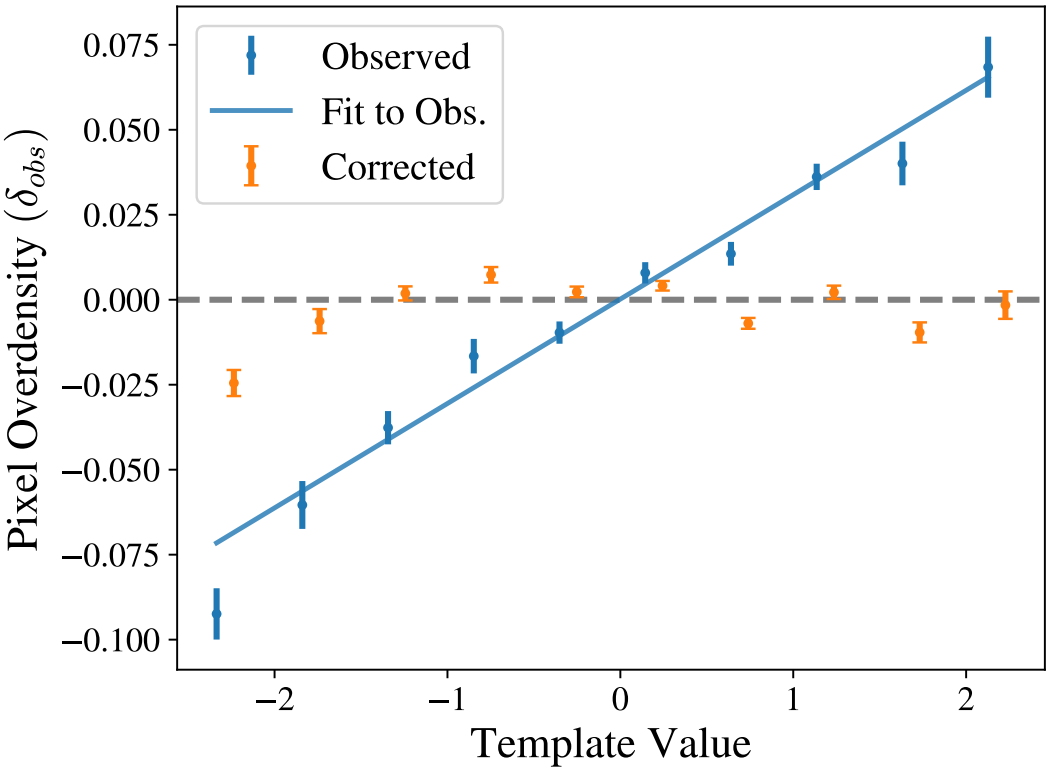

Figure 3.3: Illustration of the DES-Y1 cleaning method, showing the total observed pixel overdensity ($\delta_{\rm obs}$) as a function of a *template's* pixel overdensity, in ten evenly-spaced bins. Given the estimated covariance matrix (diagonals shown by blue error bars), the best-fit trend (blue line) can be calculated and used to reweight the observed map, producing a corrected map whose dependence on the template is removed (orange points, with corresponding standard errors on the pixel means). The process is then iterated for other templates until a satisfactory threshold is reached; see text for details.

The template with the most significant fit is used to reweight the number density in each pixel as $N'_{\rm obs}(\hat{\mathbf{n}}) = N_{\rm obs}(\hat{\mathbf{n}})/(\hat{m}_i t(\hat{\mathbf{n}}) + \hat{b}_i)$, where the significance metric is defined below. Having removed the effect of the dominant systematic, the whole process is repeated: for

[3]EP17 also use linear fits for almost all templates, with only a couple exceptions. As noted in Sec. 3.4.2, even if a template is thought to contaminate non-linearly, the relationship can usually be made linear through an appropriate transformation of the template.

each template, the pixels are assigned to bins and averaged, the new best-fit parameters are computed from Eq. (3.4), and the trend from the most significant template is removed from the data. The process stops when all templates are below a predefined significance threshold.

In general, the more contamination from template i, the stronger relationship the relationship with the observed galaxy density. However, some level of correlation is expected just by chance, and this depends on the spatial clustering of each template. The DES-Y1 method addresses this in two ways: (1) by using a different covariance matrix for the observed overdensity for each template as described above, and (2) by having a template-specific significance threshold, calibrated on mocks. Specifically, the significance statistic used is $\Delta\chi^2_i/[\Delta\chi^2_i]_{68}$, where $\Delta\chi^2_i$ is the improvement in χ^2 for the binned fit on template i, compared to a null hypothesis of $m_i = b_i = 0$. It is normalized to the $68^{\rm th}$ percentile of the same quantity measured on uncontaminated signal mocks ($[\Delta\chi^2_i]_{68}$). We use the stopping criterion $\Delta\chi^2_i/[\Delta\chi^2_i]_{68} < \Delta\chi^2_{\rm threshold} = 2$, but find that our results change little when changing this threshold between 1 and 4 (see App. B.6).

There are a number of required parameter choices in the DES-Y1-type method. These include the criterion for selecting the most significant template,[4] the significance threshold that determines when to stop weighting, the prior power spectrum for generating mocks, and choices associated with binning (e.g. number of bins, equally-spaced vs. equally-filled, etc). Here we use the fiducial choices from EP17, and investigate some of the effects of these choices in App. B.6.

3.3.2 Template Subtraction

The Template Subtraction method uses the cross-power of templates with the observed sky to estimate contamination of each template at each angular scale. Contamination is subtracted directly from the two-point clustering statistics. The method was proposed in Ho et al [190] and Ross et al [94] where it was called the "cross-correlation" technique, and we review it here.

Template Subtraction assumes the observed overdensity $d_{\rm obs}$ is a linear combination of the true galaxy overdensity s and individual template overdensities t_i:

$$d_{\rm obs} = s + \sum_{i=1}^{N_{\rm tpl}} \alpha_i t_i. \tag{3.5}$$

Any systematics or noise not accounted for by templates are subsumed into the signal s. In

[4] E.g. one could consider an R^2 statistic, the commonly-used F-statistic, Akaike or Bayesian information criteria, etc.

Ho et al [190], $d_{\rm obs}$ and t_i are taken to be in multipole space, such that $\langle ss\rangle \to \langle s_{\ell m} s_{\ell m}\rangle = C_\ell^{ss}$ (where $\langle\cdot\rangle$ is the ensemble average over many sky realizations), and $\alpha \to \alpha_\ell$ is a function of ℓ. The companion paper of Ross et al [94] works in configuration space, so in their version of Template Subtraction the data vectors are in pixel-space (they also take some additional steps; see footnote 7).

We will work in harmonic space and so follow Ho et al [190], but we will keep the notation general until dealing with the two-point functions where we will explicitly work with power spectra. The treatment for configuration space is largely identical. To apply the method, one would simply substitute the correlation function for the power spectrum $C_\ell^{ij} \to w^{ij}(\theta)$ and $\alpha_\ell \to \alpha(\theta)$. See e.g. Crocce et al [217] for an application of template subtraction to the correlation function.

If we consider just a single contaminant for simplicity ($N_{\rm tpl} = 1$), and assume that it is uncorrelated with the underlying galaxy field, then from Eq. (3.5) the two point function of the observed field is

$$\langle d_{\rm obs} d_{\rm obs}\rangle = \langle ss\rangle + \alpha^2 \langle tt\rangle. \tag{3.6}$$

Then on average,

$$\langle t d_{\rm obs}\rangle / \langle tt\rangle = C_\ell^{td}/C_\ell^{tt} = \alpha_\ell \tag{3.7}$$

and the contamination at each multipole can be estimated as

$$\hat{\alpha}_\ell = \tilde{C}_\ell^{\,td}/C_\ell^{tt} \tag{3.8}$$

where the tilde in $\tilde{C}_\ell$ indicates the power spectrum that is measured from the observed sky realization, and $\tilde{C}_\ell^{tt} = C_\ell^{tt}$ since we take templates to be fixed.

An estimate of the power spectrum can then be found to be

$$\hat{C}_\ell^{ss} = \left(\tilde{C}_\ell^{dd} - \hat{\alpha}_\ell^2 C_\ell^{tt}\right)\left(1 - \frac{1}{2\ell+1}\right)^{-1}. \tag{3.9}$$

Here $[1 - 1/(2\ell+1)]^{-1} = [(2\ell+1)/(2\ell)]$ is a factor found by Elsner et al [97] that is needed to debias the estimator.[5] The bias arises because the process is too aggressive — any chance correlation between template and the true signal is also removed, resulting in an underestimate of the true clustering power.

The Template Subtraction method is easily generalized to multiple templates by extend-

[5] In the case of the correlation function, the bias cannot be written in a signal-independent fashion, and so requires a prior signal power spectrum or simulations to estimate.

ing the dimensionality of terms as

$$\alpha_\ell \text{ (scalar)} \to \boldsymbol{\alpha}_\ell \ (N_{\rm tpl})$$
$$C_\ell^{tt} \text{ (scalar)} \to \mathbf{C}_\ell^{\mathbf{TT}} \ (N_{\rm tpl} \times N_{\rm tpl})$$

and Eqs. (3.8) and (3.9) become

$$\hat{\boldsymbol{\alpha}}_\ell = [\tilde{\mathbf{C}}_\ell^{\mathbf{TT}}]^{-1}[\tilde{\mathbf{C}}_\ell^{\mathbf{Td}}], \tag{3.10}$$

$$\hat{C}_\ell^{ss} = \left(\tilde{C}_\ell^{dd} - \hat{\boldsymbol{\alpha}}^\dagger \mathbf{C}_\ell^{\mathbf{TT}} \hat{\boldsymbol{\alpha}}\right) \left(\frac{2\ell+1}{2\ell+1-N_{\rm tpl}}\right). \tag{3.11}$$

For the cut-sky equations, we refer the reader to [97].

While previous work on Template Subtraction has focused on the cleaned power spectrum, an estimate of cleaned overdensity field itself is also of interest for cosmological study, as it contains more information than just its power spectrum.

A map estimate from the Template Subtraction method can be produced as

$$\hat{s}^{TS}(\hat{\mathbf{n}}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} \hat{s}_{\ell m}^{TS} Y_{\ell m}(\hat{\mathbf{n}}), \tag{3.12}$$

where the harmonic coefficients of the map are given by

$$\hat{s}_{\ell m}^{TS} = (d_{\rm obs})_{\ell m} - \sum_{i=1}^{N_{\rm tpl}} (t_{\ell m})_i (\hat{\alpha}_\ell)_i \tag{3.13}$$

and the (biased) power spectrum of the cleaned map is equivalent to the first factor in Eq. (3.11).

3.3.3 Mode Projection

Mode Projection (also often called Mode *De*projection [13, 195, 197, 218, 219]) assumes the same contamination model as Template Subtraction, given by Eq. (3.5). The original formulation [95, 220] cleans the map-level systematics by assigning infinite variance to contaminating templates. This procedure desensitizes the power spectrum estimate to the templates and is equivalent to marginalizing over the contamination amplitude of each template [95].

In particular, it updates the map-level covariance matrix $\mathbf{C}$ as follows

$$\begin{aligned}\mathbf{C}' &= \left[\mathbf{C} + \sum_{k}^{N_{\rm tpl}} \lim_{\beta\to\infty} (\beta_k t_k t_k^\dagger)\right] \\ &= \lim_{\beta\to\infty} \left[\mathbf{C} + \beta T T^\dagger\right]\end{aligned} \tag{3.14}$$

where t_k are the individual template maps, which can represent either real spin-0 or complex spin-2 fields [13], and which can be assembled into a matrix T, with t_k as the $k^{\rm th}$ column. In previous works with Mode Projection, the maps have been represented in pixel space, but in principle the operations can also be performed in harmonic space, e.g. representing a spin-0 field by its complex harmonic coefficients. For clarity and continuity, we will assume the maps are $N_{\rm pix}$-length vectors in what follows, as opposed to their multipole transforms. There are some benefits to performing Mode Projection in harmonic space, however, which we explore in Sec. 3.4.

The main challenge with the original formulation of Mode Projection is that it requires the construction and inversion of a covariance matrix for the whole map, which is often intractable. To remedy this, Elsner et al [99] (E16) extended Mode Projection to the popular (albeit sub-optimal) pseudo-C_ℓ estimator. In practice, this is achieved by computing the pseudo-C_ℓs of the overdensity field after first applying a filter $\mathbf{F}$, where

$$\begin{aligned}\mathbf{F} &= \lim_{\beta\to\infty} \left(I + \beta T T^\dagger\right)^{-1} \\ &= I - T(T^\dagger T)^{-1} T^\dagger,\end{aligned} \tag{3.15}$$

where the second expression follows from the Sherman-Morrison-Woodbury formula. It is easy to see that $T(T^\dagger T)^{-1}T^\dagger$ is a projection matrix, projecting an $N_{\rm pix}$-dimensional map onto a $N_{\rm tpl}$-dimensional subspace. The filter thus removes any components of the observed map within the subspace spanned by the templates (hence the alternate name of Mode *De*projection).

Taking the case of a single template map t for simplicity, $\mathbf{F}$ then takes the form $(I - (tt^\dagger)/(t^\dagger t))$, resulting in a filtered overdensity map

$$\hat{s} = \mathbf{F}\, d_{\rm obs} \tag{3.16}$$

$$= \left[I - t(t^\dagger t)^{-1} t^\dagger\right] d_{\rm obs} \tag{3.17}$$

$$= d_{\rm obs} - t\hat{\alpha}_{\rm mp} \tag{3.18}$$

where

$$\hat{\alpha}_{\rm mp} = (t^\dagger d_{\rm obs})/(t^\dagger t) = \tilde{\sigma}^2_{td}/\sigma^2_{tt}, \tag{3.19}$$

and $\tilde{\sigma}^2_{td}$ is a measure of the covariance of maps t and d. Note that this is very similar to the Template Subtraction estimate in Eq. (3.8), but here the covariances are taken over the whole footprint, rather than for a single mode ℓ. We can make the connection even more explicit by noting that in the full-sky case,

$$\tilde{\sigma}^2_{td} = \frac{1}{4\pi}\sum_{\ell=0}^{\infty}(2\ell+1)\tilde{C}_\ell^{\,td} \tag{3.20}$$

While E16 introduce this filtered map only as a means to compute the power spectrum, it can be used on its own as an estimate for the cleaned overdensity field. However, as with Template Subtraction, the power spectrum of this cleaned *map* is a biased estimate of the true power spectrum, as some of the signal is removed in the cleaning process:

$$\langle C_\ell^{\hat{s}\hat{s}}\rangle = \langle (d_{\rm obs} - t\hat{\alpha}_{\rm mp})^\dagger (d_{\rm obs} - t\hat{\alpha}_{\rm mp})\rangle \tag{3.21}$$

$$= C_\ell^{ss} - \frac{C_\ell^{tt}}{4\pi\left(\sigma^2_{tt}\right)^2}\left(2C_\ell^{ss}\sigma^2_{tt} - \frac{1}{4\pi}\sum_{\ell'}(2\ell'+1)C_{\ell'}^{ss}C_{\ell'}^{tt}\right). \tag{3.22}$$

In the full sky case, the power spectrum estimate can be debiased analytically [99]:

$$\hat{C}_\ell^{ss} = \sum_{\ell'}\left[(I+B)^{-1}\right]_{\ell\ell'} C_{\ell'}^{\hat{s}\hat{s}}, \tag{3.23}$$

where

$$B_{\ell\ell'} = \frac{C_\ell^{tt}}{4\pi\left(\sigma^2_{tt}\right)^2}\left(-2\sigma^2_{tt}\delta_{\ell\ell'} + \frac{2\ell'+1}{4\pi}C_{\ell'}^{tt}\right) \tag{3.24}$$

and $\delta_{\ell\ell'}$ is the Kronecker delta. In the presence of a mask, one can debias via iteration or assuming a prior power spectrum [99]. As we work in the full-sky case, we debias analytically via Eq. (3.23), though we do not expect an iterative or prior-based debiasing to significantly alter our conclusions.

The procedure outlined above easily generalizes to multiple maps by extending the di-

mensionality of the terms:

$$\begin{aligned}
\alpha \text{ (scalar)} &\rightarrow \alpha \ (N_{\rm tpl}), \\
t \ (N_{\rm pix}) &\rightarrow \mathbf{T} \ (N_{\rm pix} \times N_{\rm tpl}), \\
\sigma^2_{td} \text{ (scalar)} &\rightarrow \sigma^{\mathbf{2}}_{\mathbf{Td}} \ (N_{\rm tpl}), \\
C^{tt}_\ell \text{ (scalar)} &\rightarrow \mathbf{C}^{\mathbf{TT}}_\ell \ (N_{\rm tpl} \times N_{\rm tpl}), \\
\sigma^2_{tt} \text{ (scalar)} &\rightarrow \sigma^{\mathbf{2}}_{\mathbf{TT}} \ (N_{\rm tpl} \times N_{\rm tpl}).
\end{aligned} \tag{3.25}$$

Hereafter, we will use 'Mode Projection' to refer to the pseudo-C_ℓ mode projection method described above, due to the popularity of the pseudo-C_ℓ power spectrum estimator and the adoption of this version into `NaMaster`[6][13], in anticipation of LSST. We again refer the reader to E16 for the modifications necessary to account for the mask, and specifically to their Eq. (21) for the multi-template version of the debiasing matrix, which we use to correct for all Mode Projection power spectrum estimates (see [13] for the equivalent formulae for spin-2 fields).

3.4 Placing into a Common Mathematical Framework

To facilitate a comparison of the methods, it is useful to place them into a common mathematical framework. In this section, we show how all three methods presented so far can be interpreted through a regression analysis lens, and in doing so help identify different assumptions within each method and possible avenues for improvement. Moreover, we can leverage the powerful suite of tools that have already been developed and tested for regression to the task of systematics removal, facilitating and accelerating the process.

3.4.1 Connections to Regression

We have purposefully formulated the methods (e.g. Eqs. (3.19 and (3.8))in a manner designed to make the connections between Mode Projection and Template Subtraction apparent. Template Subtraction is equivalent to running the Mode Projection algorithm, but with each original template ($t_i(\hat{\mathbf{n}})$) decomposed into a set of independent templates ($t^i_\ell(\hat{\mathbf{n}})$), where

$$t^i_\ell(\hat{\mathbf{n}}) = \sum_{m=-\ell}^{\ell} t^i_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}). \tag{3.26}$$

[6] https://github.com/LSSTDESC/NaMaster

Fig. 3.4 shows this schematically. In other words, (pseudo-C_ℓ) Mode Projection can be considered a special case of Template Subtraction, where the contamination is assumed to be independent of scale *and* the full template map is used to estimate such contamination. It has been pointed out before in the context of 3D clustering estimates that Template Subtraction and Mode Projection can be related if they use equivalent templates [195].

Casting the two methods into this form allows us to make the connection to standard linear regression wherein a measured response $\mathbf{y}$ is assumed to be a linear combination of predictors given by the $\boldsymbol{\alpha}$ and a noise term $\boldsymbol{\epsilon}$:

$$\mathbf{y} = \mathbf{X}\alpha + \epsilon. \tag{3.27}$$

$\mathbf{X}$ is a $n \times p$ matrix, where p is the number of predictors (potentially including a column of ones — the intercept term), $\boldsymbol{\alpha}$ a vector of length p, and $\mathbf{y}$ and $\boldsymbol{\epsilon}$ vectors of length n.

Perhaps the most common regression method, Ordinary Least Squares (OLS), finds the vector $\hat{\alpha}$ that minimizes the squared residuals:

$$\hat{\boldsymbol{\alpha}} = \mathrm{argmin}_\alpha ||\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}||^2 \tag{3.28}$$

$$= (\mathbf{X}^\dagger\mathbf{X})^{-1}\mathbf{X}^\dagger\mathbf{y}, \tag{3.29}$$

where the second expression follows if X is full column-rank (i.e. the number of observations exceeds the degrees of freedom from the predictors). This is equivalent to the maximum likelihood solution if one assumes the noise of each element, ϵ_i, is independent and identically Gaussian distributed,

$$P(\boldsymbol{y}|\boldsymbol{X}\alpha) \sim \mathcal{N}(0, I\sigma_\epsilon). \tag{3.30}$$

such that the log-likelihood goes as $\mathcal{L} \propto |\mathbf{y} - \mathbf{X}\boldsymbol{\alpha}|^2$. Even if the assumption of Gaussianity is violated, by the Gauss-Markov theorem Eq. (3.28) still corresponds to the unbiased estimator with minimum variance if the errors $\boldsymbol{\epsilon}$ are uncorrelated and have equal variance.

We can write Eq. (3.27) in terms of the OLS estimates as

$$\mathbf{y} = \mathbf{X}\hat{\alpha} + \hat{\epsilon} = \mathbf{X}(\mathbf{X}^\dagger\mathbf{X})^{-1}\mathbf{X}^\dagger\mathbf{y} + \hat{\epsilon} \tag{3.31}$$

where the residuals are defined as

$$\hat{\epsilon} = \mathbf{y} - \mathbf{X}\hat{\alpha}. \tag{3.32}$$

The quantities of interest in the typical regression problem are the coefficients $\boldsymbol{\alpha}$ or the predicted response $\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\alpha}}$, with the goal of understanding the influence of predictors or to predict future observations, and hence the residuals are largely used to assess whether the

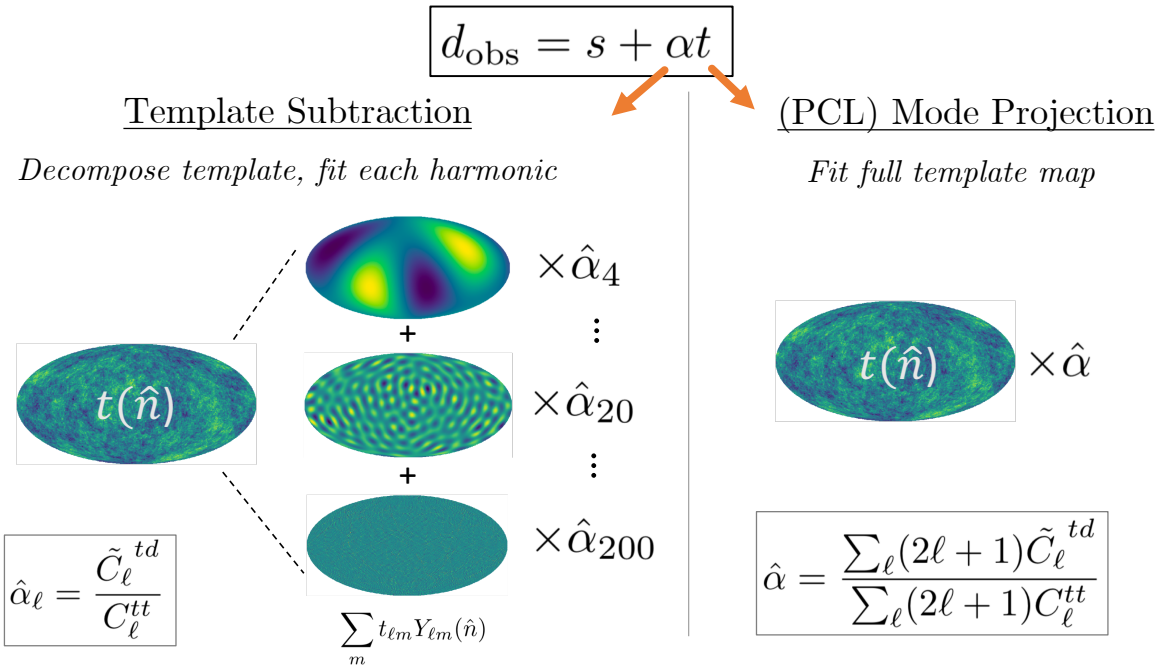

Figure 3.4: Schematic illustration of the difference between the Template Subtraction and (pseudo-C_ℓ) Mode Projection methods. Template Subtraction allows templates to have different levels of contamination at each scale. This is analogous to performing Mode Projection, but first decomposing each template map into a series of derived templates, each corresponding to a different harmonic ℓ. See Sec. 3.4.1 for details.

basic OLS assumptions hold. However comparing Eq. (3.32) to Eqs. (3.13) and (3.18), we see that **both Mode Projection and Template Subtraction can be interpreted as OLS regression methods where the observed overdensity signal is regressed onto the templates, and the reconstructed overdensity signal $\hat{s}$ and power spectrum $C_\ell^{\hat{s}\hat{s}}$ correspond to the map and power spectrum of the residuals $\hat{\boldsymbol{\epsilon}}$.**

Mode Projection uses the full map footprint, with each pixel corresponding to a single observation, for a total of $N_{\rm tpl}$ fit coefficients. In contrast, Template Subtraction can be interpreted as performing multiple OLS regressions in parallel on smaller subspaces — one at each multipole in our case — for a total of $N_\ell \times N_{\rm tpl}$ fit coefficients (see Fig. 3.4).

We can write the Template Subtraction amplitudes computed by Eq. (3.10) in OLS form as

$$\hat{\alpha}_\ell = (\mathbf{T}_\ell^\dagger \mathbf{T}_\ell)^{-1} \mathbf{T}_\ell^\dagger d_\ell, \tag{3.33}$$

where $\mathbf{T}_\ell$ is a $(2\ell+1) \times N_{\rm tpl}$ matrix, with each column corresponding to a template, consisting

of all the harmonic coefficients for a fixed ℓ:

$$\mathbf{T}_\ell = \begin{pmatrix} t^1_{\ell,-\ell} & t^2_{\ell,-\ell} & \cdots & t^{N_{\rm tpl}}_{\ell,-\ell} \\ t^1_{\ell,-\ell+1} & t^2_{\ell,-\ell+1} & \cdots & t^{N_{\rm tpl}}_{\ell,-\ell+1} \\ \vdots & \vdots & \ddots & \vdots \\ t^1_{\ell,\ell} & t^2_{\ell,\ell} & \cdots & t^{N_{\rm tpl}}_{\ell,\ell} \end{pmatrix}. \tag{3.34}$$

In cases where the multipoles (or angular scales) are binned, the number of fit coefficients is reduced to $N_{\rm bins} \times N_{\rm tpl}$, which reduces the variance of the contamination estimate. Indeed, Mode Projection corresponds to a limiting case, where the modes of each template are averaged with equal weight before fitting. However in principle one could apply weights differently across scales, and as we will show, this can produce improved coefficient estimates. Alternatively, one could fit individual modes as in Template Subtraction but combine at the coefficient level — potentially useful if certain scales are of particular interest for a given analysis.[7]

An immediate consequence of the OLS interpretation of these methods is in making explicit the assumptions that Mode Projection and Template Subtraction are making about the underlying density field — they are exactly the "OLS" assumptions for the error term ϵ in the regression model: independent, Gaussian and of equal variance, in whatever basis the map is represented. These assumptions hold well for Template Subtraction, which performs a separate regression at each multipole ℓ. In this case, the assumed OLS "noise" terms are the set of harmonic coefficients of the map ($s_{\ell m}$) at that multipole, which have $\mathrm{Cov}[s_{\ell m_1}, s_{\ell m_2}] = C^{ss}_\ell \delta_{m_1 m_2}$. For Mode Projection, these assumptions are violated, as the covariance matrix between overdensity pixels is not diagonal, $\mathrm{Cov}[s(\hat{\mathbf{n}}_i), s(\hat{\mathbf{n}}_j)] \neq \sigma^2_{\rm sig}\delta_{ij}$.

Since the primary contribution to the "noise" of the OLS fit is the clustering signal itself, we can diagonalize it by performing Mode Projection in multipole space, with the maps d, s and t_i becoming complex column vectors comprised of the map spherical harmonic coefficients. The noise of the observed overdensity $d_{\ell m}$ is then $\mathrm{Cov}[s_{\ell_1 m_1}, s_{\ell_2 m_2}] = C^{ss}_\ell \delta_{\ell_1 \ell_2} \delta_{m_1 m_2}$. While diagonal, this varies strongly with ℓ and therefore violates the assumption of equal variance, a property known as 'heteroskedasticity' in the statistics literature.

However once the noise is diagonal, we can improve the Mode Projection estimate of $\hat{\boldsymbol{\alpha}}$

[7] In their real-space analysis of SDSS galaxies, Ross et al [94] seem to implement a version of this. They use Template Subtraction to produce fit coefficients for a large number of scales and templates, but ultimately select one coefficient to apply to all scales for each template. However it is unclear how they compute the single summary coefficient.

by weighting the observed data and template modes by (a prior-inferred) $1/\sqrt{C_\ell^{ss}}$:

$$\hat{\alpha} = \frac{\sum_{\ell=0}^{\infty}(2\ell+1)\tilde{C}_\ell^{\,td}/C_\ell^{ss}}{\sum_{\ell=0}^{\infty}(2\ell+1)\tilde{C}_\ell^{\,tt}/C_\ell^{ss}}. \tag{3.35}$$

This is equivalent to a weighted least-squares approach and recovers the maximum likelihood estimate of $\hat{\boldsymbol{\alpha}}$, eschewing the erroneous assumption of a flat signal power spectrum. This of course only works in the ideal full-sky case, but in principle it should not be difficult to extend to a masked sky, e.g. using a predicted cut-sky C_ℓ^{ss} computed using the standard coupling matrix from the mask [e.g. 99, 221] along with the cut-sky harmonics of the templates and datavector, or appropriate binning of modes. This can be viewed as a form of 'prewhitening' the data, which accounts for the off-diagonal pixel covariance in the likelihood through an appropriate transform. We explore the potential improvement from such prewhitening in App. B.2, finding that it improves cleaning, but is subdominant to differences between cleaning methods and higher-order corrections we discuss below.

Finally we note that both the Template Subtraction bias from Elsner et al [97], as well as the pseudo-C_ℓ Mode Projection bias from E16 result trivially when interpreting them through the OLS lens, in which the variance of observed residuals is well-known to be biased low:

$$\langle\hat{\epsilon}^\dagger\hat{\epsilon}\rangle = \left(\frac{N_{\rm data}-p}{N_{\rm data}}\right)\epsilon^\dagger\epsilon. \tag{3.36}$$

For Template Subtraction, the regression at each harmonic has $N_{\rm data} = 2\ell+1$ and number of predictors $p = N_{\rm tpl}$, leading exactly to the debiasing terms for the signal power estimate in Eqs. (3.9) and (3.11). The debiasing terms for Mode Projection in Eq. (3.22) are more complicated and dependent on the signal and template clustering, but if we take both C_ℓ^{ss} and C_ℓ^{tt} to be independent of ℓ, Eq. (3.22) reduces to

$$\langle\hat{s}\hat{s}\rangle = C^{ss}\left(1 - \frac{1}{\sum_{\ell'=0}^{\ell_{\rm max}}(2\ell'+1)}\right), \tag{3.37}$$

where $p = N_{\rm tpl} = 1$ and $N_{\rm data} = \sum_{\ell'=0}^{\ell_{\rm max}}(2\ell'+1) = (\ell_{\rm max}+1)^2$ is the total number of Fourier modes in the map. This is in keeping with the interpretation of E16, wherein each template removes one degree of freedom from the number of observed Fourier modes. This interpretation can help to assess the risk of overfitting based on the size of the template library.

By making connections between current methods and linear regression explicit, we not only facilitate their interpretation, but can more easily identify the tacit assumptions within

these methods, as well as readily improve upon them, drawing on the large body of research into the statistical properties of various regression approaches.

3.4.2 Additive vs. Multiplicative Treatment

The systematic contamination described in Eq. 3.2 results in both additive and multiplicative contributions to the overdensity field. One way in which the methods described here differ is whether or not they ignore the multiplicative contributions. These take the form $\delta(\hat{\mathbf{n}})f_{\rm sys}(\hat{\mathbf{n}})$ and, if unaddressed, can bias cosmological constraints in upcoming surveys [102]. Here we show how so-called "additive methods" like Mode Projection (or any other regression method) and be readily adapted to account for these multiplicative terms and so lead to improved map and power spectrum estimates.

For clarity, we reformulate Eq. (3.2) in the general notation of Sec. 3.3 for direct comparison with the additive methods, taking $\delta_{\rm obs} \to d_{\rm obs}$ and $\delta \to s$, such that

$$d_{\rm obs} = \gamma(1+s)(1+f_{\rm sys}) - 1 \tag{3.38}$$

where again we have suppressed the pixel index. $\gamma = \bar{N}_{\rm true}/\langle N_{\rm obs}\rangle_{\rm pix}$ accounts for the so-called integral constraint, wherein the mean *observed* number density is used to compute the overdensity field, rather than the true full-sky mean density (see App. B.4 for a more in depth look at the impact of this monopole term).

To compare with the additive methods, it is convenient to define a zero-centered systematic as:

$$f_i' \equiv \frac{f_i - \bar{f}_i}{1+\bar{f}_i}, \tag{3.39}$$

and write Eq. (3.38) in an equivalent but zero-centered form:

$$d_{\rm obs} = \gamma'(1+s)\left(1+f'_{\rm sys}\right) - 1, \tag{3.40}$$

with the new prefactor

$$\gamma' \equiv \gamma(1+\bar{f}_{\rm sys}) = \left(1 + \langle s' f'_{\rm sys}\rangle_{\rm pix} + s_0\right)^{-1} \tag{3.41}$$

ensuring that the monopole in $d_{\rm obs}$ is zero, and having the property that $\langle\gamma'\rangle \approx 1$.[8] Here $s_0 \equiv \langle s\rangle_{\rm pix}$ characterizes the global overdensity in which the footprint resides, and $s' \equiv s - s_0$

[8]Here the approximation stems from making the assumption $\langle x^{-1}\rangle \approx \langle x\rangle^{-1}$, which holds very well for the cases we are studying where the mean is taken over a footprint with $N_{\rm pix} \gtrsim 10^5$ and shot noise is subdominant.

is the deviation from that local overdensity.

Thus the observed overdensity field contaminated with a generic systematic f_i can be equivalently written as contamination from a *zero-centered* systematic with a rescaled amplitude, f_i'.

Expanding Eq. (3.40), we have

$$d_{\rm obs} = s + \gamma' f'_{\rm sys} + \gamma' s f'_{\rm sys} + (\gamma' - 1)(s + 1), \tag{3.42}$$

Comparing to additive models like Mode Projection and Template Subtraction, which take

$$d^{\rm add}_{\rm obs} = s + f'_{\rm sys} = s + \sum_{i=1}^{N_{\rm tpl}} f'_i, \tag{3.43}$$

we see that they assume that $\gamma' = 1$ (no mean local overdensity and a vanishing correlation between signal and systematics over the footprint), as well also that $s f'_{\rm sys} = 0$ *for every pixel*, a much stronger assumption. Despite these assumptions, additive estimates of the total contamination are unbiased, provided the templates T span the space of the true contamination:

$$\langle \hat{f}_{\rm sys} \rangle = \langle T\hat{\boldsymbol{\alpha}} \rangle = \langle T(T^\dagger T)T^\dagger d_{\rm obs} \rangle \tag{3.44}$$

$$\approx T(T^\dagger T)T^\dagger f'_{\rm sys} = f'_{\rm sys}. \tag{3.45}$$

Intuitively this makes sense, since in the ensemble average the multiplicative term sf' will vanish.

From Eq. (3.40), we can then make an improved estimate of the signal map as

$$\hat{s} = \frac{1 + d_{\rm obs}}{1 + \hat{f}_{\rm sys}} - 1, \tag{3.46}$$

$$= \frac{d_{\rm obs} - \hat{f}_{\rm sys}}{1 + \hat{f}_{\rm sys}} \tag{3.47}$$

where the second form makes clear that this is a simple rescaling of the additive signal estimate, $d_{\rm obs} - \hat{f}_{\rm sys}$. Therefore, *in a model with multiplicative contamination, signal estimates from additive methods can be improved by weighting the estimated signal map by* $1/(1 + \hat{f}_{\rm sys})$. Such reweighting should be avoided for contaminants that are thought to contribute additively to the number density (such as stellar contamination), as these modify γ and result in an additive contribution to the overdensity, but no direction-dependent multiplicative terms (see e.g. [217, 219]).

To explicitly close the loop on the aforementioned methods, the DES-Y1 method performs a series of 1-D regressions and iteratively weights the observed overdensity in a manner equivalent to Eq. (3.46) for each template, whereas Mode Projection estimates contamination via a single $N_{\rm tpl}$-dimensional regression, with a signal estimate that can be improved via Eq. (3.47).[9] Applying the multiplicative correction makes Mode Projection equivalent to the Weights model where the coefficients are derived from a simultaneous multiple regression on all the templates (such as in [196, 203]), but with an additional correction to debias the inferred two-point function. Thus a pixelized weights map for Mode Projection can be produced[10] as

$$w(\hat{\mathbf{n}}) = (\hat{s}(\hat{\mathbf{n}}) + 1)/(d_{\rm obs}(\hat{\mathbf{n}}) + 1), \tag{3.49}$$

Fig. 3.5 illustrates the effect of the multiplicative terms — as well as the impact of neglecting them — on the residuals of a map with a single linear, multiplicative contaminant. The diagonal, dotted line shows the expected relation that would be precisely followed by a purely additive contaminant. A multiplicative contamination adds significant scatter around this relation, shown as the gray points. This scatter remains when the contamination is cleaned with an additive method (orange), but is effectively removed when the multiplicative component is taken into account (blue). Fig. 3.6 shows how errors on the estimated overdensity field are drastically reduced when applying the multiplicative correction of Eq. (3.47) to a realistic use case with multiple contaminating systematics (see Sec. 3.6 for details of implementation).

[9] As noted in Sec. 3.2 even linear contaminants will have interaction terms up to order $N_{\rm tpl}$, such that in principle, for Eq. (3.44) to fully capture $f_{\rm sys}$, additional templates up to $t_i t_j t_k ... t_{N_{\rm tpl}}$ would need to be included in the template library. A more precise and efficient approach would be to not add any interaction templates, but instead combine the base systematic estimates as

$$\hat{f}_{\rm sys,alt} = \prod_{i=1}^{N_{\rm tpl}} (1 + \hat{f}_i) = \prod_{i=1}^{N_{\rm tpl}} (1 + \hat{\alpha}_i t_i), \tag{3.48}$$

where recall t_i corresponds to the $i^{\rm th}$ template and $i^{\rm th}$ column of T, and $\hat{\alpha}_i$ the $i^{\rm th}$ element of $\hat{\alpha}$. This is closer to the treatment of the DES-Y1 method, wherein weightings for each $\hat{f}_i$ are applied in series and thus cumulatively.

In practice we find that Eq. (3.44) is a very good approximation since $\sigma^2_{\rm sys} \lesssim \mathcal{O}(10^{-2})$, so the nonlinear interaction contributions to $f_{\rm sys}$ *due to each systematic acting as its own multiplicative screen* are fairly negligible, i.e. $f'_{\rm sys} \approx \left(\sum_{i=1}^{N_{\rm tpl}} f'_i\right)$, as long as the templates sufficiently capture the form of contamination: $f'_i = \alpha_i t_i$. Of course this latter condition is a basic requirement of all of the methods we describe here, one that can and should be verified through standard residual plots and other regression diagnostic techniques to ensure an appropriate contamination model for each template. Methods that incorporate template selection criteria, such as the proposed Elastic Net, can help to satisfy this by allowing a large number of templates to be included in order to address potential higher order terms with little penalty.

[10] This can be released on its own or, as with the DES-Y1 data release, as an additional column at the catalogue level. c.f. https://des.ncsa.illinois.edu/releases/y1a1/key-catalogs/key-redmagic).

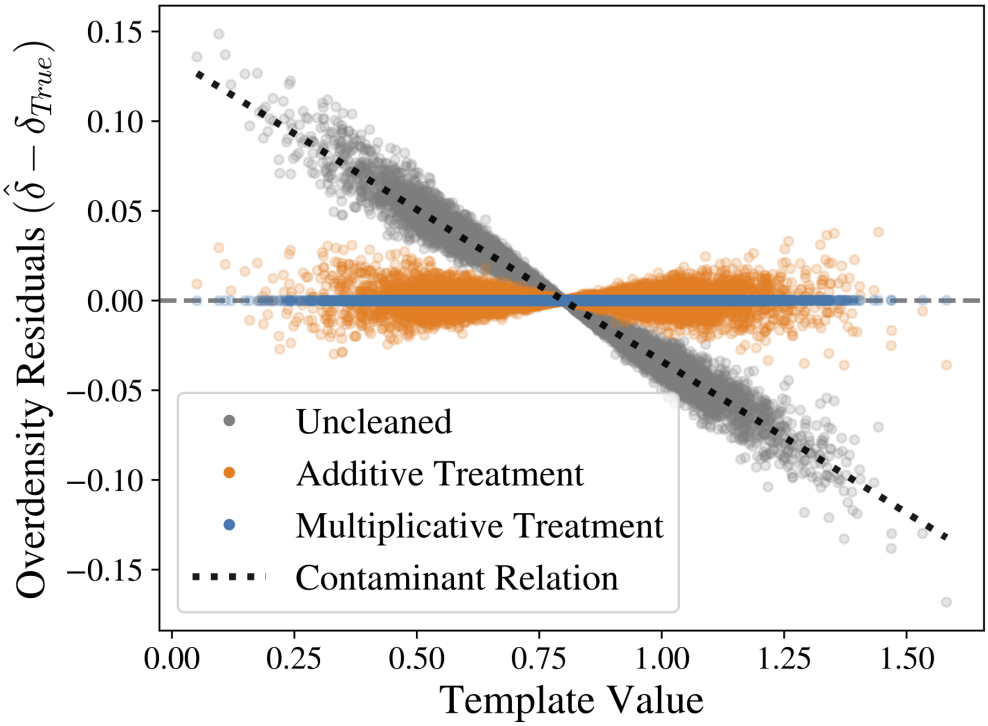

Figure 3.5: The error in estimates of the overdensity δ in a toy Gaussian map when contaminated with a single template. Gray points indicate the pixel-based difference between the observed, uncleaned overdensity and the true overdensity when the contamination is multiplicative (additive contamination would lie directly along the dotted line). Orange points are the result when erroneously assuming the contamination is only additive. Blue points are the result when correctly treating the multiplicative component.

3.4.3 Multiplicative Effect on Likelihood

While the multiplicative term vanishes in the ensemble average, resulting in the same ensemble pixel mean as the additive-only prediction ($\langle d_{\rm obs}\rangle = f'_{\rm sys}$), the pixel variance is modulated:

$$\begin{aligned}\mathrm{Var}[d_{\mathrm{obs}i}] &\approx \langle [s\gamma'(1+f'_{\rm sys})]^2\rangle \\ &\approx \langle s^2\rangle (1+f'_{\rm sys})^2 \\ &= \sigma^2_{\rm sig}(1+f'_{\rm sys})^2,\end{aligned} \tag{3.50}$$

where for large $N_{\rm pix} \gtrsim 10^5$, $\langle\gamma'\rangle \approx \langle\gamma'^2\rangle \approx 1$. The corresponding covariance between pixels is

$$\mathrm{Cov}(d_{\rm obs}, d_{{\rm obs}j}) \approx \left\langle (\gamma')^2 \left[s_i(1+f_{{\rm sys}_i}')\right]\left[s_j(1+f_{{\rm sys}_j}')\right]\right\rangle . \tag{3.51}$$

This is the source of the systematic-dependent scatter in Fig. 3.5, which will result in biased two-point statistics from additive methods. Because the contamination estimate is unbiased, the correction of Eq. (3.47) almost fully suppresses this variance, but the multiplicative terms

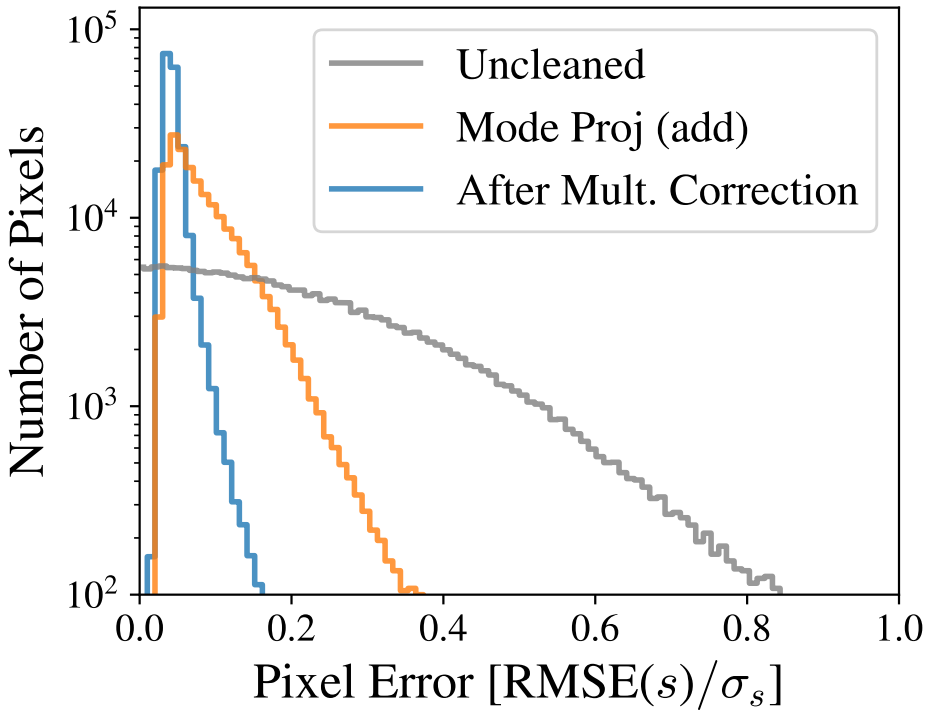

Figure 3.6: Distribution of pixel errors before cleaning (gray), after cleaning with Mode Projection but *before* multiplicative correction (orange), and *after* multiplicative correction (blue). The errors have been calculated as the RMSE of each pixel across 100 cleaned mocks in our fiducial configuration of a DES-like survey as described in Sec. 3.6, and have been normalized to the expected dispersion from the true overdensity field.

also impact the likelihood when performing the regression. In pixel-space a simple fix would be to iterate: use an initial estimate of $\langle \hat{f}_{\rm sys} \rangle$ with Eq. (3.50) to apply inverse variance weights to the maps before making a second estimate of $\langle \hat{f}_{\rm sys} \rangle$. In practice these are "errors on the errors" and so the impacts will be subdominant to the multiplicative correction to the datavector itself.

3.5 Applications

We can use the insights of the previous sections to propose two additional methods, as well as to estimate the errors on the cleaned map. We now describe these in turn.

3.5.1 Iterative Forward Selection

We include an iterative Forward Selection method that incorporates some of the main features of the DES-Y1 method, but adopts some of the simplifying assumptions of Mode Projection. The result is greatly simplified and easier to implement than the full DES-Y1 method.

We keep the core of the template selection algorithm, but modify the fit procedure and significance criterion to eliminate the need to generate mocks. We do this by adopting the same implicit assumptions of Mode Projection: that pixels are uncorrelated and have equal variance. This allows for an analytical solution for the best-fit parameters $\theta = \{m_i, b_i\}$ and their covariance Cov_θ for each template, which we obtain using `numpy.polyfit`[11]. We then adopt a simplified significance criterion of $\Delta\chi^2_{\rm FS} = \theta^T[\mathrm{Cov}_\theta]^{-1}\theta$, and use the same stopping threshold as the DES-Y1 method.[12]

This Iterative Forward Selection method is a fast and simple method that incorporates some of the key aspects of the DES-Y1 method, the iterative weighting and template selection, while avoiding the most computationally expensive parts, the generation of mocks. We expect some loss of precision by not including a covariance matrix in the fitting step, but on the other hand to gain some precision by not having to bin pixels, so this method can help to benchmark the importance of including the covariance matrix in a DES-Y1-like method.

3.5.2 Elastic Net

We also propose a method that closely mimics Mode Projection but incorporates template selection, thereby reducing the impact of overfitting when the template library is large.[13] Having shown that Mode Projection is equivalent to linear regression, we adopt a regression method specifically designed to automatically select predictors based on the data.

This selection is accomplished by modifying the *Loss* function that is optimized when fitting, which is equivalent to applying a prior to the template coefficients and finding their maximum a-posteriori (MAP) estimate. Specifically, instead of finding $\hat{\alpha}$ that minimizes the

[11] To estimate Cov_θ, `numpy.polyfit` assumes a diagonal Gaussian covariance of the pixels, scaled so that the best-fit model has reduced χ^2 of $\chi^2_{\rm red} = \chi^2/(N_{\rm pix} - 2) = 1$

[12] As with the DES-Y1 method, this method can suffer from a lack of convergence when the theshold is low, where chance correlations between the signal realization and templates result in a loop of the same series of templates being repeatedly reweighted. We adopted a limit of $10 \times N_{\rm tpl}$ reweightings for each signal realization before breaking the loop and using the resulting signal estimate as is. This occurred occasionally and at very low thresholds, with no discernible effect on the estimated maps or power spectra.

[13] See [96] for an alternative approach that pre-selects templates for projection using a χ^2 threshold.

square of the residuals ($||d_{\rm obs} - T\alpha||^2$), we instead minimize

$$\text{Loss} = \frac{1}{2N_{\rm pix}}||d_{\rm obs} - T\alpha||_2^2 + \lambda_1||\alpha||_1 + \frac{\lambda_2}{2}||\alpha||_2^2, \tag{3.52}$$

where

$$||\alpha||_1 = \sum_i^{N_{\rm tpl}} |\alpha_i| \tag{3.53}$$

is the L1-norm of α, and

$$||\alpha||_2 = \left(\sum_i^{N_{\rm tpl}} |\alpha_i^\dagger \alpha_i|\right)^{\frac{1}{2}} \tag{3.54}$$

is the usual vector L2-norm of α. Here λ_1 and λ_2 are hyperparameters that are tuned from the data, which we now discuss in turn:

1. The L1-norm term incentivizes sparsity in α by penalizing non-zero coefficients of templates, thus naturally performing template *selection*. This is useful because the number of templates in modern surveys can be enormous — e.g. [96] produce ~ 3700 templates for their analaysis of SDSS quasars — and so it is common to pre-select only a handful to use, for fear of removing true signal. Since we don't know *a priori* which templates are contaminating, the incorporation of an automated selection scheme enables a more agnostic, data-centric approach to cleaning a large library of templates, while mitigating the risk of overfitting. The use of this penalty term in isolation (i.e. setting $\lambda_2 = 0$) is often called the Least Absolute Shrinkage and Selection Operator (LASSO)[222], and has a Bayesian interpretation of applying a zero-centered Laplace prior on the elements of $\boldsymbol{\alpha}$, with a width $\propto 1/\lambda_1$ (see e.g. [223] for a discussion). L1 priors to induce sparsity have been used in a variety of astrophysical problems, such as for source separation in cosmic microwave background analyses [224–226] or in reconstructing mass maps from weak lensing data [227–229].

2. The L2-norm term helps address collinearity (i.e. correlation) between template maps which, when present, can cause the matrix $T^\dagger T$ to be ill-conditioned and the variance of contamination estimates to be large. When it is the only additional penalty term (i.e. $\lambda_1 = 0$), this is often called Ridge Regression, or Tikhonov Regularization. It is straightforward to show that, from a Baysian perspective, this method is equivalent to placing a zero-centered Gaussian prior on the elements of α, with a width $\propto 1/\lambda_2$.

Since each penalty term addresses a different issue with standard regression, it is not uncommon to combine them, as proposed by Zou and Hastie [230], in a method known as

the "Elastic Net". We use the `scikit-learn` [14] implementation, `ElasticNetCV`, with a hyperparameter space of $\lambda_1/(\lambda_1+\lambda_2) \in \{0.1, 0.5, 0.9\}$ and 100 values of $(\lambda_1+\lambda_2)$ spanning three orders of magnitude, which are automatically determined from the input data (the default setting). We use 5-fold cross-validation to determine the best λ_1 and λ_2, trained on a random selection of 30% of the input map pixels.[14]

In this 5-fold cross-validation scheme, the training sample (30% of the map) is itself partitioned into five equal subsamples. For each combination of hyperparameters, one subsample is withheld for validation, while the other four are used to train the model by minimizing Eq. (3.52). The mean squared error (MSE) of the validation sample is then computed and stored (i.e. the first term in Eq. (3.52)). One of the four training subsamples is then withheld as the new validation set, and the process is repeated until each of the five subsamples has been used exactly once for validation, with their average MSE used to characterize the goodness-of-fit for the given set of hyperparameters λ_1 and λ_2.

Setting $\lambda_1 = \lambda_2 = 0$ reduces to OLS regression and hence to the pseudo-C_ℓ Mode Projection method, while sampling extreme values for the relative weight of the L1 vs. L2 penalty allows for the effective use of only one of the penalty terms, if preferred by the data. The use of cross-validation on a subset of the map allows the data to dictate which model is most appropriate, with minimal risk of overfitting. We illustrate the utility of this in Fig. 3.7, which shows how the cross-validation scheme naturally increases the L1 penalty when fitting for more (*un*contaminating) templates. We found that the L2 penalty became increasingly important when the correlation between templates increased beyond $\rho_{\rm tpl} \gtrsim 0.9$.

3.5.3 Map Errors

We can use the regression framework to gain insight into how errors in the estimated overdensity map are distributed across pixels. This aids the propagation of map errors in cross-correlation studies and summary statistics beyond the two-point functions, as well as helps to identify regions that may benefit from masking.

For simplicity, we assume additive contamination and correction and ignore higher-order terms:

$$d_{\rm add} = s + f_{\rm sys} = s + T\alpha \tag{3.55}$$

[14] We performed the cross-validation procedure on a subset rather than the full footprint as further protection against overfitting, but this is likely overly cautious and subsequent tests showed little difference in performance between training on 30% as compared to the full footprint.

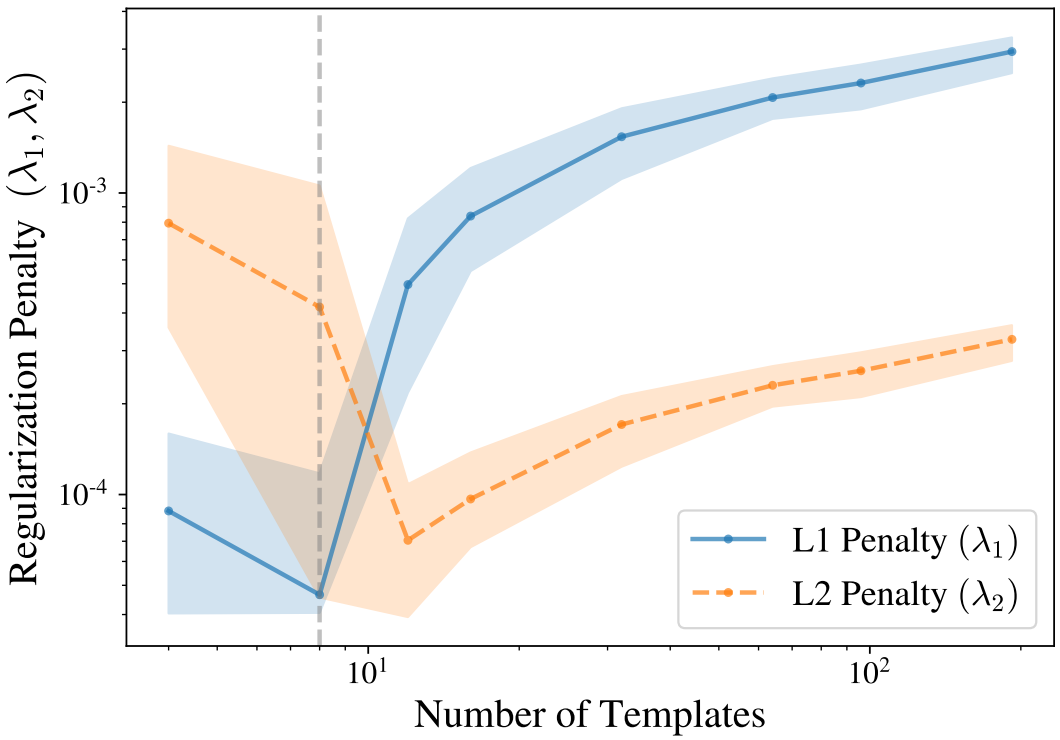

Figure 3.7: Best-fit L1 and L2 penalty coefficients in the regularization technique described in Sec. 3.5.2, as a function of the number of templates used for cleaning, $N_{\rm tpl}$ (new signal and template maps are generated at each value of $N_{\rm tpl}$). In all cases, 12 templates are contaminating the observed data (vertical dashed line). The importance of the L1 penalty, facilitating template selection, becomes increasingly important as more templates are included for cleaning. Lines and shaded region indicate the median and central 68% probability mass of 50 mocks at each $N_{\rm tpl}$ for the central bin of our fiducial DES-like survey. Here, $\rho_{\rm tpl} = 0.2$ within template groups, though plots for other $\rho_{\rm tpl}$ look similar. See Sec. 3.6 for details of implementation.

The estimated contamination amplitude is then

$$\hat{\alpha}_{\rm mp} = (T^\dagger T)^{-1} T^\dagger d_{\rm add} \tag{3.56}$$
$$= \alpha + (T^\dagger T)^{-1} T^\dagger s \tag{3.57}$$

such that our signal estimate is

$$\hat{s}_{mp} = d_{\rm add} - T\hat{\alpha}_{\rm mp} \tag{3.58}$$
$$= s - T(T^\dagger T)^{-1} T^\dagger s \tag{3.59}$$
$$\equiv (I - H)s \tag{3.60}$$

where the matrix $H \equiv T(T^\dagger T)^{-1} T^\dagger$ is often called the 'Hat' or 'Projection' matrix in the

statistics literature. Then

$$\mathrm{Var}[(\hat{s}_{\mathrm{mp}} - s)_i] = \mathrm{Var}[(Hs)_i] = [H\mathrm{Var}[s]H^\dagger]_{ii} \tag{3.61}$$

If we make the assumption that the signal covariance is diagonal, then $\mathrm{Var}[s] \approx \sigma^2_{\mathrm{sig}} I$ and

$$\mathrm{Var}[(\hat{s}_{\mathrm{mp}} - s)_i] \approx \sigma^2_{\mathrm{sig}}(HH^\dagger)_{ii} = \sigma^2_{\mathrm{sig}} H_{ii} \tag{3.62}$$

where we have used the fact that H is both Hermitian and idempotent so that $HH^\dagger = HH = H$.

Despite a number of simplifying assumptions and the fact that some of the methods only fit for some of the templates, we find that with the exception of Template Subtraction, H_{ii} is a remarkably good predictor[15] of how the errors in the overdensity estimates are distributed for all the methods. The errors arise from removing real signal during the cleaning process, with H_{ii} as a measure of how susceptible pixel i is to such overcorrection. This also indicates that while to first order all correlation with templates is removed from the estimated overdensity field, the templates remain imprinted on the map through their absence; there is missing signal in precisely their spatial configuration.

Intuitively, H_{ii} as a distance measure of pixel i from the center of mass of other pixels in the N_{tpl}-dimensional space spanned by the templates. This is sometimes referred to as 'leverage', as pixels with higher H_{ii} have larger impact when performing a regression.[16] This can be seen by observing that the estimated systematic field can be written as

$$\hat{f}_{\mathrm{sys}} = Hd_{\mathrm{obs}} \tag{3.63}$$

such that the leverage

$$H_{ii} = \frac{\partial \hat{f}^{(i)}_{\mathrm{sys}}}{\partial d^{(i)}_{\mathrm{obs}}} \tag{3.64}$$

encodes the sensitivity of the contamination estimate to an observed over- or underdensity at pixel i. Because pixels with high leverage can have an outsized effect on the estimated contamination, we expect leverage to be a useful tool for identifying potentially problematic pixels that should be masked before cleaning, in addition to providing error estimates for those pixels that remain.

[15] Note that H_{ii} only requires the diagonal elements of H, which are far more tractable to calculate than the full $N_{\mathrm{pix}} \times N_{\mathrm{pix}}$ matrix.

[16] This phenomenon is very familiar from the simple case of fitting a 1D line to a scatter of 2D points $\{x, y\}$, where the best-fit line is 'pulled' preferentially to points that lie farther from $\bar{x}$.

It is straightforward to derive the mean leverage value as

$$\bar{H}_{ii} \leq N_{\rm tpl}/N_{\rm pix}, \tag{3.65}$$

with the equality holding if T is full rank, since

$$\sum_{i}^{N_{\rm pix}} H_{ii} \;=\; {\rm Tr}(H) \leq N_{\rm tpl}, \tag{3.66}$$

providing a basis on which to determine extreme leverage values.

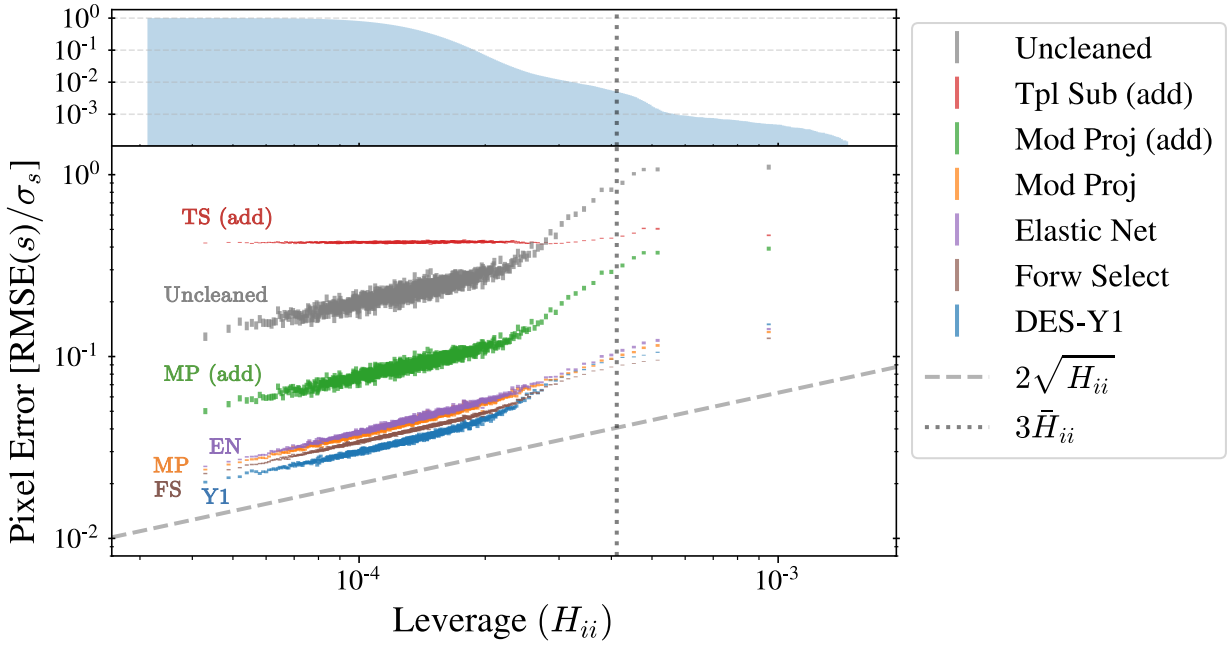

Figure 3.8: Root-mean-square error of pixel overdensity estimates, normalized to expected dispersion from the true overdensity due to cosmic variance, vs. pixel leverage for 100 signal realizations. The vertical axis shows the standard error across pixels in 1000 equal-sized bins (in this case 197 pixels per bin at `Nside`=128). The error in both observed and estimated overdensity scales as roughly $\propto H_{ii}^{1/2}$ for all methods (dashed line, to guide the eye). The dotted vertical line indicates a commonly used threshold of $3\times$ the mean leverage across pixels to identify pixels that may have an undue impact on regression fit parameters. The histogram in the top panel indicates the number of pixels at a given leverage. $N_{\rm tpl} = 27$, $N_{\rm sys} = 11$, $\sigma^2_{\rm sys} = 0.01$.

The main panel of Fig. 3.8 shows the RMS error (RMSE $= \sqrt{\langle(\hat{s}-s)^2\rangle}$) of each pixel computed over 100 cleaned DES-like mock maps plotted against leverage H_{ii} from 27 cleaning templates, 11 of which are contaminating. Pixels are grouped into 1000 bins of 197 pixels, according to their leverage value, and we show the mean and standard error of the RMSE for each bin. We see that pixels with low leverage value have much smaller error in the

estimated overdensity map, and that the error goes roughly as $\propto H_{ii}^{1/2}$ (diagonal dashed line), as predicted by Eq. (3.62). The Template Subtraction method is an exception to this trend likely because the regression happens in a different space, at each harmonic separately, and so does not relate cleanly to the pixel leverage.[17]

The top panel of Fig. 3.8 shows the fraction of map pixels below a given leverage (note the log scale), with the vertical dotted line indicating $3 \times \bar{H}_{ii}$, which is one of two common thresholds used in statistics to flag points that may bias a regression analysis ($2 \times \bar{H}_{ii}$ being the other). Here, 0.5% of map pixels exceed $3 \times \bar{H}_{ii}$; these pixels potentially merit further inspection or masking, as they are particularly prone to biasing the regression. The trend of the uncleaned data may be surprising, but as noted in Sec. 3.4.2, because of the integral constraint, $d_{\rm obs}$ is insensitive to a monopole in $f_{\rm sys}$ and so as long as templates approximately trace the true contamination, overdensities near the mean of the templates (i.e. low H_{ii}) will be most accurately measured, even if contamination is greater than at other points in the map (see App. B.4).

A complementary statistic is the 'Cook's distance'[231, 232] for each pixel, which uses H_{ii} and $\hat{s}_i$ to provide a measure of the total change in the $\hat{s}$ map if pixel i were to be masked (assuming additive contamination and correction). Along with the leverage, we expect this to be a useful tool when performing template-based mitigation of spatial systematics and for mask creation. We leave further investigation of these as diagnostic tools, as well as generalization to the multiplicative case, to a later work.

We next describe the fiducial survey on which we test the performance of foreground-cleaning methods.

3.6 Evaluating Performance

Our analysis is fully synthetic, with the procedure depicted in Fig. 3.2. We compare the cleaning methods described, including results for both the standard additive Mode Projection case (denoted 'Mode Projection (add.)') as well as one with the multiplicative correction from Eq. (3.47) (denoted simply 'Mode Projection'). For the Elastic Net, we only show results that include the multiplicative correction.

We only consider full-sky maps in this study. Extension to partial-sky surveys should be fairly straightforward, requiring the usual correction of cut-sky power spectra, but this applies equally across the full-sky spectra estimated with each method here and so we do not expect it to qualitatively change the main results.

[17] In principle, one could construct the analogous leverage quantity $H_{\ell m} = t_{\ell m}[\mathbf{C}_\ell^{\mathbf{TT}}]^{-1} t_{\ell m}^\dagger$ in harmonic space for the analysis of errors in $\hat{s}_{\ell m}$, which may be useful for cross correlation analyses in harmonic space.

3.6.1 Templates

We first describe the fiducial set of templates that we use, for both contamination and cleaning purposes. We adopt several classes of templates in order to span a range of possible contaminants and their spectral behavior. In most cases, we use multiple templates of the same class by generating Gaussian realizations of maps from the same theoretical power spectrum. The classes of template we use are

- $C_\ell \propto (\ell+1)^0$ (white noise)
- $C_\ell \propto (\ell+1)^{-1}$
- $C_\ell \propto (\ell+1)^{-2}$
- $C_\ell \propto \exp\left[-(\ell/10)^2\right]$
- a "Cat-scratch" map, with 128 horizontal stripes to model a basic scanning pattern and/or differences in depth due to overlapping tiles
- a 2D Gaussian "spot" map
- a E(B-V) extinction map, with dependence on latitude removed.

The last three correspond to static maps which do not change throughout the analysis. We use the full-sky E(B-V) map[18] from Planck [233], but since this is dominated by emission near the galactic plane, which LSS surveys typically avoid, we reweight the map to remove its major latitudinal dependence.

We normalize the individual templates to the same overall variance, and construct a total systematic map as a product of some or all of the individual template maps:

$$1 + f_{\rm sys} = \prod_{i=1}^{N_{\rm sys}} (1 + \alpha_i t_i) \tag{3.67}$$

Note that this model can generally encompass contamination to any polynomial order simply by including templates that are products of others (e.g. $t_{\rm new} \equiv t_i^2$), and incrementing $N_{\rm sys}$ accordingly. Similarly, nonlinear contamination can often be made linear through an appropriate transformation of the template map.[19] This total systematic map is then scaled to a

[18] https://wiki.cosmos.esa.int/planckpla/index.php/CMB_and_astrophysical_component_maps#The_.5Bmath.5DE.28B-V.29.5B.2Fmath.5D_map_for_extra-galactic_studies

[19] E.g. EP17 fit linear models to the square root of exposure time and sky brightness, based on how how they contribute to the depth map.

desired overall map variance $\sigma^2_{\rm sys}$, thus determining the overall contamination field $f_{\rm sys}$. We use a fiducial level of contamination of $\sigma^2_{\rm sys} = 0.01$, as we found this to produce fluctuations similar to those seen in the DES-Y1 data [212]; this corresponds to an RMS error on δ of $\sim 10\%$. Changing the level of contamination $\sigma^2_{\rm sys}$ did not significantly alter our results.

We perform the contamination and cleaning procedure shown in Fig. 3.2 on each redshift bin and for each cleaning method over many sky realizations, and plot the mean and central 68% probability mass of the relevant quality statistic. We use the same set of templates and total systematic map for across redshift bins and sky realizations, but generate a new set for each unique combination of parameter choices (e.g. level of cross-correlation between templates, number of templates used, etc.) in order to minimize any effects from specific template realizations.

We use `CLASS` [234] to compute theoretical galaxy clustering power spectra for a mock LSS survey, including contributions from redshift-space and Doppler distortions and lensing. We found gravitational potential terms to contribute $\lesssim 1\%$ to the resultant C_ℓ for $\ell > 7$ but increased computation time by an order of magnitude, so we neglect them. Since we find the cleaning procedures are not strongly sensitive to the signal power spectrum, this should not impact our results. We then use `Healpy`[11] to generate full-sky Gaussian realizations of large-scale structure overdensity ($\delta \equiv \delta\rho/\rho$) maps for each redshift bin with `NSIDE` $= 128$. We compare the impact of using lognormal maps in App. B.1, finding it does not change our results.

3.6.2 Cosmological Model and Simulated Survey

We assume a standard ΛCDM cosmological model with one species of massive neutrino and parameter values from best-fit Planck 2018:

$$\{\Omega_c, \Omega_b, h, n_s, \sigma_8, \tau, m_\nu/\mathrm{eV}\} = \{0.26499, 0.04938, 0.6732, 0.96605, 0.8120, 0.0543, 0.06\}.$$

Given the precise parameter constraints from current probes, the dependence of our results on cosmological parameters is expected to be very minimal. In contrast, the choice of the parameter set to be *determined* from the survey may be highly dependent on the residual systematics.

In general for comparing the methods, the exact form of the galaxy power spectra is not very consequential, so we use a fiducial survey comparable to the completed Y5 Dark Energy Survey, for which a realistic level of contamination can be estimated based on existing data. We assume the number density distribution of galaxies to be in the form

$$\frac{dn}{dz} \propto \left(\frac{z}{z_0}\right)^\alpha \exp\left[-(z/z_0)^\beta\right], \tag{3.68}$$

where $z_0 = 0.55$, $\alpha = 2.65$, and $\beta = 3.34$. We assume five redshift bins centered at redshifts $\{0.225, 0.375, 0.525, 0.675, 0.825\}$, with galaxy bias of $\{1.4, 1.6, 1.6, 1.95, 2\}$, respectively, and containing galaxies with Gaussian redshift dispersion of $\sigma_z = 0.05$. These values were chosen to closely approximate the `REDMAGIC` redshift distribution given in EP17.

We choose to work primarily in harmonic space. Therefore, starting with some map with overdensity $\delta \equiv \delta N/N$, where N is the galaxy count over some patch, the expansion in spherical harmonics gives

$$\delta(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{3.69}$$

and the angular power spectrum is given by

$$C_\ell = \sum_{m=-\ell}^{\ell} \frac{|a_{\ell m}|^2}{2\ell + 1}. \tag{3.70}$$

Because we are working in the full-sky limit, all well-known estimators of power return the same result, so here we make use of the `anafast` and `alm2cl` functions in `Healpy`. To more accurately account for the cosmological impact of the cleaning methods on data from a DES Y5-like survey, we divide the assumed sample variance $\sigma^2_{C_\ell}$ by a factor of $f_{\rm sky} = 0.116$.

We add shot noise to the theoretical power spectrum as $C_\ell \to C_\ell + \bar{n}^{-1}$, with $\bar{n} = 1.5 \times 10^8$, but this is negligible at the large scales we work with ($\ell \leq 350$). We are primarily interested in studying the systematic impacts of cleaning (or not) using spatial templates, so it is reasonable to focus on cases where the signal-to-noise is large (i.e. shot noise is negligible).[20]

3.7 Simulation Results

To compare methods, we compare the fidelity of the cleaned data products to the truth, either at the map level or at the level of the power spectrum, rather than look for cosmological-parameter biases. We do this for a few reasons: (1) the map and power spectrum are more general, being independent of (but easily mapped to) any specific cosmological model one wants to test, or summary statistic one wants to use; (2) while we primarily study applications to galaxy clustering data here, the methods themselves are quite general and can easily be applied to other data sets for which one has tracers of potential contamination,

[20] Shot noise may have the effect of (1) rendering the the regression residuals more diagonal in pixel-space (or flattening them in harmonic space), which could actually improve the regression procedure, and/or (2) introduce significant skewness in the distribution. We would expect the impacts of these to be similar to those of prewhitening the data or using lognormal mocks, and so based on our results in Apps. B.1 and B.2, we do not expect shot noise to significantly impact on our findings.

such as shear or convergence maps; (3) galaxy clustering *alone* leads to relatively weak cosmological constraints and is rarely used on its own to constrain cosmology.

We therefore limit ourselves to investigating biases in data space and leave the investigation of impacts on cosmological constraints to a later work when weak lensing data can be incorporated in a more realistic fashion. At this stage, the test bed is sufficiently representative to compare foreground-cleaning methods in a manner to inform future LSS analyses.

3.7.1 Characterizing Performance

We first study the impacts of the different methods on the estimated maps and power spectra for a single configuration and compare the residual biases of each. For this fiducial comparison, we generate 50 mocks for each redshift bin and contaminate them with 11 systematics, two from each of the four Gaussian classes, plus the three static templates. We construct a template library that contains the contaminating templates, plus four additional realizations from each Gaussian class, for a total of 27 cleaning templates. Each method uses this library to produce estimates of the overdensity field and power spectra.

We show map residuals of each cleaning method for the lowest redshift bin in Fig. 3.9, where the residuals are binned into deciles of the true overdensity. Results for other redshift bins are similar. From left to right, in approximate order of performance, the figure shows the Template Subtraction method (red), Mode Projection without (green) and with (orange) multiplicative correction, the Elastic Net method (purple), Forward Selection (brown) and the DES-Y1 method (blue).

The overcorrection of Template Subtraction is evident, with density fluctuations consistently *under*-estimated (i.e. peaks and voids are both less extreme than they should be). The other methods are all very close to unbiased with respect to the true overdensity field, with bias of the mean $\lesssim 0.001$ for each bin. The multiplicative methods show significantly reduced within-bin scatter (i.e. smaller error bars) compared to the additive ones — the additive Template Subtraction and Mode Projection methods (leftmost, red and green) have typical errors in the overdensity of $\sigma_s \sim 0.1$ and $\sigma_s \sim 0.01 - 0.05$, respectively, compared to the errors of $\sigma_s \sim 0.005 - 0.02$ for the multiplicative methods. This suggests that applying the multiplicative correction results in significantly improved map estimates, making them excellent candidates for map-based analyses, such as as counts-in-cells or density-split statistics.

While the signal estimates are unbiased (with the exception of Template subtraction), the errors of the additive methods increase near extremes of the density field. This is similar

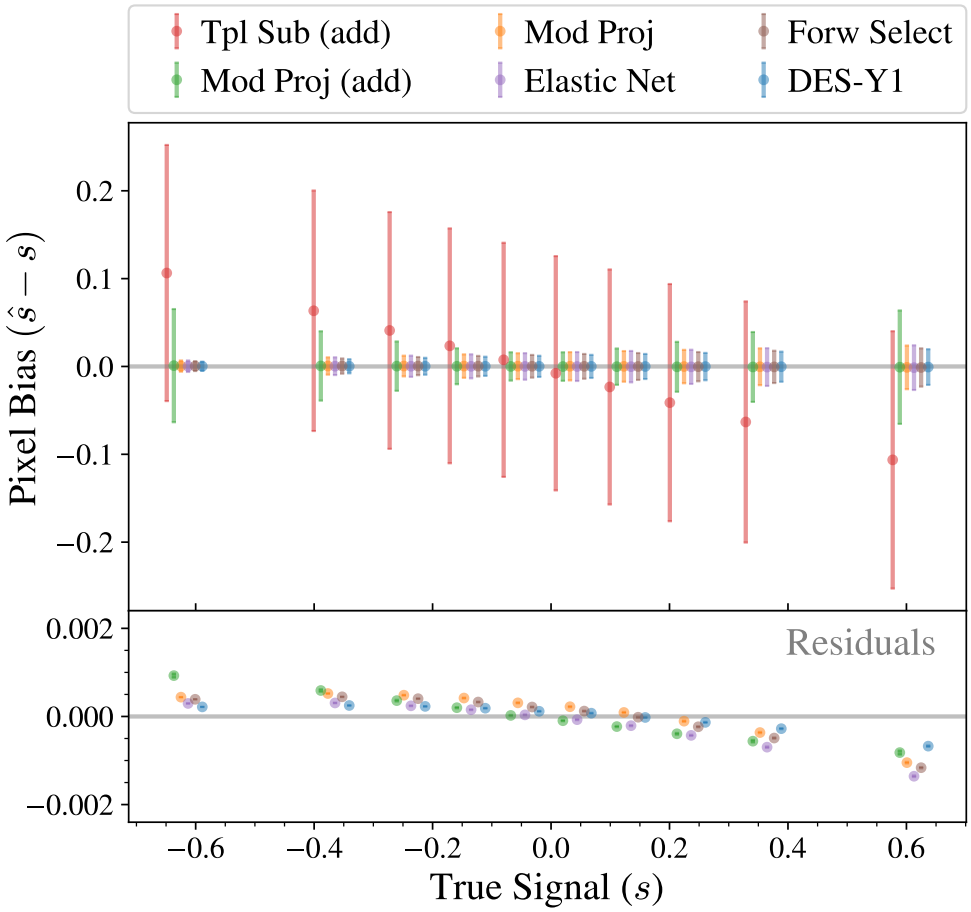

Figure 3.9: Error in overdensity estimates for different cleaning methods, binned in deciles of the true overdensity and with points offset for clarity. The top plot includes error bars indicating the standard deviation of pixel errors in each bin, while the bottom plot is a zoomed-in version to better display how the means deviate from zero. Overcorrection at the map-level is only significant for Template Subtraction, which under-estimates the magnitude of both peaks and voids, while other methods are very close to unbiased. See text for details.

to the result in Fig. 3.6, which showed larger errors at extreme template values, in part for the same reasons. Both Figs. 3.8 and 3.9 indicate a clear stratification of the methods, with the methods that fail to treat the multiplicative component of contamination showing significantly larger error.

We also compare the maps in harmonic space. The left panel of Fig. 3.10 shows the per-multipole performance of the cleaning algorithms as $(1 - C_\ell^{\hat{s}s}/C_\ell^{ss})$ vs. the multipole ℓ, where

$$C_\ell^{s\hat{s}} = \langle s_{\ell m} \hat{s}^*_{\ell m} \rangle \tag{3.71}$$

This quantifies the fractional *missing* cross-power between the the true and estimated maps, such that a perfect reconstruction corresponds to 0, and pure noise corresponds to 1 (note

the log scale). This conveys the approximate level of error expected when using cleaned maps for cross-correlation studies.

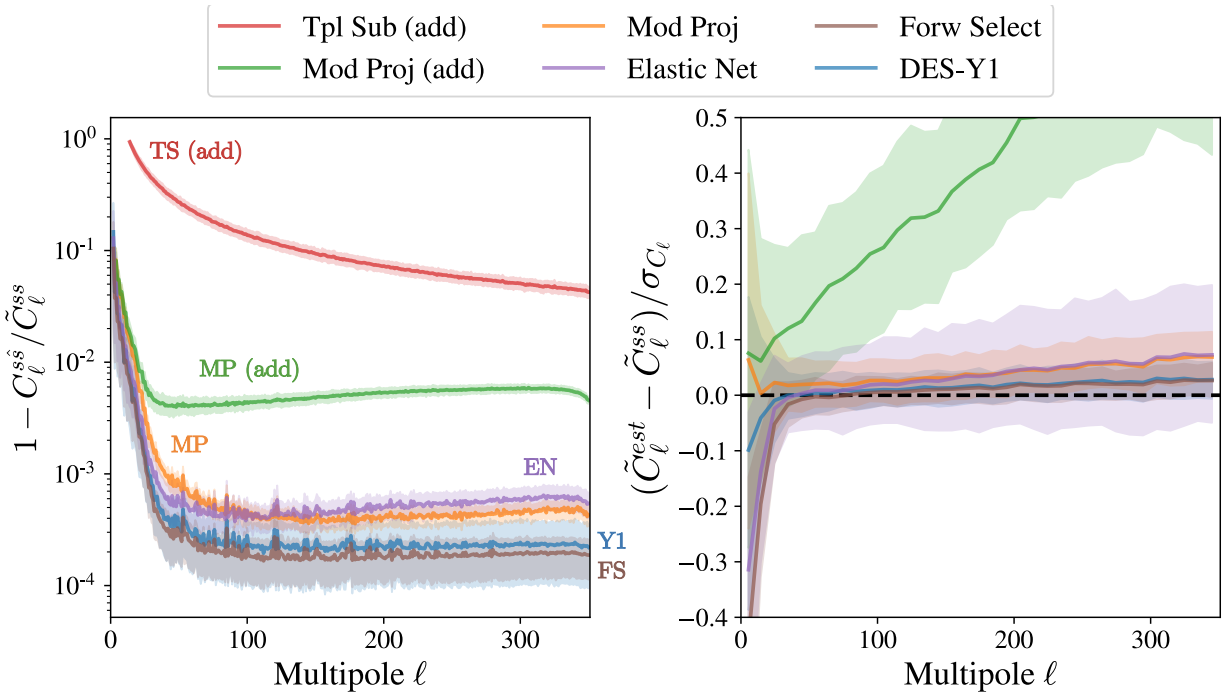

Figure 3.10: *Left*: Error in map reconstruction for each method as a function of multipole ℓ in a DES-like survey, shown as the deficit in correlation at each multipole between the true and cleaned maps $(1 - C_\ell^{s\hat{s}}/C_\ell^{ss})$. A perfect reconstruction corresponds to 0, whereas pure noise corresponds to 1. For all methods (except perhaps Template Subtraction), the cleaned map is a good approximation of the true map for cross-correlation purposes, especially at scales $\ell \gtrsim 30$. *Right*: Error in power spectrum estimation, shown as the residual angular power relative to sample variance $(\tilde{C}_\ell^{\rm est} - \tilde{C}_\ell^{ss})/\sigma_{C_\ell}$ in bins of $\Delta\ell = 10$. Solid lines indicate means of cleaning performed on 50 signal realizations of each bin and shaded regions indicate the central 68% probability mass of the 250 total realizations. The multiplicative correction applied to Mode Projection removes most of the bias of the method (green to orange). Here we use 27 templates, of which 11 are contaminating the data.

All of the methods that treat the multiplicative contamination perform significantly better than the additive methods. The corrected Mode Projection and Elastic Net, and the DES-Y1 method all have excellent performance at $\ell \gtrsim 30$ or scales below about 0.2 degrees on the sky, showing $\lesssim 0.1\%$ error. Maps cleaned with these methods should therefore be excellent candidates for cross-correlation studies. Even the additive mode projection method performs quite well with error of $\lesssim 1\%$ in this case, and as such it may be adequate for many studies.

In the right panel of Fig. 3.10 we show the error in the power spectrum estimate as the difference between the estimated (after cleaning) and true angular power in bins of $\Delta\ell = 10$

and normalized to sample variance $(\tilde{C}_\ell^{\rm est} - \tilde{C}_\ell^{ss})/\sigma_{C_\ell}$, where for σ_{C_ℓ} we use the standard Gaussian approximation for cosmic variance, scaled by $1/f_{\rm sky}$. Unlike $C_\ell^{s\hat{s}}$, this quantity is insensitive to phase-differences between the true and reconstructed maps of the map.

To lowest order, all of the methods work well, and the residual biases are below cosmic variance for the large angles studied here (note that systematic shifts will become more significant with larger multipole bins). For Mode Projection, the performance is satisfactory only once it is corrected for the multiplicative bias via Eq. (3.46), which both reduces bias and uncertainty in the estimated power spectra. We do not show Template Subtraction on the right for clarity — its mean traces the mean for the additive Mode Projection method, but the dispersion is very large, exceeding the plot limits.

The Elastic Net and Forward Selection methods show a similar deficit at large scales as does Mode Projection before it is debiased via Eq. (3.47). This is because the power spectra of the clustering signal and most of the cleaning templates peak at low ℓ, such that more power is removed from large scales. The contribution from the signal power spectrum to this effect (i.e. heteroskedasticity) is mitigated for the DES-Y1 method, which uses the signal covariance. In practice, biases exhibited by any of the methods for the signal power spectrum could be estimated and removed by running on realistic contaminated mocks.

3.7.2 Susceptibility to Overfitting

Any template-fitting model faces a challenge to neither underfit nor overfit the data. In the case of underfitting, residual contamination will be left over in the map and inferred to be signal. In the case of overfitting, a portion of the signal will be inadvertently removed from the map, having been mistaken for systematics. Additionally, increasing the number of fitted templates increases the variance of the estimated power spectrum, which will increase the error of $\tilde{C}_\ell^{\rm est}$ in a mean-squared sense [99].

Mode Projection and Template Subtraction address the risk of overfitting by estimating how much signal power is lost from over-correction given the template library and scaling the power spectrum accordingly (Eqs. 3.11 and 3.23). In contrast, the DES-Y1 and Forward Selection methods use thresholds to limit the templates used for cleaning to only those that are most significant, an approach that was also implemented in the *Extended* Mode Projection method of [96] for the QML power spectrum estimator (though as shown by [97], this comes at the cost of an unknown bias in the power spectrum). As described in Sec. 3.5.2, the Elastic Net reduces overfitting by adding a prior on the template coefficients to reduce the number of templates used.

While each of the methods addresses overfitting in its own way, the library of templates fed

to them has in most cases already been narrowed from a much larger set of possible templates through decisions made by researchers. For example, almost all modern surveys observe any given patch of sky multiple times, resulting in multiple values for each observing condition for every pixel. To produce a scalar template map requires compressing these values into a summary statistic and, as it isn't known a priori which statistic will best capture systematic contamination of the data, multiple statistics may be computed, each corresponding to its own template [see e.g. 191]. If just one statistic (such as the mean) is chosen as representative as is often done, there is the very real risk of discarding potential templates that more accurately capture the contamination, resulting in residual contamination, or *under*fitting.

Therefore, one of the key performance metrics for these methods is their ability to handle

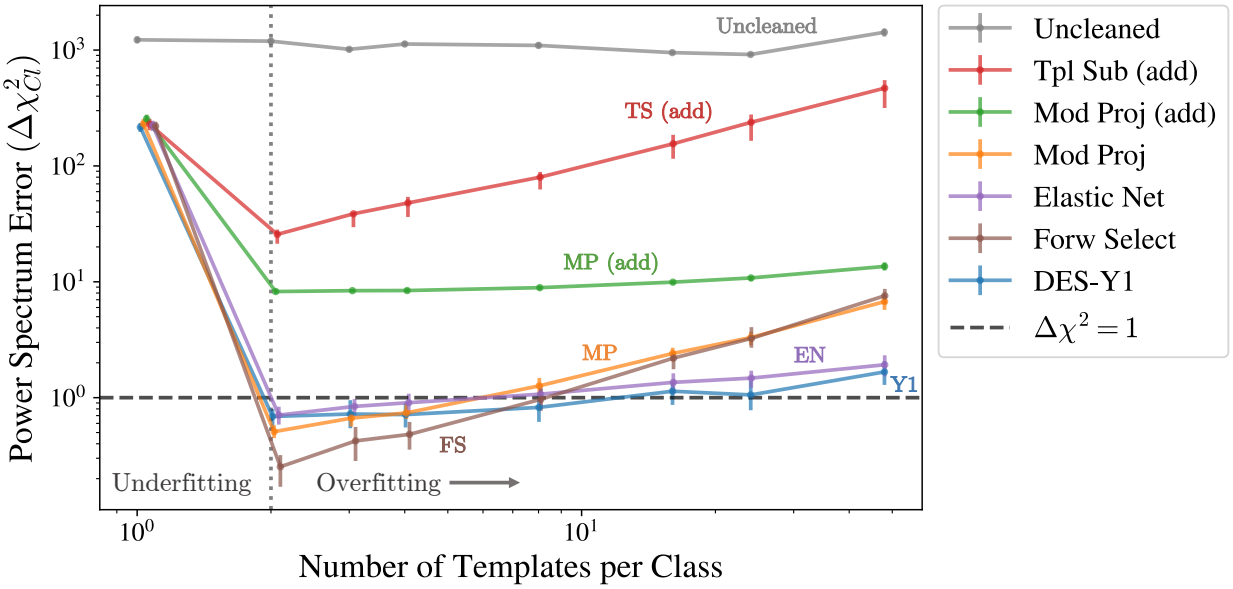

Figure 3.11: Bias in the angular power spectrum, $\Delta\chi^2_{C_\ell}$, as a function of the number of templates fit to the map. We consider Gaussian templates which have a correlation of $\rho_{tpl} = 0.2$ within each of the four template classes, as defined in Sec. 3.6.1. We generate two template realizations per class with which to contaminate each signal map ($N_{\rm sys} = 2$, denoted by the vertical dotted line). The templates used to perform the cleaning vary from one to 48 for each type of template spectrum, for a total of four to 196 templates, with new realizations generated for each $N_{\rm tpl}$ (this is the source of the noise in the 'Uncleaned' line). The Template Subtraction, Mode Projection, and Forward Selection methods are all mildly susceptible to overfitting — signaled by the increase in $\Delta\chi^2_{C_\ell}$ for $N_{\rm tpl} > 2$ — though only Template Subtraction to a degree where it overcomes the penalty for neglecting a contaminating template ($N_{\rm tpl} = 1$). For the *additive* Mode Projection method, $\Delta\chi^2_{C_\ell}$ is dominated by the bias from not addressing the multiplicative contribution to the power spectrum (see Fig. 3.10, right panel), while the other methods are dominated by increased variance from chance correlations. The bias from failing to correct for the multiplicative term dominates even when fitting for ~ 200 templates. The DES-Y1 and Elastic Net display a lesser dependence on $N_{\rm tpl}$, and so are more robust to overfitting. See Sec. 3.7.2 for details.

increasing numbers of non-contaminating templates without degrading map or power spectrum estimates, and so simultaneously mitigate the risks of under- and over-fitting.

To characterize the error in the reconstructed angular power spectrum, we use the sum of squared errors between the true and reconstructed power spectra, normalized by sample variance:

$$\Delta\chi^2_{\mathrm{C}_\ell} = \sum_{z\mathrm{bins}} \sum_{\ell=\ell_{\min}}^{350} \frac{\left(\tilde{C}_\ell^{\mathrm{est}}(z) - \tilde{C}_\ell^{ss}(z)\right)^2}{\sigma^2_{C_\ell^{ss}(z)}}, \tag{3.72}$$

where $\ell_{\min} = 2$, except for Template Subtraction where $\ell_{\min} = \mathrm{Ceil}[(N_{\mathrm{tpl}} - 1)/2]$, since with N_{tpl} templates, all signal is removed for $\ell \leq (N_{\mathrm{tpl}} - 1)/2$.

In Fig. 3.11 we show $\Delta\chi^2_{\mathrm{C}_\ell}$ as a function of the number of templates used to clean the maps (Fig. B.5 shows the same plot for map-level statistics, which demonstrate very similar behavior to $\Delta\chi^2_{\mathrm{C}_\ell}$). We generate two template realizations per class with which to contaminate each signal map, and vary the number of templates used to perform the cleaning from one to 24 for *each template class*. The true contaminants are always 'selected first', such that $N_{\mathrm{tpl}} = N_{\mathrm{sys}} = 2$ represents correctly fitting for the two contaminating templates from each class (vertical dotted line), whereas $N_{\mathrm{tpl}} > 2$ indicates the penalty for overfitting of non-contaminating templates. The error bars come from many signal realizations for the same template maps, and different template and signal map realizations are used for each value of N_{tpl}.

Fig. 3.11 demonstrates that all methods are susceptible to overfitting, as indicated by the fact that $\Delta\chi^2_{\mathrm{C}_\ell}$ increases for $N_{\mathrm{tpl}} > 2$, but that some are more susceptible than others. Template Subtraction and additive Mode Projection are the worst-performing methods with $\Delta\chi^2_{\mathrm{C}_\ell} \gtrsim 10$ for all cases, with Template Subtraction showing a strong dependence on N_{tpl}. Multiplicative Mode Projection and Forward Selection display approximately the same $\Delta\chi^2_{\mathrm{C}_\ell} \tilde{\propto} N_{\mathrm{tpl}}$ scaling as Template Subtraction, whereas The Elastic Net and DES-Y1 methods show a much weaker scaling, indicating that they are much more robust to a larger number of templates.

The trend for *additive* Mode Projection method indicates the importance of the multiplicative correction. Here, the error in the power spectrum does not scale with N_{tpl} as strongly as that of Template Subtraction or the multiplicative Mode Projection method because it is dominated by the bias from not addressing the multiplicative contribution to the power spectrum (see Fig. 3.10, right panel), not the increased variance from a larger number of templates. The bias from failing to correct for the multiplicative term dominates the additive Mode Projection error even when overfitting by ~ 200 templates (or equivalently, roughly 19 templates to quadratic order). Were the plot to continue to the right, we would

expect the error to begin to scale similarly to the other Mode Projection and Template Subtraction methods.[21]

Another key point is that for all cases except Template Subtraction, the penalty for overfitting is dwarfed by the penalty for neglecting contaminating templates ($N_{\rm tpl} = 1$ on the x-axis). This suggests that the researchers should err on the side of overfitting, rather than risk removing contaminating templates from the cleaning library. This is especially true if using a method that is more robust to overfitting, such as the Elastic Net or DES-Y1 method. In sum, the DES-Y1, Forward Selection, Mode Projection with multiplicative correction, and Elastic Net methods all perform very well relative to the uncleaned case, with the Elastic Net and DES-Y1 methods being most robust to overfitting and achieving the best performance with $\Delta\chi^2_{{\rm C}_\ell} \simeq 1$ even when $N_{\rm tpl} \gg N_{\rm sys}$.

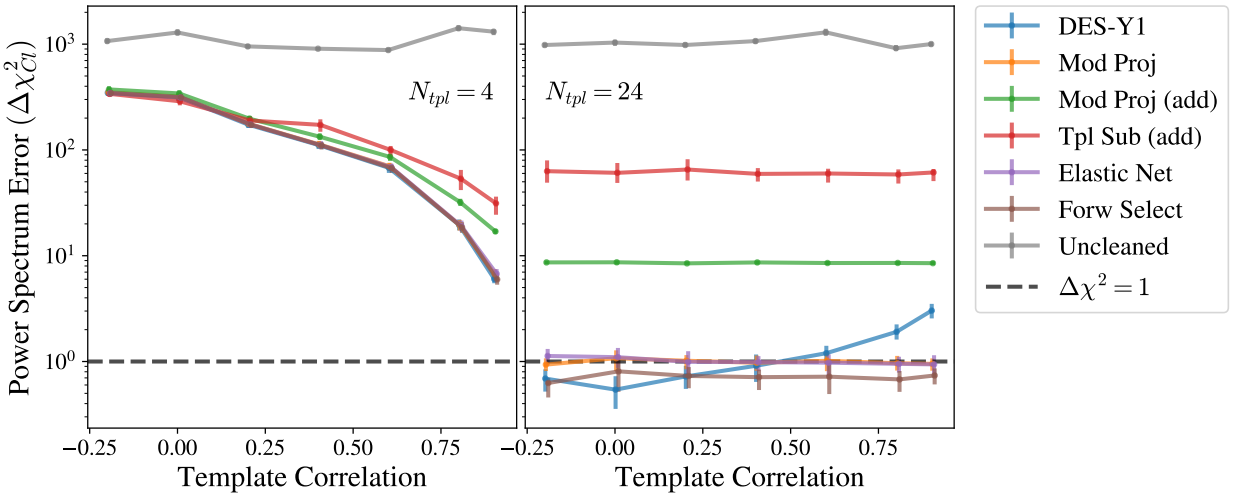

Figure 3.12: Error in the angular power spectrum, $\Delta\chi^2_{{\rm C}_\ell}$, as a function of the level of cross-correlation imposed between the templates within the same class. We assumed contamination from two realizations from each of the four classes (i.e. $N_{\rm sys} = 8$). The left panel assumes cleaning with only one of the contaminating templates from each class, while in the right panel we clean for four templates from each class, including the contaminating ones. Note that in the case where template correlation $\rho_{\rm tpl} \to 1$, the two templates are identical and it is equivalent to cleaning only for one contaminating templates, an ideal scenario. In the right panel we see that while the DES-Y1 outperforms others when templates are completely orthogonal, it suffers as the level of correlation between templates increases. The Elastic Net method mitigates this problem.

[21] The multiplicative bias is not the dominant contribution for Template Subtraction because its *effective* number of templates is much larger, since it performs N_ℓ regressions for each template.

3.7.3 Impact of Correlated Templates

Real templates often have groups of templates that are highly similar to one another in their spectral behavior and/or in their correlation to one another, which we have modeled here as different template classes. The same tracer/property measured in different wavelength bands, or different summary statistics (e.g. the mean vs. median) for the same tracer in a multi-epoch survey are both common examples that can result in very similar templates. We wish to investigate the impact of selecting a non-optimal template for cleaning, which only partially describes the true systematic. This could be either through the choice of a non-optimal summary statistic, or through the apriori choice of a 'representative' template from a group of similar templates in order to mitigate the risk of overfitting, as is commonly done in current surveys.

We test this by cleaning with sets of templates that have varying levels of within-class correlation. For each template class (corresponding to one of the spectra listed in Sec. 3.6.1) we use `Healpy.synfast` to generate template realizations with off-diagonal covariance terms between templates i and j of

$$C_\ell^{ij} = \begin{cases} \rho_{\rm tpl}\sqrt{C_\ell^{ii}C_\ell^{jj}}, & \text{if } i \text{ and } j \text{ in same class} \\ 0, & \text{if } i \text{ and } j \text{ in different classes} \end{cases}$$

We only use the first four classes from Sec. 3.6.1, which are defined by their spectrum and from which we can generate multiple Gaussian realizations with defined levels of cross-correlation.

Fig. 3.12 shows the performance of the methods when the within-class correlation between templates is varied. We again consider the case of two contaminating systematics from each of the four Gaussian template classes. The left panel shows the case where for each class we have chosen only one of the templates to clean with, deeming it "representative" of the template group. As within-class correlation between the systematics increases, the cleaning templates are more representative and can increasingly remove more of the unaccounted-for contamination. At $\rho_{\rm tpl} = 0.9$, the multiplicative methods are able to reduce the error to $\Delta\chi^2_{{\rm C}_\ell} \sim 6$ compared to $\Delta\chi^2_{{\rm C}_\ell} \sim 300$ for the uncorrelated case.

Despite the additional freedom of the Template Subtraction method to fit multipoles independently, it does not do a better job than the other methods of correcting for the "unknown" systematics. The multiplicative methods have almost identical performance, with the dominant contributions to residual errors in the power spectrum resulting from the unaccounted-for systematics and, to a lesser extent, failing to treat the multiplicative term of the contamination.

The right panel in Fig. 3.12 illustrates the other approach of including many possible templates rather than preselecting a few: we use six cleaning templates from each class: the two true systematics, plus four more that are uncontaminating, for a total of 24. We find that with the exception of the DES-Y1 method, performance of the methods is largely independent of the correlation between templates.[22]

Comparing the panels, even if using a high threshold of similarity of $\rho_{\rm tpl} = 0.9$ to discard templates, significantly more error is introduced through neglecting a contaminating template than through overfitting, so it is better to not pre-select templates solely on the basis of similarity to others and instead err on the side of too many templates rather than too few. Template Subtraction is the one exception to this, where each additional template results in N_ℓ additional fits. While the additional freedom does not substantially protect against unknown systematics, it does result in a much steeper penalty for overfitting from higher $N_{\rm tpl}$.

3.7.4 Extensions

By interpreting current LSS systematics cleaning methods in the context of regression, we have facilitated their comparison and interpretation, as well as motivated several possible extensions to them. We have explored some of these extensions in this work, such as the Elastic Net method in Sec. 3.5.2, and the use of the leverage statistic to predict overdensity errors and aid mask creation, but with the extensive body of regression methods, there are many more that we must leave to future work. For example, one promising avenue for regression methods that use a threshold for template selection would be to motivate that threshold by controlling the ratio of Type I (false correction) to Type II (false omission) errors in the selection process via the False Discovery Rate [235], based on the relative impact of each type of error on the analysis.

We have noted individually multiple cases where the assumptions made by the methods do not hold and how they might be improved. A full treatment of these effects is beyond the scope of this study and would include the full non-Gaussian likelihood of $P(d_{\rm obs}|\hat{f}_{\rm sys})$, including contributions from systematics, but as we show in Apps. B.1 and B.2, the corrections from these are minor compared to the methodological differences and the improvements

[22] We found this to be true for both map-level and 2-pt reconstruction statistics, though we only show the latter here. It is not obvious from the outset that this would be the case — Forward Selection methods are often criticized for being less reliable when predictors are correlated, though this is in the context of the more typical regression scenario where it is the predictors themselves that are of interest, as opposed to the residuals which is our focus here. The source of the dependence of the DES-Y1 results on $\rho_{\rm tpl}$ is not entirely clear, but our investigations found it to be mildly impacted by both binning choices and the total monopole of systematic maps.

we suggest. Generalized linear models may be a promising compromise for future mitigation routines, preserving off-the-shelf implementation and diagnostic tools, while providing greater specificity for the likelihood and relaxing some of the tacit assumptions of Mode Projection and OLS regression.

The methods presented here are general enough to be applicable in any situation where one has an external prediction (template) for systematic contamination of observational data, and is equally applicable to spin-2 fields. The insights gained can be used to further extend linear models like the ones in this work, or inform the formulation of nonlinear contamination models, non-parametric methods, or machine learning approaches such as that of Rezaie et al [200].

3.8 Summary of Methods

Here we summarize our findings about the performance of systematic-cleaning methods.

- **DES-Y1 method**: The most complicated method of the ones we studied, the DES-Y1 method resulted in some of the lowest biases in the cleaned maps. It usefully includes prior information about the covariance between pixels in the fitting procedure, albeit in a coarse way. However it is also somewhat complicated to implement, as it requires a large number of parameter choices on the part of the researcher (binning number and procedure, significance statistic and threshold, power spectrum prior) and the generation of realistic mocks. We observed some degradation of its performance as the correlation between templates increased. It is one of the two methods most robust to overfitting when using a large library of templates that are not actually contaminating the data (the other being Elastic Net).

- **Mode Projection:** The standard pseudo-C_ℓ Mode Projection method, as introduced in [99] and implemented in `NaMaster`[13]. We showed that it is equivalent to removing the result of an ordinary least squares regression of the observed data onto the template maps (thus providing a map estimate), with an additional step to debias the power spectrum. This removes most of the contamination present, but can be simply adapted to, and significantly improved by, treating the multiplicative component of contamination instead of just the additive term. We demonstrate how to do this in Sec. 3.4.2. In all cases we studied, the error from not correcting the multiplicative term dominated over error induced from overfitting — as Fig. 3.11 illustrates, in the *ideal*

case where our templates exactly matched the systematics, not treating the multiplicative term introduced as much error as using $\sim 30\times$ more templates than systematics in the cleaning procedure.

- **Template Subtraction:** Equivalent to performing an individual OLS regression at each multipole, resulting in large variance and significant loss of signal from overfitting. As a result, it does not reconstruct maps well and generally performs most poorly in all of our tests. However, our implementation is a limiting case, where each harmonic from each template is allowed to contaminate independently, in contrast to Mode Projection where all modes contaminate identically. The work here should make it straightforward to construct a hybrid method where all modes contribute identically like in Mode Projection (as is physically motivated) and hence have small variance, but where certain modes are prioritized for cleaning, based on the analysis case.

- **Iterative Forward Selection:** This is a method we propose, which is a much simpler version of the DES-Y1 method that requires only a single tunable parameter (a significance threshold) and no mocks. We found that it produces excellent results and is robust to correlation between templates, but is not as robust to overfitting, displaying the same dependence of roughly $\Delta\chi^2_{\mathrm{C}_\ell} \propto N_{\mathrm{tpl}}$ as the Mode Projection and Template Subtraction methods.

- **Cross-Validated Elastic Net**: A cleaning method we introduce, which we find has the best overall performance, being consistently low error and robust to overfitting. It is equivalent to Mode Projection, but with the amplitude of contamination for each template having a mixed Gaussian/Laplace prior applied to encourage sparsity and thus automatically select the important templates. The 'priors' are not strictly such in a Bayesian sense, as their strengths are determined by the data through cross-validation. It is easy to implement using out-of-the-box software and doesn't require a user-defined prior for the power spectrum or debiasing step, providing the best balance of performance, ease of implementation, interpretability and robustness.

3.9 Conclusions

In this study, we carried out a broad comparison of methods used to remove astrophysical, atmospheric, and instrumental systematic errors that affect galaxy-clustering measurements. We have generalized previous work by 1) showing how different methods can be interpreted under a common regression framework, 2) jointly assessing the robustness of methods on

simulated data, 3) investigating the reconstruction fidelity of LSS *map(s)*, rather than just their clustering statistics, as the maps are useful points of departure for numerous other analyses (e.g. summary statistics beyond the power spectrum, cross-correlations, searches for signatures of dark matter or exotic new physics); and 4) proposing improvements to current methods, as well as new, hybrid and efficient methods for the systematics cleaning.

We employed a simple and general model for systematics, given in Eq. (3.1), which allows for spatially varying multiplicative and additive systematic errors with a range of clustering properties to any generic cosmological field. Equipped with that model, we defined a testing procedure that attempts to mimic real-world conditions for LSS surveys, where the true galaxy map is contaminated with an unknown set of systematics and a set of known templates is used to model and correct for the contamination. Given our methodology (pictorially described in Fig. 3.2) and a set of assumptions about the fiducial DES Y5-like survey used to generate the maps, we studied the performance of the systematics-cleaning methods under different conditions.

We showed that both Template Subtraction and Mode Projection, while developed independently, can be interpreted through a regression framework where the signal of interest corresponds to the noise term of a regression model. This allowed us to straightforwardly apply known statistical results and techniques to these methods. We used this to adapt additive methods to account for multiplicative errors (Fig. 3.6), and identify potentially highly contaminated map pixels as a function of their "leverage" (Fig. 3.8), while opening up avenues for further improvement. One such avenue we touched on was to optimize Mode Projection (or other regression methods) by prewhitening the maps in harmonic space. Recognizing that the noise of the regression is the clustering signal itself, we proposed that the maps could be efficiently and optimally inverse-variance weighted in harmonic space, where the clustering signal is diagonal. This is equivalent to accounting for the off-diagonal pixel covariance in the pixel-based regression methods, which is rarely done for tractability reasons (but see Wagoner et al [201] for one approach). We found this to improve results (Fig. B.3), but be subdominant to the multiplicative correction and differences between the cleaning methods.

We introduced two new methods for cleaning: (1) the 'Forward Selection' method, which is a greatly simplified version of the DES-Y1 method that achieves similar performance albeit being less robust to a large number of templates; and (2) the 'Elastic Net' method, a simple out-of-the-box method that implements Mode Projection, but which automatically selects important templates. We found that the Elastic Net method is very robust, with strong performance even when there is a large number of templates (Fig. 3.11) or templates are highly correlated (Fig. 3.12); both are cases where other methods display weaknesses. This

method is very easy to implement, and we recommend it for future surveys.

On the whole, we found that all of the methods perform quite well, dramatically improving the chi-squared difference between the cleaned and true (uncontaminated) angular power spectrum. At the map level, Template Subtraction was the only method that did not significantly reduce the RMS overdensity error across pixels (Figs. 3.8 and B.5), and so we do not recommend the version implemented here for map reconstruction. Once we adopted only the algorithms that take into account both additive and multiplicative errors, all of the methods improved $\Delta\chi^2_{C_\ell}$ by three orders of magnitude relative to the uncleaned case. Moreover, overfitting did not lead to large degradation in the reconstructed power spectra (see Fig. 3.11), which is encouraging. Finally we found that the performance of the various systematics-cleaning methods is very weakly dependent on the level of cross-correlation between the template maps used for the cleaning, with the DES-Y1 method being mildly more susceptible.

We end with several recommendations based on this work:

I. Current and future cleaning methods should account for multiplicative contamination. 'Weights' methods like the DES-Y1 method already do this and other methods like (Pseudo-C_ℓ) Mode Projection can easily do so via Eqs. (3.46)-(3.47).

II. Cleaning methods based on a single Ordinary Least Squares regression are equivalent to (Pseudo-C_ℓ) Mode Projection and so should debias inferred two-point functions accordingly. For more complicated methods where the bias cannot be determined analytically, it can be characterized and removed through performing cleaning on mock catalogs.

III. Analyses should err on the side of overfitting rather than underfitting for templates, as the error from the former tends to be small provided templates do not contain any more information about the true density field than would occur by chance. Researchers should avoid arbitrarily removing templates from the library prior to cleaning based solely on their similarity to other templates. Larger template libraries result in increased variance of the map and power spectrum estimators, especially with the very large number of templates that will be available to future surveys. Therefore:

IV. In scenarios where a very large template library is available, the data itself should be used to select a subset for cleaning. Among the methods that we studied, this is accomplished by either a DES-Y1 type method or the Elastic Net with cross validation. Both show good robustness, and the latter is simple to implement with common software. The theoretical connections we have made between methods should make

alternative template selection routines such as those in [191] and [200] simple to adapt and implement.

V. The cleaning methods used thus far can — and should — be viewed in the context of regression, with the estimated overdensity field corresponding to the regression residuals. Researchers should make use of the powerful suite of existing tools and diagnostic measures to assess the validity of regression models when cleaning LSS data (e.g. leverage for outlier detection, Q-Q plots, partial regression/residual plots) and to aid mask creation. This is applicable to all methods studied in this chapter.

Chapter 4

Chasing the Inflationary Spectral Runnings in the Presence of Systematic Errors

4.1 Introduction

Cosmological inflation [236–238] has passed observational tests with flying colors: the combination of the cosmic microwave background (CMB) with measurements of large-scale structure (LSS) confirms that the geometry of the universe is nearly flat and that the spectrum of density fluctuations is almost scale-invariant [239]. The super-horizon fluctuations observed in the temperature-polarization cross-correlation in the CMB behave in precisely the way that inflation predicts [240]. Beyond these successes, the most important upcoming test of inflation is the search for the signature of primordial gravitational waves, which inflation generically predicts, in the CMB polarization. In this work we study the prospects of another important test of inflation: the search for the running of the scalar spectral index.

The primordial power spectrum of curvature fluctuations can be parameterized by Taylor expanding about a pivot scale k_*

$$\frac{k^3}{2\pi^2}P_s(k) = A_s\left(\frac{k}{k_*}\right)^{(n_s-1)+\frac{1}{2}\alpha_s\ln(k/k_*)+\frac{1}{6}\beta_s(\ln(k/k_*))^2+\ldots}, \tag{4.1}$$

where A_s is the scalar amplitude, n_s is the spectral index, and α_s and β_s are its first and second derivatives, respectively, evaluated at the pivot scale k_*. Single-field slow-roll inflation models predict the power spectrum to be nearly scale invariant, i.e. $n_s \approx 1$, a prediction borne out through measurements of the CMB. The *Planck* experiment [241] has constrained

these parameters for the ΛCDM+α_s model, measuring $n_s = 0.968 \pm 0.006$ and $\alpha_s = -0.003 \pm 0.007$ at the pivot $k_* = 0.05$ Mpc^{-1}. Expanded to include the second running, the *Planck* constraints become $n_s = 0.959 \pm 0.006$, $\alpha_s = 0.009 \pm 0.010$, and $\beta_s = 0.025 \pm 0.013$ (see also [242]).

In single-field slow-roll inflationary models, the runnings are of the order $\alpha_s \sim (1-n_s)^2 \sim 10^{-3}$ and $\beta_s \sim (1-n_s)^3 \sim 4 \times 10^{-5}$ [243] (see also [244, 245]), levels far below the sensitivity of the Planck satellite mission. However the orders of runnings are potentially reachable with new generations of CMB and LSS surveys. Detection of the runnings with magnitudes larger than these values would indicate that the mechanism that generated the primordial fluctuations cannot just be described by a single-field slow roll model [246, 247]. It is possible, for example, for large runnings to be generated by modulations to the inflationary potential [248, 249]. It has also been proposed that modulations resulting in a large value of $\beta_s \sim 10^{-3}$ could produce an appreciable number of primordial black holes (PBHs) [250]; at $\beta_s \approx 0.03$, these PBHs would be large enough to be a dark matter candidate [27, 251]. Thus, even bounds on inflationary spectral runnings that are above the level needed to test single-field slow-roll inflation can provide valuable information.

Munoz et al [27] (M17) investigated how well future surveys will be able to measure α_s and β_s, using a CMB Stage 4 (CMB-S4) experiment in combination with various LSS surveys. They find that even with the combination of a billion-object survey such as SKA, the runnings will only be measured to $\sigma_\alpha = 9.3 \times 10^{-4}$ and $\sigma_\beta = 2 \times 10^{-3}$, levels insufficient for a significant detection if the values are near those predicted by single-field slow-roll inflation (see [252–254] for other forecasts on spectral runnings constraints using CMB and future large-scale structure surveys). It is worth noting, however, that these forecasts only make use of LSS data that is comfortably in the linear regime ($k \lesssim 0.1h\,\mathrm{Mpc}^{-1}$). LSS surveys measure tracers of the matter power spectrum $P_m(k, z)$, and in principle can access information deep in the nonlinear regime, up to $k \simeq 1h\,\mathrm{Mpc}^{-1}$ and beyond. Fig. 4.1 shows the characteristic shape of the matter power spectrum (black) and how a non-zero second running has the greatest impact at the largest and smallest scales (green). The combination of large scales accessed by the CMB and small scales accessed by the LSS is particularly important for constraining the spectral index and its running, as the long lever arm in wavenumber helps to break degeneracies with other cosmological parameters.

Using information from small scales (large k) introduces significant challenges, however. Fluctuations in matter density become large at small scales, so at some scale linear perturbation theory becomes insufficient to describe their evolution. There is a significant ongoing effort to improve our understanding of structure growth in this non-linear regime [255–259]. Baryonic effects also become important at these scales and affect the power spectrum of

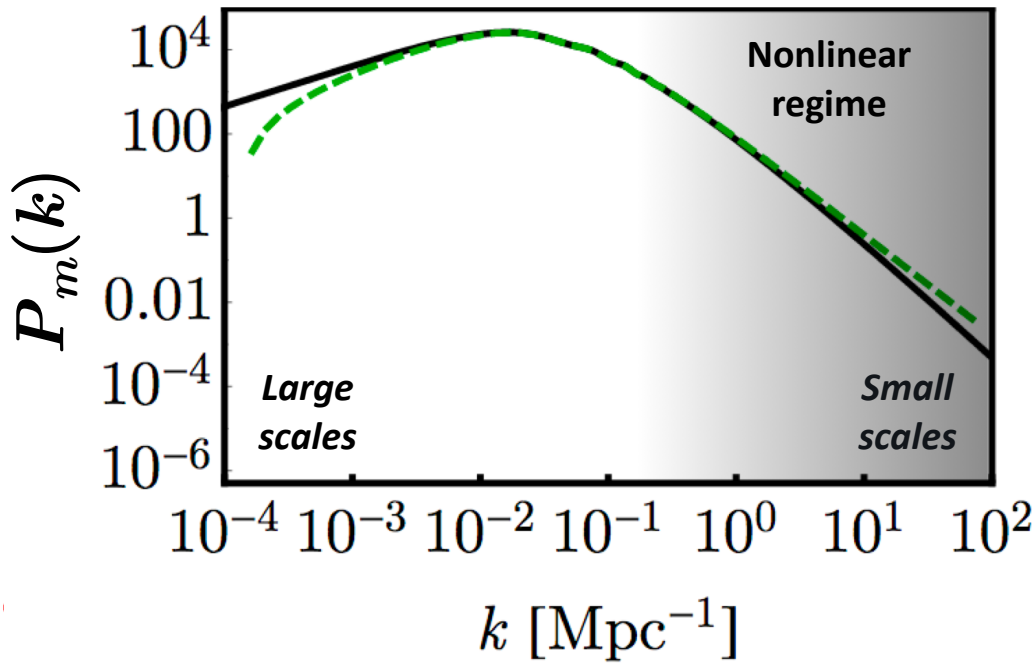

Figure 4.1: Matter power spectrum at the Planck ΛCDM best-fit values (black) and with a non-zero second running of $\beta_s = 0.03$ (green). The largest and smallest scales are those most sensitive to a change in the spectral runnings. While the information content from large scales is limited by cosmic variance, small-scales in the nonlinear regime offer an intriguing means to further constrain the spectral runnings. Shading indicates scales that are increasingly in the nonlinear regime (above $k \gtrsim 0.1h\,\mathrm{Mpc}^{-1}$). Adapted from Ref. [27].

large-scale structure tracers [260–262]. In addition, nonlinearities at small scales induce correlations between wavenumbers [263], so that the covariance of power spectra evaluated at two wavenumbers depends on the nontrivial matter trispectrum.

It is therefore of fundamental importance to understand to what extent the small-scale systematics in the LSS can be parameterized and self-calibrated in order to utilize those scales in the search for α_s and β_s. The main goal of this chapter is to assess how constraints on the runnings improve as LSS information at higher wavenumbers is added to the analysis. We investigate how the results are biased when the nonlinear regime is mismodeled, and how well this bias can be mitigated through the inclusion of nuisance parameters at small scales.

The chapter is organized as follows. In Sec. 4.2, we describe our methodology in detail: our fiducial cosmological model, the CMB and LSS surveys considered, and the Fisher matrix formalism we use for forecasting constraints. In Sec. 4.3, we present and discuss our forecast for the spectral running α_s constraints using future galaxy surveys alone and in combination with CMB-S4. We then introduce the Fisher bias formalism for modeling systematic bias in cosmological parameters, and discuss the corresponding results for α_s in Sec. 4.4. In Sec.

Table 4.1: Cosmological parameters, their fiducial values, and numerical derivative step sizes used for the Fisher matrix calculation. The last two parameters correspond to the Mead model for describing nonlinear effects.

Parameter (p_i)	Fiducial Value	Step Size (Δp_i)
$\Omega_b h^2$	0.02222	$\pm 1\%$
$\Omega_c h^2$	0.1197	$\pm 1\%$
τ	0.06	$\pm 1\%$
H_0	67.5	$\pm 1\%$
n_s	0.9655	$\pm 1\%$
$10^{10} A_s$	21.96	$\pm 1\%$
α_s	0	$\pm 1 \times 10^{-3}$
β_s	0	$\pm 1 \times 10^{-3}$
$A_{\rm bary}$	3.13	$\pm 5\%$
η_0	0.6044	$\pm 5\%$

4.5, we present our constraints and systematic bias results for the second spectral running, β_s. We summarize our findings and conclude in Sec. 4.6.

4.2 Methods

In this section we describe our fiducial model for CMB and LSS observations and describe our forecasting methodology, which makes use of the Fisher matrix formalism to forecast the precision of measurements of the runnings.

4.2.1 Fiducial Model

We assume a flat ΛCDM cosmology with six parameters in addition to the spectral runnings: the physical baryon and CDM densities $\Omega_b h^2$ and $\Omega_c h^2$, the reionization optical depth τ, the Hubble constant H_0, the scalar spectral index n_s, and the primordial power spectrum amplitude A_s. The values of these parameters in our fiducial model are listed in Table 4.1.

4.2.2 Modeling the CMB

The CMB fluctuations have a wealth of information about the early universe, providing some of the tightest constraints for cosmology to date [264]. The observed CMB angular power spectrum can be related to the primordial power spectrum $P_s(k)$ that sourced those fluctuations via

$$C_\ell^{XY} = \frac{2}{\pi} \int dk\, k^2 P_s(k) \Delta_\ell^X(k) \Delta_\ell^Y(k). \tag{4.2}$$

Labels X and Y can refer to temperature (T), polarization modes (E, B), or lensing potential (d), and Δ_ℓ^X is the transfer function which encompasses both source and projection terms integrated over the line-of-sight.

Taking T and E as our observables, the observed angular power spectra can be represented as a vector (C_ℓ^{TT}, C_ℓ^{EE}, C_ℓ^{TE}) with covariance matrix

$$\mathrm{Cov}_\ell = \left(\frac{2}{(2\ell+1) f_{sky}}\right) \begin{bmatrix} (\tilde{C}_\ell^{TT})^2 & (\tilde{C}_\ell^{TE})^2 & \tilde{C}_\ell^{TT}\tilde{C}_\ell^{TE} \\ (\tilde{C}_\ell^{TE})^2 & (\tilde{C}_\ell^{EE})^2 & \tilde{C}_\ell^{EE}\tilde{C}_\ell^{TE} \\ \tilde{C}_\ell^{TT}\tilde{C}_\ell^{TE} & \tilde{C}_\ell^{EE}\tilde{C}_\ell^{TE} & \frac{1}{2}[(\tilde{C}_\ell^{TE})^2 + \tilde{C}_\ell^{TT}\tilde{C}_\ell^{EE}] \end{bmatrix}$$

where the auto power spectra include contributions from noise:

$$\begin{aligned} \tilde{C}_\ell^{TT} &= C_\ell^{TT} + N_\ell^{TT} \\ \tilde{C}_\ell^{EE} &= C_\ell^{EE} + N_\ell^{EE} \\ \tilde{C}_\ell^{TE} &= C_\ell^{TE} \end{aligned} \tag{4.3}$$

We adopt the same noise properties for a CMB-S4 experiment used by [27]:

$$N_\ell^{TT} = \Delta_T^2 \exp\left[\frac{\ell(\ell+1)\theta_{\mathrm{FWHM}}^2}{8 \ln 2}\right] \tag{4.4}$$

and

$$N_\ell^{EE} = 2 \times N_\ell^{TT}\ , \tag{4.5}$$

where the temperature sensitivity is $\Delta_T = 1\mu$K-arcmin and the beam full-width-half-maximum is $\theta_{\mathrm{FWHM}} = 8.7 \times 10^{-4}$ radians [265]. We assume a sky coverage $f_{\mathrm{sky}} = 0.4$ and that the usable range of multipoles are $\ell \in [30, 3000]$ for C_ℓ^{TT} and C_ℓ^{TE}, and $\ell \in [30, 5000]$ for C_ℓ^{EE}. To represent additional constraints coming from low-ℓ polarization (e.g. from the Planck

High Frequency Instrument) which break the degeneracy between τ and A_s [266], we include a Gaussian prior on τ with width $\sigma(\tau) = 0.01$.

The nonlinearity of matter fluctuations affects the CMB power spectrum at small scales mainly through lensing. While the effect on the CMB lensing power spectrum from the large-scale structure bispectrum can be significant [267], the corresponding changes in the TT, EE and TE angular power spectra are negligible [268]. We therefore do not consider the modeling uncertainties from nonlinear lensing effects on the CMB power spectra in this work.

4.2.3 Modeling Large Scale Structure Surveys

LSS surveys utilize a variety of tracers in order to probe the growth of structure in the universe as a function of cosmic time, such as galaxies, quasars, and the Lyman-alpha forest. These measurements, in turn, enable strong constraints to be placed on both early- and late-universe parameters [269–271].

In the linear regime, the matter power spectrum can be computed for a given cosmology using Boltzmann codes such as `CAMB` [272] or `CLASS` [273]. On smaller scales where linear perturbation theory breaks down, one must resort to other methods. These may include N-body or hydrodynamical simulations, or else semi-analytic prescriptions, for example ones based on the halo model of LSS [274–276]. However these methods are not guaranteed to capture all the relevant physics. The presence of redshift space distortions (RSD), which render the power spectrum observed in redshift space anisotropic, further complicates matters.

Because we aim to investigate the impact of systematic errors on constraints from LSS, and those are mainly due to modeling uncertainties at small scales, we parameterize the observed galaxy power spectrum in a way that allows us to generically encapsulate modifications to our fiducial power spectrum due to nonlinear effects. Following Seo and Eisenstein [277], we write the redshift-space power spectrum of tracer X as

$$P_{\rm obs}(k,\mu,z) = b_X^2(z)\left(1+\frac{f(z)}{b_X(z)}\mu^2\right)^2 P_m(k,z)\exp\left[-\frac{k^2\mu^2\sigma_v^2}{H_0^2}\right] M_{\rm nl}(k,z) \tag{4.6}$$

where $P_m(k,z)$ is the matter power spectrum from `CAMB` with nonlinear corrections from `HMcode` [28], μ is the cosine of the angle between the line connecting galaxy pairs and the line of sight and $f(z) = \mathrm{d}\ln D/\mathrm{d}\ln a$ is the logarithmic derivative of the linear growth factor. The exponential term, featuring the velocity dispersion σ_v, models the power suppression along the line-of-sight at small scales due to redshift-space distortions (the so-called Fingers-of-God effect). Here σ_v is calculated using the virial scaling relation from [278], evaluated

at the characteristic mass of collapsed halos (M_*). We find that the effect only has a minor impact, slightly increasing the forecasted errors at $k_{\rm max} > 1h\,{\rm Mpc}^{-1}$. The impact of baryons and other effects on nonlinear scales (henceforth *nonlinear effects*) are accounted for by the as-yet undefined function $M_{\rm nl}(k,z)$. The term b_X^2 describes the linear galaxy bias for galaxy population X, which we define to have the redshift dependence of $b^2(z) = b_0^2(1+z)$. We marginalize over the amplitude b_0 when determining cosmological parameter constraints, and absorb any scale-dependent bias effects into $M_{\rm nl}(k,z)$.

We now turn to the "nonlinear" function $M_{\rm nl}(k,z)$. We consider three treatments, in order of increasing complexity:

1. **No Nuisance model:** The simplest case is the trivial one where the nonlinear power is assumed to be modeled perfectly by the modified halo model prescription in `HMCode` and there is no scale-dependent bias. This corresponds to $M_{\rm nl}(k,z) = 1$, with no additional nuisance parameters. We refer to this as the *No Nuisance* model.

2. **Mead model:** The next model for $M_{\rm nl}(k,z)$ is the one presented by Mead et al [28], in which the modifications to nonlinear power due to baryonic feedback effects are parametrized using two parameters [A_b and η_0] (*Mead parameters*). In this case,

$$M_{\rm nl}(k,z) \;=\; \frac{P_{\rm Mead}(k,z,A_b,\eta_0)}{P_{\rm Mead,\ DMonly}(k,z)} \tag{4.7}$$

   where "DMonly" refers to the default Mead parameter values of $A_b = 3.13$ and $\eta_0 = 0.6044$.

3. **Many Free Parameter (MFP) model:** The final model for the nonlinearities is a much more agnostic prescription similar to Bielefeld et al [279], in which $M_{\rm nl}(k,z)$ is allowed to float freely in bins of wavenumber k and smoothly, as a low-order power-law, in redshift. Since at low k the power spectrum is well determined theoretically, we allow $M_{\rm nl}(k,z)$ to vary only for k at the quasi-linear regime and above, setting it to unity at large scales.

   We therefore have

$$M_{\rm nl}(k,z) \;=\; \begin{cases} (1 + c_{1,k}z + c_{2,k}z^2)B_k & \text{if } k > 0.1 \\ 1 & \text{if } k \leq 0.1 \end{cases} \tag{4.8}$$

   where k has units $h\,{\rm Mpc}^{-1}$, and B_k, $c_{1,k}$ and $c_{2,k}$ are free parameters. One set of $\{B_k,\ c_{1,k},\ c_{2,k}\}$ is specified in each wavenumber bin of width $\Delta \ln k = 0.05h\,{\rm Mpc}^{-1}$. This bin width is fixed, so as the maximum wavenumber $k_{\rm max}$ is raised, the number of k

bins increases, and consequently so does the number of nuisance parameters. The total number of nuisance parameters in $M_{\rm nl}(k,z)$ thus ranges from 0 to 279 as $k_{\rm max}$ is varied from 0.1 to $10h\,{\rm Mpc}^{-1}$, and hence we refer to this as the *Many Free Parameter* (*MFP*) model.

The covariance between the observed power spectrum at wave numbers k_α and k_β is given as the sum of the "unconnected" part, which is diagonal in the two wavenumbers, and the connected contribution given by the full trispectrum:

$$[{\rm Cov}]_{k_\alpha,k_\beta} = \frac{8\pi^2[P_{\rm obs}(k_\alpha,\mu,z)]^2}{V_{\rm eff}(k_\alpha,\mu,z)k_\alpha^2\Delta k_\alpha}\delta_{k_\alpha,k_\beta} + T_{k_\alpha,k_\beta}. \tag{4.9}$$

The effective volume of the survey varies with redshift and is given by

$$V_{\rm eff}(k,\mu,z) = V(z)_{\rm survey}\left[\frac{n(z)P_{\rm obs}(k,\mu,z)}{1+n(z)P_{\rm obs}(k,\mu,z)}\right]^2 \tag{4.10}$$

where $n(z)$ is the galaxy number density of each redshift bin and $V(z)_{\rm survey}$ is the volume in $[h^{-1}{\rm Mpc}]^3$,

$$V(z)_{\rm survey} = \int_{z_{\rm min}}^{z_{\rm max}} \Omega_{\rm survey}\frac{r(z')^2}{H(z')}dz'. \tag{4.11}$$

Here, $r(z)$ is the comoving distance, $H(z)$ is the Hubble parameter, and $\Omega_{\rm survey}$ is the sky coverage of the survey in steradians. The term T_{k_α,k_β} is the contribution from the trispectrum due to the non-Gaussian nature of the matter field,

$$T_{k_\alpha,k_\beta} = \int_{k_\alpha}\frac{d^3\boldsymbol{k_1}}{4\pi k_\alpha^2\Delta k_\alpha}\int_{k_\beta}\frac{d^3\boldsymbol{k_2}}{4\pi k_\beta^2\Delta k_\beta}T(\boldsymbol{k_1},-\boldsymbol{k_1},\boldsymbol{k_2},-\boldsymbol{k_2}). \tag{4.12}$$

We obtain T_{k_α,k_β} with the same calculation method described by Wu and Huterer [280], who use the halo model to calculate the trispectrum, showing that it is dominated by the one-halo term. We refer the interested reader to that work for details.

In their spectral running constraint forecasts, M17 consider a wide survey like the Dark Energy Spectroscopic Instrument (DESI) [281] as well as a deep and narrow survey similar to the Wide Field Infrared Survey Telescope (WFIRST) [282], finding that they improve constraints on the runnings by $\sim 20\%$ and 30%, respectively, when added to data from a CMB-S4 experiment. Here we take a Euclid-like survey to be our fiducial survey, and we include a DESI-like survey for comparison.

Euclid: Euclid [283] is a proposed space-based LSS survey with large sky coverage and a deep redshift distribution, which should provide excellent constraints on the evolution of dark

energy. We use the spectroscopic sample defined in [283], assuming 15000 deg^2 ($f_{\rm sky} \approx 0.36$) and a total of 50 million galaxies. We use the redshift bins given in Table VI of [284], with thickness $\Delta z = 0.1$ in the range $z \in [0.6, 2.1]$. We infer the effective number density in each bin as $n(z) = \bar{n}P_{0.14,0.6}(z)/P_{\rm obs}(k = 0.14h{\rm Mpc}^{-1}, \mu = 0.6, z)$, where $\bar{n}P_{0.14,0.6}(z)$ is a quantity reported by [284] and $P_{\rm obs}$ is calculated via Eq. (4.6). The resulting $n(z)$ is shown in Fig. 4.2.

DESI: The Dark Energy Spectroscopic Instrument (DESI [281]) is a Stage-IV ground-based spectroscopy experiment at Mayall telescope in Arizona, which will target multiple tracer populations over 14,000 deg^2 ($f_{\rm sky} \approx 0.34$) with good signal to noise out to $z \lesssim 1.5$. Here too we adopt the distribution given in [284], which combines projections for the populations of Luminous Red Galaxies (LRGs), Emission Line Galaxies (ELGs) and quasars (QSOs) into estimates of $\bar{n}P_{0.14,0.6}(z)$ in redshift bins of $\Delta z = 0.1$ in the range $z \in [0.1, 1.9]$. We calculate an effective $n(z)$ for each bin in the same way as with the Euclid-like projections, and show them in Fig. 4.2. We assume that the Euclid-like and DESI-like experiments do not overlap and we combine their information by summing the Fisher matrices, as we describe below.

4.2.4 Forecasting

We forecast uncertainties of cosmological parameters as a function of $k_{\rm max}$ using a Fisher matrix analysis. The Fisher matrix formalism is an extremely simple and efficient method to estimate the errors on model parameters given a set of data [285, 286]. If one approximates the likelihood as a multi-variate Gaussian in the parameters around its peak, the resulting Hessian (matrix of second derivatives) can be used to calculate the forecasted uncertainties in the cosmological parameters. The better the actual constraints on the parameters are, the closer the likelihood function is to a Gaussian distribution, and the more accurate the Fisher matrix approximation is. To the extent that we are assuming powerful future surveys with small errors on most parameters, the Fisher matrix approximation should be excellent. More importantly, given that our MFP systematics case contains up to ∼300 parameters, a Fisher forecast is the *only* reasonably straightforward way to estimate the errors.

Under the assumption of Gaussian perturbations and Gaussian noise, the Fisher Matrix for CMB temperature and polarization anisotropies [287–289] can be written as

$$F_{ij}^{\rm CMB} = \sum_{\ell} A_{\ell,i} ({\rm Cov}_\ell)^{-1} [A_{\ell,j}]^T, \tag{4.13}$$

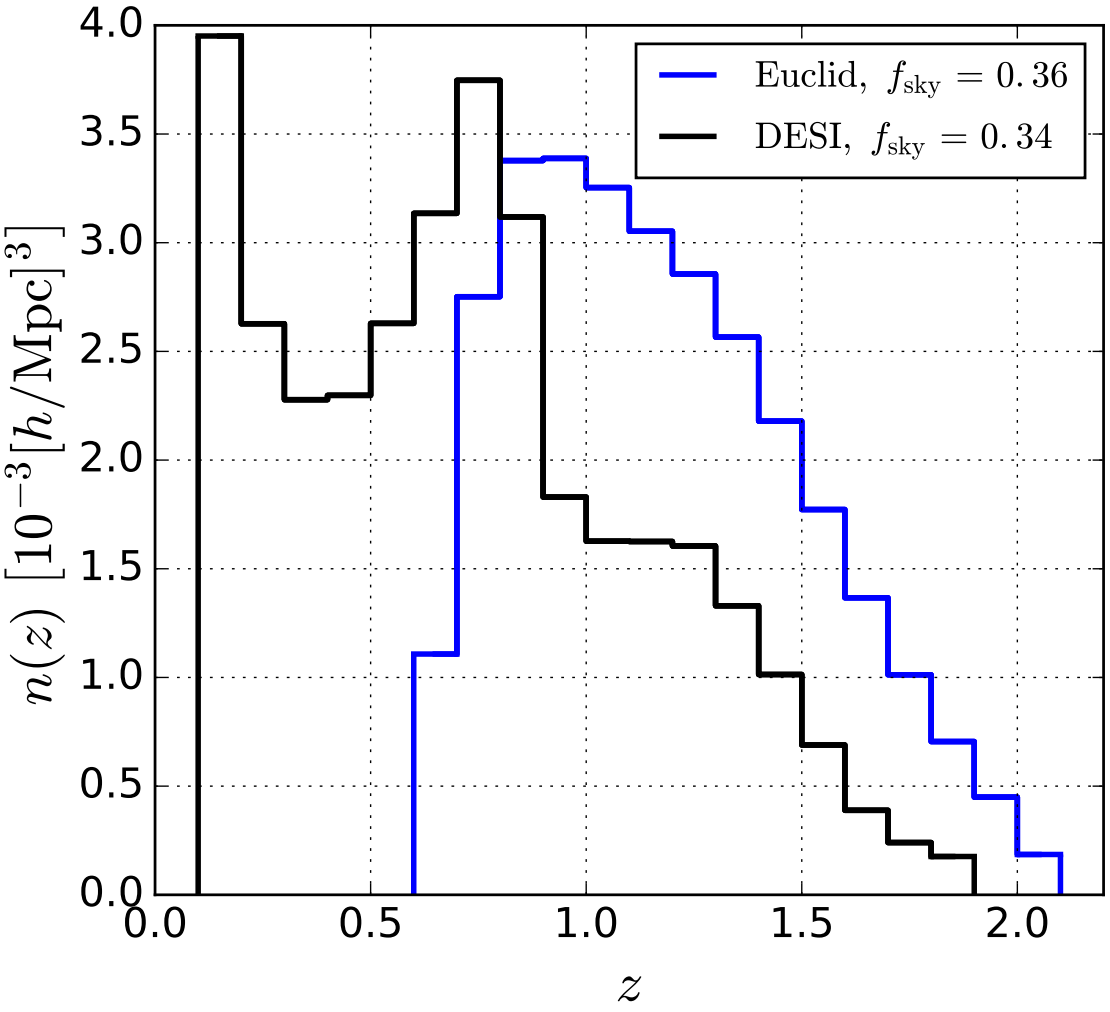

Figure 4.2: Galaxy number density $n(z)$ of Euclid and DESI in each redshift bin. The features in the DESI $n(z)$ are due to the fact that the sample is a combination of several populations of sources.

where

$$A_{\ell,i} = \left(\frac{\partial C_\ell^{TT}}{\partial p_i}, \frac{\partial C_\ell^{EE}}{\partial p_i}, \frac{\partial C_\ell^{TE}}{\partial p_i} \right) \tag{4.14}$$

and the covariance is given by Eq. (4.3). The Fisher matrix for the observed LSS power spectrum is

$$F_{ij}^{\rm LSS} = \sum_z \sum_\mu d\mu \sum_{k_\alpha, k_\beta} \frac{\partial P(k_\alpha, \mu, z)}{\partial p_i} \left[{\rm Cov}^{-1}\right]_{k_\alpha, k_\beta} \frac{\partial P(k_\beta, \mu, z)}{\partial p_j}, \tag{4.15}$$

where the sums are over all bins in z, μ, and k, and p_i runs over the cosmological parameters $\{\Omega_b h^2, \Omega_c h^2, \tau, H_0, n_s, A_s, \alpha_s, \beta_s\}$ as well as the linear bias parameter b_0 and the nuisance parameters in every k-bin, $\{B_k, c_{1,k}, c_{2,k}\}$. We define k bins logarithmically, with $\Delta \ln k = 0.05$ in the range $k_{\rm max} \in [0.1, 10] h\,{\rm Mpc}^{-1}$, and bin μ in 11 evenly spaced bins from -1 to 1.

Forecasts for a combination of experiments can be calculated by summing their Fisher matrices, and a forecast for the lower bound on the the error for a given parameter is given by the Cramer-Rao inequality

$$\sigma(p_i) \geq \begin{cases} \sqrt{(F^{-1})_{ii}} & \text{(marginalized)} \\ 1/\sqrt{F_{ii}} & \text{(unmarginalized).} \end{cases}$$

4.3 Results

We now present the principal results. To give an idea of the approximate overall level of constraint on the cosmological parameters, we summarize the fiducial constraints for our CMB-S4 forecast on the spectral runnings: when fixing $\beta_s = 0$, we obtain marginalized error on the spectral running of $\sigma_\alpha = 3.0 \times 10^{-3}$. When allowing β_s to vary, we find $\sigma_\alpha = 3.4 \times 10^{-3}, \sigma_\beta = 8.0 \times 10^{-3}$. (All constraints listed are the marginalized error, unless otherwise noted.) These constraints are similar to those of M17, although slightly weaker because we do not include lensing information.

We now turn to the main goal of this work: exploring whether and how adding information from LSS improves constraints on the spectral runnings. We first consider galaxy clustering alone, and then in conjunction with CMB-S4.

4.3.1 Galaxy Clustering

To see how information from LSS data at small scales impacts constraints on the first spectral running, we forecast the marginalized 1σ constraints on α_s as a function of $k_{\rm max}$. For the moment, we hold the second running β_s fixed at 0; we will let β_s vary further below, in Sec. 4.5.

Fig. 4.3 shows the increase in constraining power when we include clustering information at small scales, comparing the performance of the No Nuisance (blue), Mead (red), and MFP models (black) for nonlinear effects. We also show constraints for the No Nuisance and MFP cases without the trispectrum contribution to the covariance (dashed), to demonstrate that its contribution to the error budget is minor (see Appendix C.1 for a case where shot noise is suppressed and the trispectrum dominates the error budget).

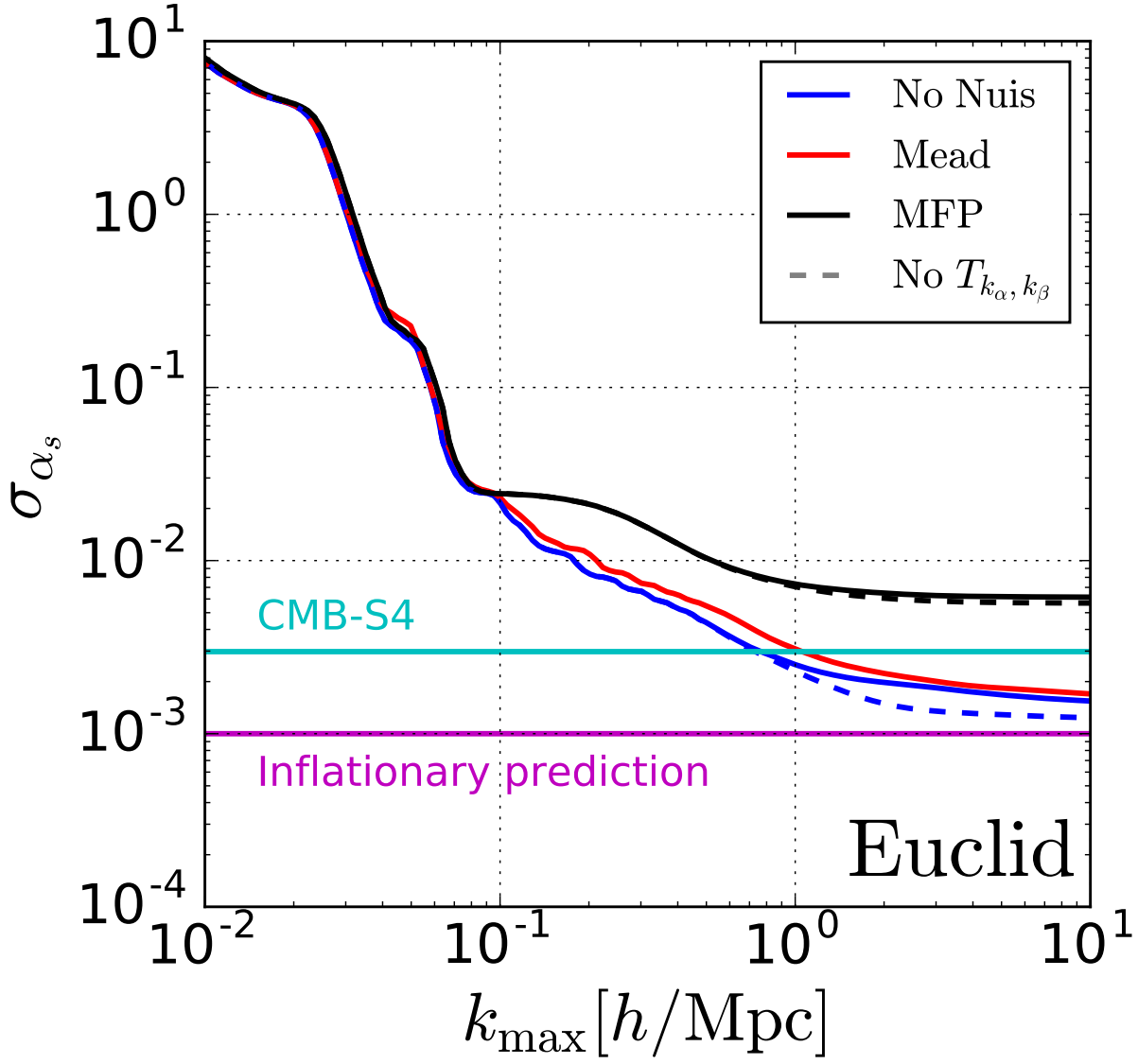

Figure 4.3: Error (1σ here and everywhere) in the spectral running α_s as a function of $k_{\rm max}$, evaluated for our fiducial Euclid-like survey, assuming $\beta_s = 0$. The legend shows our assumption about modeling of the systematics, while T_{k_α,k_β} refers to the inclusion of the trispectrum to the data covariance. Note that here and in subsequent plots, the value of the running denoted as the "Inflationary prediction" (purple horizontal line) is only approximate.

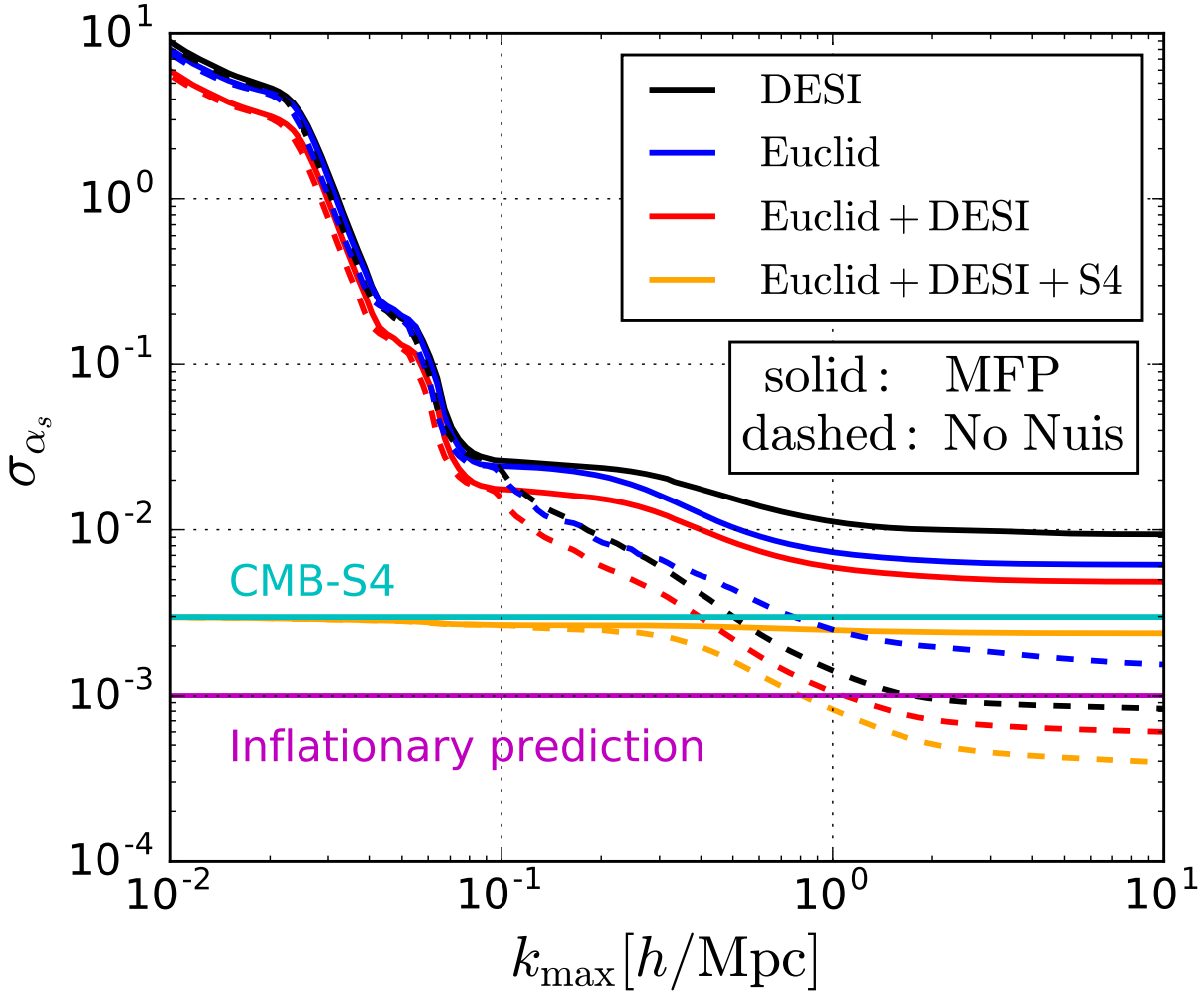

Figure 4.4: Marginalized constraints on α_s when combining information from different surveys (DESI-like, Euclid-like, CMB-S4, and also in combination). Solid curves include the MFP description of the systematic errors in galaxy surveys (see Eq. (4.8), while the dashed curves do not. Results using the Mead parameterization are similar to the No Nuisance (No Nuis) case, and so we we omit them here for clarity.

In the No Nuisance case, that is, the forecast for constraints if no parameters need to be introduced to model nonlinear effects, we find a large gain in constraining power for high $k_{\rm max}$. This gain remains whether or not we include the trispectrum contributions to the power spectrum's covariance at small scales (solid and dashed curves, respectively). However, the overall gain with increasing $k_{\rm max}$, even in this no-systematics case, is not as significant as might be expected based on the behavior at linear scales, because the slope of the σ_α vs $k_{\rm max}$ curve changes at scales where nonlinearities become important, $k_{\rm max} \simeq 0.1h\,{\rm Mpc}^{-1}$. This flattening in $\sigma_\alpha(k_{\rm max})$ implies that, even in the optimistic no-systematics scenario and pushing out to $k_{\rm max} = 10h\,{\rm Mpc}^{-1}$, the Euclid constraint on the running would only be comparable to the expected inflationary signal, $\sigma_\alpha \simeq 10^{-3}$ and so be insufficient for a statistically significant detection of α of that size.

The red and black curves in Fig. 4.3 show how these constraints respond to the addition of nuisance parameters intended to capture nonlinear effects, corresponding to the Mead and MFP models, respectively. The Mead model, which introduces only two new parameters, produces results similar to the No Nuisance case. In contrast, constraints become considerably weaker (e.g. by a factor of $\sim$5 at $k_{\rm max} = 10h\,{\rm Mpc}^{-1}$) for the MFP model, which captures nonlinear effects via an agnostic, piecewise-in-k $M_{\rm nl}(k,z)$ with many free parameters (up to 279 for the highest $k_{\rm max}$). Thus, in this more conservative treatment of small-scale systematics, the gains from including high-k modes are rendered modest at best, particularly for $k_{\rm max} \gtrsim 1h\,{\rm Mpc}^{-1}$. We will show below in Sec. 4.4, however, that the MFP parametrization does protect the constraints against the systematic biases due to modeling uncertainties in the high-k power spectrum.

Clearly, in the comparison of forecasted constraints, the more gentle treatment of systematics (with fewer free parameters) in the Mead model produces more favorable results than the more agnostic MFP case. However, this comparison of statistics-only errors alone is not enough to answer the question of whether a given treatment of systematics is sufficient for an analysis. Rather, modeling choices must be made by balancing the consideration of expected constraining power with the need for nuisance parameters to protect against biases to the best-fit cosmological parameters. Accordingly, we compare our three $M_{\rm nl}(k,z)$ treatments by studying their relative ability to protect against biases in Sec. 4.4.

4.3.2 Galaxy Clustering and CMB

The large lever arm provided by the combination of CMB and LSS allows for much tighter constraints on the running than using LSS data alone. We illustrate this in Fig. 4.4, which gives the marginalized 1-σ constraints on α_s for different $k_{\rm max}$ when combining LSS infor-

mation from a Euclid-like and/or a DESI-like survey with that from a CMB-S4 experiment. We now show only forecasts which include the trispectrum contribution to the covariance, and use the comparison between the solid and dashed curves to compare the performance of the No Nuisance and MFP models, respectively. For clarity, we do not diplay the curves for the Mead model, which are similar to those for the No Nuisance case.

The curves for LSS data alone show results similar to those in Fig. 4.3. We find that DESI and Euclid yield comparable errors in the running (with a ∼30% smaller error for Euclid), with their combination giving a slight improvement over Euclid alone.

As in the Euclid-only case, we see a 5–10× degradation in constraints if the MFP treatment of systematics is adopted compared to the No Nuisance case. We note that this degradation is greater for DESI (black) than for Euclid (blue).

Next we consider the effect of adding CMB-S4 information to the Euclid+DESI combination, which is shown in orange in Fig. 4.4. When large and mildly nonlinear scales of the LSS ($k_{\rm max} \lesssim 0.5h\,{\rm Mpc}^{-1}$) are used, the CMB information dominates the (CMB+LSS) constraining power, and the combined error is essentially equivalent to that from CMB-S4 alone. At smaller scales, the LSS surveys help tighten constraints, but only in the No Nuisance case. In the MFP case, where many nuisance parameters are marginalized over, LSS data adds little constraining power on α_s compared to CMB-S4 data alone.

4.4 Systematic Biases in Model Parameters

The fact that there are significant modeling uncertainties associated with the theoretical prediction of galaxy clustering at small scales is our primary motivation for studying different choices of the $M_{\rm nl}(k,z)$ function to describe nonlinear effects. Any analysis will have to make simplifying choices for how to model the physics of nonlinear structure growth, baryonic effects, and scale-dependent galaxy bias. To the extent that those choices provide an incomplete description of the underlying physics there will be inaccuracies in the theoretical prediction for the observed galaxy power spectrum. Here we examine how these systematic errors—that is, residuals between the true and assumed power spectrum—impact parameter estimation for the spectral running.

In order to characterize this, we represent a typical form for the residuals due to systematic errors by taking the *difference* between two commonly used parameterizations of the matter power spectrum on small scales. Specifically, we subtract the nonlinear prescription by Takahashi et al [6] from that of Mead et al [28].[1] The power spectra generated with these

[1]We take the default parameter values of $A_{\rm bary}$ and η_0 corresponding to the DMONLY case in `HMcode` as of Feb. 2018, which includes the updates of Ref. [7]

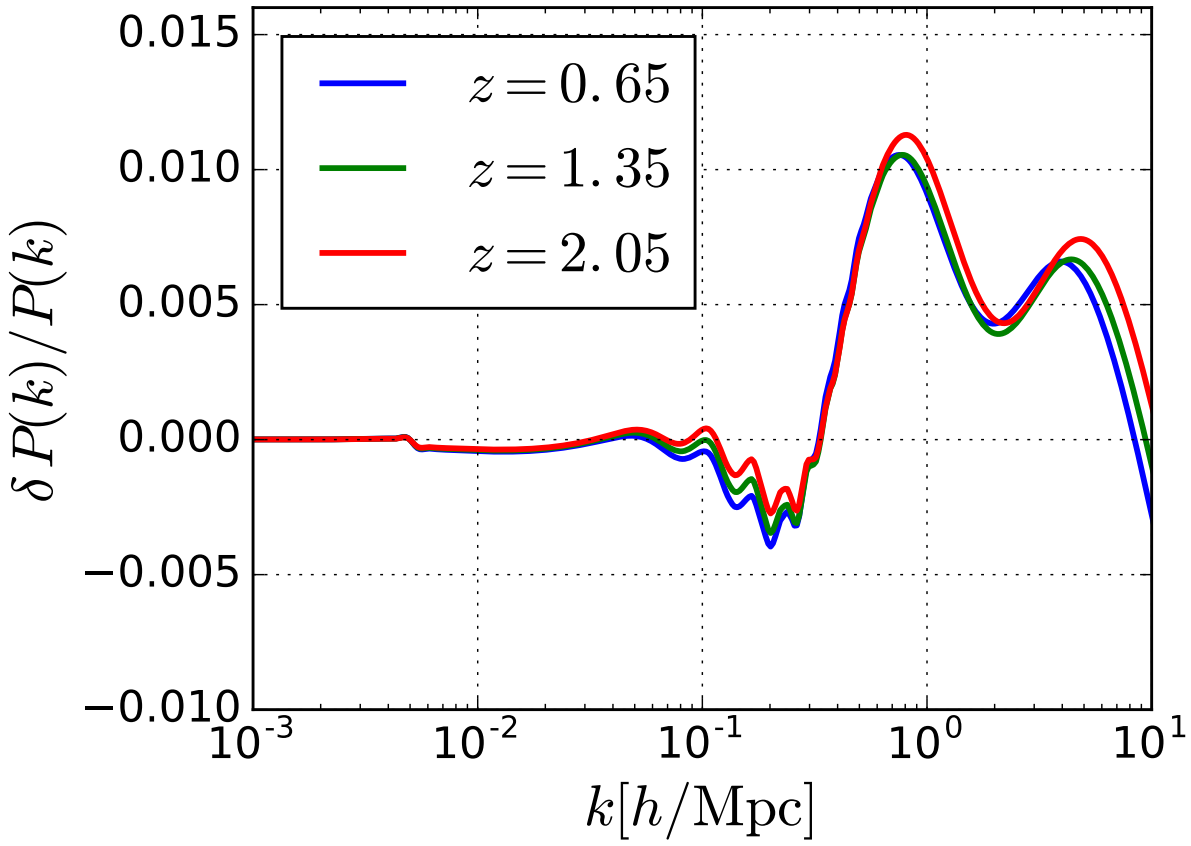

Figure 4.5: Relative difference between the nonlinear predictions from two popular fits: that of Takahashi et al [6] and of Mead et al [28] scaled so as to correspond to about 1% maximum difference at small scales. The quantity shown, $\delta P/P = 0.2(P_{\rm taka} - P_{\rm mead})/P_{\rm mead}$, is the fiducial model for the small-scale systematics that we employ in subsequent plots to gauge the protection offered by our systematics parametrizations.

two codes differ by up to ∼5%, roughly independent of redshift for the range considered. For the future surveys we consider, we optimistically assume that that theoretical advances will allow the small-scale power spectrum to be computed to an accuracy of about 1%. We therefore adopt a fifth of the Takahashi-Mead difference as our fiducial model for residual systematics, that is,

$$\delta P(k,\mu,z) \ = 0.2\left[P_{\rm Taka}(k,\mu,z) - P_{\rm Mead}(k,\mu,z)\right], \tag{4.16}$$

which we show in Fig. 4.5 as a fraction of our fiducial power spectrum.

We use the Fisher matrix formalism to predict the bias that the residuals in Eq. (4.16) will produce in cosmological parameters [290, 291]. In the limit where changes to best-

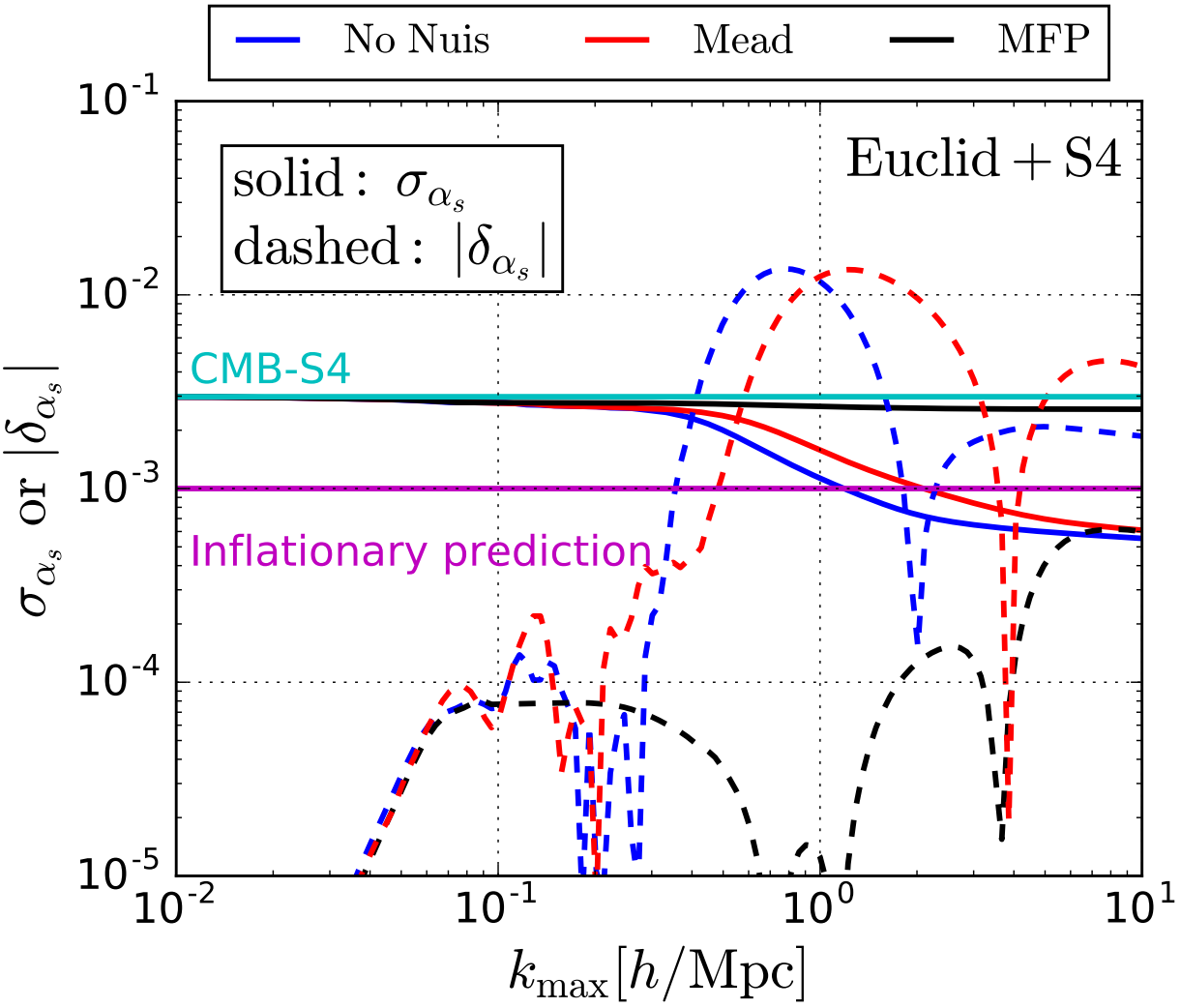

Figure 4.6: 1-σ statistical errors (solid curves) and bias (dashed) in the first spectral running, as a function of $k_{\rm max}$. We adopt the Euclid+CMB-S4 combination of surveys. The legend on top denotes three alternate assumptions about the systematic error modeling: none (blue), Mead (red), and MFP (black).

fit parameters can be expanded linearly in small changes to the observations, the bias in parameter p_i can be written as

$$\delta p_i \approx \sum_j (F^{-1})_{ij} G_j, \tag{4.17}$$

where

$$G_j = \sum_{z,\mu,k_\alpha,k_\beta} d\mu \frac{\partial P(k_\alpha,\mu,z)}{\partial p_j} \left[\mathrm{Cov}^{-1}\right]_{k_\alpha,k_\beta} \delta P(k_\beta,\mu,z),$$

and Cov is the same covariance matrix defined in Eq. (4.9). This formula is only accurate when the biases are small compared to the forecasted errors — that is, $|\delta p_i| \ll \sigma_{p_i} = \sqrt{(F^{-1})_{ii}}$ — so we use it to determine the approximate threshold at which the bias on p_i becomes unacceptably large.

We plot both the bias $|\delta_{\alpha_s}|$ (dashed) and marginalized uncertainty σ_{α_s} (solid) for Euclid+CMB−S4 constraints on α_s in Fig. 4.6. The value of $k_{\rm max}$ where the bias and uncertainty become comparable tells us roughly the smallest scales that can be in included in an analysis without the systematic effects in δP adversely biasing the results for α. We see that though the MFP nuisance parameter prescription (black) has weaker constraints than the No Nuisance and Mead cases, it also is significantly better at protecting against bias. That is to say, on all $k_{\rm max}$ scales we examined, the bias in α for the MFP case is well below its statistical uncertainty. In contrast, the No Nuisance and Mead prescriptions have $\delta_{\alpha_s} \approx \sigma_{\alpha_s}$ at $k_{\rm max} \approx 0.4h\,\mathrm{Mpc}^{-1}$ and $k_{\rm max} \approx 0.6h\,\mathrm{Mpc}^{-1}$ respectively. Comparing the value for σ_{α_s} at these $k_{\rm max}$ values, we see that if we restrict ourselves to scales with $\delta_{\alpha_s} < \sigma_{\alpha_s}$, the improvement from adding high-k LSS data is marginal for all three $M_{\rm nl}(k,z)$ treatments.

To confirm that these results are robust against changes to the shape of our residual function $\delta P(k,z)$, we compared the same bias projections for a variety of other $\delta P_{i,j}(k,z) \propto P_i(k,z) - P_j(k,z)$, where $i,j \in \{$Mead [28], Takahashi [6], Bird [292], Peacock[2], Halomodel [275]$\}$ runs over a subset of possible prescriptions for the nonlinear matter power spectrum in `CAMB`. We normalized these so that the relative difference $\delta P_{i,j}(k,z)/P_{\rm Mead}$ had the same RMS as our fiducial case[3] (see Appendix C.2). Thus the fiducial $\delta P(k,\mu,z)$ given in Eq. 4.16 and the magnitude of resulting biases derived therefrom should be fairly representative of possible errors in modeling $P(k,z)$, while also aligning with the oft-quoted baseline assumption that uncertainties have to be controlled to 1% or better in order to not degrade the accuracy of future cosmological measurements of dark energy (e.g. [293]).

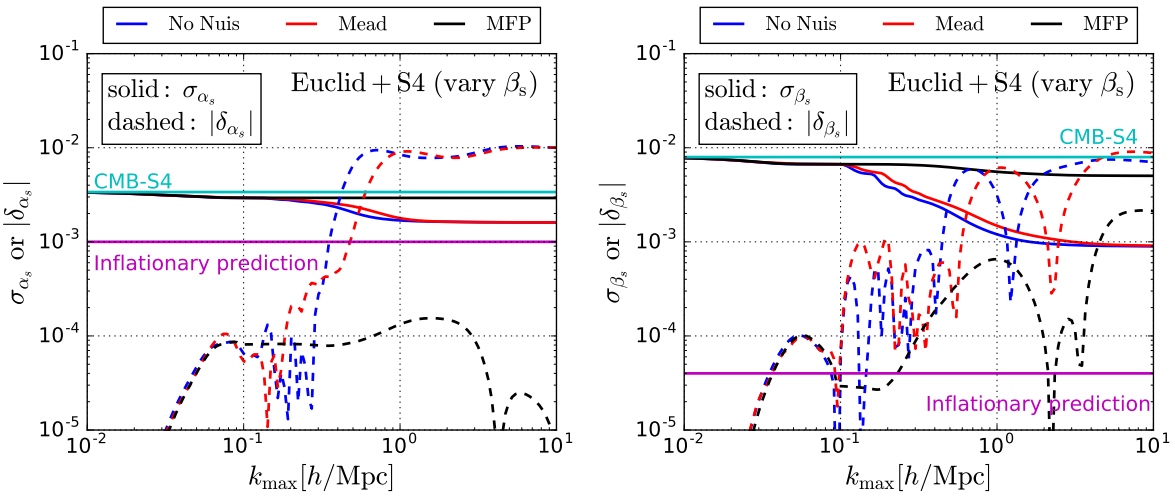

Figure 4.7: Similar to Fig. 4.6, except now β_s is allowed to vary. The left panel shows the 1-σ error and parameter bias in α_s as a function of $k_{\rm max}$, while the right panel shows the same for β_s. The curves have the same meaning as in Fig. 4.6.

4.5 Constraining the Second Running: ΛCDM $+\alpha_s+\beta_s$

We now expand the cosmological parameter space to include the second running β_s—that is, we extend our expansion of the spectral index to second order in $\ln k$. This is a parameter for which constraints from LSS data have the potential to be particularly interesting. Recent Planck results have suggested a positive second running β_s at nearly 2σ confidence which, if it persists, will help to discriminate between inflationary models [242, 245]. Additionally, as mentioned in Sec. 4.1, the current best-fit of $\beta_s = 0.025\pm0.013$ has important implications for physics of the late universe, as it makes primordial black holes a viable dark matter candidate (albeit with the requirement of a negative third-order running to avoid overproduction [251]).

The left panel of Fig. 4.7 shows that, when β_s is allowed to vary, combined constraints from CMB-S4 and LSS are no longer able to reach the inflationary prediction for α_s at any $k_{\rm max} < 10h\,{\rm Mpc}^{-1}$, even when the non-linear $P(k,z)$ is modeled perfectly and with no nuisance parameters (solid blue curve). On the other hand, the right panel of Fig. 4.7 shows that β_s itself benefits greatly from the addition of the LSS data. While CMB-S4 is expected to improve constraints on β_s by a factor of $\sim$4 over current levels, our results indicate that LSS data in the nonlinear regime from Euclid or DESI has the potential to improve this significantly up to $k_{\rm max} \sim 2h\,{\rm Mpc}^{-1}$, at which point shot noise limits the information that can be gained.

[2]http://www.roe.ac.uk/ jap/haloes

[3]For $k > 0.005h\,{\rm Mpc}^{-1}$, corresponding to the minimum k for which `CAMB` calculates nonlinear modifications to the power spectrum.

We next consider the systematic biases in β_s using the same prescription as in Sec. 4.4. Using our fiducial model for power spectrum residuals due to unaccounted-for systematics [Eq. (4.16)], Fig. 4.7 shows that, without introducing undue bias, adding data from a Euclid-like survey can improve constraints on β_s by a factor of 3–4 compared to the CMB-S4 only case.[4] While still an order of magnitude too large to reach β_s predicted by standard single-field slow-roll inflation, this level of precision is in the regime necessary to test for models relevant for PBH formation [27, 251, 294].

4.6 Conclusions

In this work, we have investigated how small-scale information from large-scale structure surveys can improve constraints on the first [α_s] and second [β_s] runnings of the scalar spectral index [n_s]. Previous analyses have been limited to the linear regime where the matter power spectrum is accurately described by theory, but the possibility of extending analyses to nonlinear regimes in the future is attractive. This is for two reasons: First, there are many more modes at small scales and hence statistical errors from cosmic variance are greatly reduced. Second, accessing high k values provides a longer lever arm when combined with CMB constraints, which increases the sensitivity to variations in the spectral index and its runnings.

Attempts to include small-scale information are limited by challenges associated with theoretical modeling of the nonlinear power spectrum. Nonlinear clustering of dark matter, baryonic effects, and scale-dependent galaxy bias all contribute to modeling uncertainties on small scales. Therefore, it is critical to not only calibrate models for these effects as accurately as possible, but also to carefully characterize how analyses' cosmological results are affected by residual errors in predictions for small-scale power.

Motivated by these considerations, we compare forecasted constraints for spectral runnings from a few different parameterizations intended to capture the effects of systematics in the nonlinear regime. Specifically, we study cases where small scales are modeled using the fiducial halo model code (No Nuisance case), the parameterization from Mead et al [7] which introduces two nuisance parameters (Mead case), and an agnostic treatment adapted from Bielefeld et al [279] with up to a few hundred parameters, depending on $k_{\rm max}$ (Many Free Parameters, or MFP case).

[4] This was the one case where our fiducial $\delta P(k,\mu,z)$ differed somewhat in its bias forecast from the ensemble of other $\delta P(k,\mu,z)$ tested, with $\delta_\beta/\sigma_\beta = 1$ occurring at $k_{\rm max} \approx 0.5$ and $0.7h\,{\rm Mpc}^{-1}$ for the No Nuisance and Mead models, respectively ($\sim 4\times$ improvement in σ_β), compared to $k_{\rm max} \approx 0.2$ and $0.4h\,{\rm Mpc}^{-1}$ for the typical $\delta P(k,\mu,z)$ ($\sim 3\times$ improvement in σ_β). The results are still qualitatively similar, however.

We first study the forecasts for statistical errors on the first spectral running α_s for future LSS surveys like Euclid and DESI alone, as well as in combination with CMB-S4. We find that in the No Nuisance and Mead cases, the constraints from large-scale structure surveys tighten substantially as $k_{\rm max}$ is raised to include nonlinear scales. The MFP case also shows improvement, but with a flatter dependence on $k_{\rm max}$ and weaker constraints overall. It is also at nonlinear scales where constraints using LSS and CMB data begin to improve α_s constraints compared to CMB-S4 data alone. The tightest constraints come from the Euclid+DESI+CMB-S4 combination, for which our No Nuisance forecasts for statistical errors reach a value about a third of the α_s predicted by single-field slow-roll inflation at $k_{\rm max} \gtrsim 3h\,{\rm Mpc}^{-1}$. This could be precise enough to achieve a $\sim 3\sigma$ detection. These results become less promising, however, when we investigate the extent to which mismodeling of the nonlinear power spectrum biases cosmological parameter estimation. Using the difference between two commonly used nonlinear prescriptions as an example of expected modeling uncertainties, we determine the highest $k_{\rm max}$ we can use in an analysis before the resulting systematic bias in α_s becomes comparable to its statistical errors.

We find that for 1% errors in the power spectrum, in the No Nuisance case both α_s and β_s remain unbiased (i.e. bias is smaller than the 1σ statistical error) up to $k_{\rm max} \approx 0.3 - 0.4h\,{\rm Mpc}^{-1}$. Including these smaller scales results in significant improvements in σ_β, but only marginal improvements in σ_α. Adopting the Mead parametrization of the systematics leads to very similar results indicating that the two free parameters from Mead et al [28], motivated to account for baryonic feedback, are not sufficient to offer protection against the 1%-level residual small-scale systematics in the power spectrum we might expect to encounter. In contrast, for the MFP parametrization α_s and β_s are unbiased for all $k_{\rm max}$ studied, but the statistical error on the runnings in the CMB+LSS scenarios is only marginally better than that of CMB-S4 alone.

Our level of optimism regarding future measurements of the spectral runnings using LSS data is therefore mixed. The values of α_s and β_s predicted by standard, single-field slow-roll models of inflation seem out of reach even when CMB-S4 information is combined with that of most powerful future LSS surveys once the small-scale systematics in the galaxy surveys are taken into account. To improve upon this, it will be necessary to more precisely characterize small-scale systematics such as through a suite of high-quality hydrodynamical simulations that adequately spanned the range of possibilities, which can then be sampled over (e.g. via emulation). If the range in uncertainty were found to be largely orthogonal to the cosmological parameters of interest, then a detection of the inflationary α_s may be possible even at the level of 1% uncertainty, e.g. via marginalizing over the principle modes of uncertainty as is done for baryons in Ref. [295]. In either case, larger values of spectral

runnings predicted by other classes of inflationary models, as well as those motivated by other physics (e.g. primordial black holes) are within reach, and should be testable with the next generation of surveys.

Chapter 5

Closing Remarks

The large-scale structure of the Universe is a key cosmological probe for understanding the nature of dark energy and testing ΛCDM as it is sensitive to both the late-time expansion history and the rate of the growth of structure. The next generation of LSS surveys will see dramatic increases in both the volumes probed and the number of galaxies observed, resulting in extremely powerful tests of the current cosmological paradigm. These gains in statistical precision create heightened risk of unmodeled systematic errors that could bias cosmological inference, necessitating strategies to identify and mitigate such sources of bias. This thesis presents three different investigations into sources of systematic error for analyses using large-scale structure data.

In Chapter 2, we investigated how residual calibration errors in the measured LSS density field propagate into reconstructions of the Integrated Sachs-Wolfe effect. The ISW effect is sourced from the decay of gravitational potentials in the presence of dark energy and so is sensitive to dark energy and modifications to General Relativity. As the ISW is the dominant late-time contribution to the cosmic microwave background anisotropies, being able to reconstruct the ISW map offers the tantalizing possibility of separating the early- from the late-time contributions to the CMB. Previous studies have attempted to do exactly that, but recent work demonstrated that photometric calibration errors in modern LSS surveys are a limiting systematic. Using simulated analyses, we investigated how the fidelity of the ISW map reconstruction depends on the level of photometric calibration achieved for two upcoming Stage IV LSS surveys, SPHEREx and Euclid. We found that both including tomographic information from a single survey and using data from multiple, complementary galaxy surveys improve the reconstruction by mitigating the impact of spurious power contributions from residual calibration errors. We also found that failing to account for spurious power from calibration errors in the ISW estimator significantly degraded results, with the potential to actually *worsen* estimates when adding information from LSS surveys to in-

formation from just CMB temperature anisotropies. We proposed a modified estimator to address this, and found that if the photometric calibration errors in galaxy surveys can be independently controlled at the level required to obtain unbiased dark energy constraints, then it will be possible to reconstruct ISW maps with excellent accuracy using a combination of maps from Euclid and SPHEREx.

In Chapter 3, we turned our attention to the primary methods to calibrate galaxy samples in LSS surveys through detecting and mitigating systematic contamination from a large variety of astrophysical and observational effects. We compared several existing methods for removing such systematics from galaxy clustering measurements. We showed how all the methods, including the popular pseudo-C_ℓ Mode Projection and Template Subtraction methods, can be interpreted under a common regression framework, and used this to suggest improved approaches. We showed how methods designed to mitigate systematics in the power spectrum can be used to produce clean maps, which are necessary for cosmological analyses beyond the power spectrum and other map-based analyses like the one studied in Chapter 2. We extended current methods to treat the next-order multiplicative contamination in observed maps and power spectra, which reduced power spectrum errors from $\Delta\chi^2_{\mathrm{C}_\ell} \simeq 10$ to $\simeq 1$ in simulated analyses. Two new mitigation methods were proposed, which incorporate desirable features of current state-of-the-art methods while being simpler to implement. Investigating the performance of all the methods on a common set of simulated measurements from Year 5 of the Dark Energy Survey, we tested their robustness to various analysis cases. Our proposed methods produced improved maps and power spectra when compared to current methods, while requiring almost no user tuning. Finally, we made recommendations for systematics mitigation in future surveys, noting that the methods presented are generally applicable beyond the galaxy distribution to any field with spatial systematics.

Our final investigation in Chapter 4 pivoted from spatially-dependent observational systematics to studying the effect of general modeling uncertainties at small scales on the inference of cosmological parameters carrying crucial information about inflation. Specifically, we studied the feasibility of measuring the spectral runnings α_s and β_s using data from simulated surveys similar to the upcoming LSS surveys DESI and Euclid, in combination with data from a Stage IV CMB experiment. Since these measurements will be sensitive to modeling uncertainties for the nonlinear high-k power spectrum, we examined how three different ways of parameterizing those systematics—introducing zero, two, or several hundred nuisance parameters—affect constraints and protect against parameter biases. Considering statistical errors alone, we found that including strongly nonlinear scales can substantially tighten constraints. However, these constraints weakened to levels not much better than

those from a CMB-S4 experiment alone when we limited our analysis to LSS scales that are large enough that estimates were not strongly affected by systematic biases. Given these considerations, near-future large-scale structure surveys are unlikely to add much information to the CMB-S4 measurement of the first running α_s. We found that there is more potential for improvement for the second running, β_s, for which large-scale structure information will allow constraints to be improved by a factor of $3-4$ relative to using the CMB alone. We concluded that while these constraints are still above the value predicted by slow-roll inflation, they do probe regions of parameter space relevant to nonstandard inflationary models with large runnings, for example those that can generate an appreciable abundance of primordial black holes.

These studies contribute to the vital work of detecting, assessing, and mitigating the impact of unmodeled systematic errors in analyses using upcoming large-scale structure surveys. This is critical both to avoid potential biases in future results as well as to provide guideposts for future model development. The last several years have seen a flurry of proposals for new cosmological models as cracks in the standard ΛCDM paradigm have grown wider. Most notable of these cracks is the so-called Hubble tension, in which early Universe measurements infer an H_0 under ΛCDM that differs from model-agnostic late-time measurements by $4-6\sigma$. These discrepancies have withstood a wide battery of independent tests and grown more significant over time. The failure to as yet identify any astrophysical or observational systematic errors that can explain the discrepancy has led to a growing focus on novel cosmological models to remedy the disagreement. This in turn has spurred a rapid growth in the development and application of statistical techniques for model comparison in a cosmological context, a trend that looks likely to continue. Unfortunately, no beyond-ΛCDM cosmological model proposed thus far has provided a decisive resolution of the H_0 problem (see e.g. [296, 297] for recent reviews), but the increasing focus on radical new models speaks to a broader weakening of the implicit priors that most cosmologists have for the standard ΛCDM paradigm, a shift that has only occurred as more sources of systematic error are ruled out.

Another outstanding but less prominent discrepancy concerns the ISW effect. Several studies have reaffirmed discrepancies between the predictions of ΛCDM and the observed strength of the ISW effect as measured on stacked voids (e.g. [76, 81, 298, 299]). However more traditional projected cross-correlation analyses of LSS with CMB that include less extreme over- and under-densities do not find substantial disagreement with ΛCDM (e.g. [26, 300, 301]). While the ISW measurement discrepancy has received considerably less attention than the growing H_0 tension, it offers a tantalizing and complementary means of probing physics beyond ΛCDM, in particular because of the ISW's sensitivity to changes in

the physics of dark energy. It was recently shown by Ref. [299] that the novel (albeit radical) AvERA model of Ref. [302] could in theory simultaneously resolve the Hubble tension and the abnormally strong signature of the CMB cold spot. Because this model has no dark energy, it leaves unique signatures in the ISW that can serve as a 'smoking gun', such as a negative correlation between matter and the CMB at redshifts of $1.5 \lesssim z \lesssim 4.4$, a range that should be accessible by both the Dark Energy Survey and the Dark Energy Spectroscopic Instrument making it ripe for near-future experimental tests [303], and Chapter 2 contains much of the framework that would be necessary for testing the significance of such tomographic measurements under a baseline ΛCDM cosmology. From the measurement side, the tools from Chapter 3 will be crucial for applying the cleaning necessary for samples at such high redshift where systematic contamination is large, lest the fog in our lenses be mistaken for the smoking gun, so to speak.

As the field continues to grow and we gain access to improved measurements and new datasets, inconsistencies inevitably arise and must be reconciled. The Hubble tension and abnormal ISW void measurements join other peculiarities (such as an anomalously low lensing signal [304]) that, having withstood easy systematic explanation, have led to a greater questioning of the standard cosmological model itself and primed us for the next revolution in our cosmological understanding.

It is important to note, however, that the field is increasingly dominated by large collaborations which function to stitch together the products of many different teams working with different data products at diverse stages in the analysis pipeline. This helps to drive the enormous progress the field has made, but also renders analyses vulnerable, as each step is built on the scaffolding of others' decisions, allowing potential errors to have far-reaching impacts but without the commensurate visibility. It is difficult if not impossible for any one researcher to understand and vet all aspects of the analysis, let independently replicate it. As the field moves forward, it will be increasingly important to account for *researchers* as a source of systematic error; to characterize and report uncertainties in the choices that were made throughout the analysis pipeline and develop streamlined methods to marginalize over and account for this human element.

The fundamental goal of modern cosmology is to understand the Universe we inhabit at its most basic level. Unlike many fields, our ability to manipulate the conditions of our experiments is extremely limited, forcing us to rely on ever-improving (but passive) observations, against which we compare carefully-constructed theoretical models. We have gained an inordinate amount of knowledge over the last 40 years through the iterative process of testing, refining, and extending our models with each new wave of hard-won data. However, upcoming surveys like Euclid, SPHEREx, DESI, LSST and Roman Space Telescope will provide

high-resolution measurements of the large-scale structure with up to 20 billion galaxies over large fractions of the sky and beating down statistical errors. If we can adequately control the myriad systematic effects that loom large, then these surveys will test our understanding of cosmology like never before and truly herald the age of Precision Cosmology. Through such a dance between theory and observation, carefully choreographed by the quantifiable unknown, the Universe grows its self-awareness.

Appendix A

ISW Reconstruction Appendices

A.1 Equivalence with Other Estimators

Here we demonstrate the equivalence between our estimator for the ISW map coefficients $\hat{a}_{\ell m}^{\rm ISW}$ in Chapter 2, and the estimators proposed by Manzotti et al [80] (MD14) and Barreiro et al [75] (B08).

Our estimator in Eq. (2.7) is based on a version of the likelihood from MD14 that has been reformulated to handle observed CMB maps like any other input map. MD14 derive their estimator using the likelihood

$$\mathcal{L}(T^{\rm ISW}) \propto \frac{1}{\sqrt{\det(CD)}} \times \exp\left\{-\frac{1}{2}\boldsymbol{d}^T D^{-1}\boldsymbol{d}\right\} \times \exp\left\{-\frac{1}{2}\left(T^{\rm obs}-T^{\rm ISW}\right)C^{-1}\left(T^{\rm obs}-T^{\rm ISW}\right)\right\}, \tag{A.1}$$

where $C \equiv C^p + C^n$ is the angular power spectrum of the primordial C^p and noise C^n contributions to CMB temperature fluctuations, $\boldsymbol{d}$ is a vector of ISW and LSS tracer maps, and D is the covariance matrix between the ISW and LSS tracers (see [80] Eqs. (4-6)), with ISW maps associated with the first (1) index. This likelihood is a product of the independent likelihoods for $(T^{\rm obs} - T^{\rm ISW})$ and for the input maps in d.

Instead of explicitly including independent terms for the primordial CMB and LSS tracers (which are assumed to have no cross-correlation), we include the total observed CMB temperature,

$$T^{\rm obs} = T^p + T^{\rm ISW}, \tag{A.2}$$

where T^p includes both the primordial CMB temperature as well as any instrumental noise

terms. We then expand the data vector to include $T^{\rm obs}$:

$$\mathbf{d}_{\ell m} = (a^{\rm ISW}_{\ell m}, g^1_{\ell m}, \ldots, g^n_{\ell m}) \rightarrow (a^{\rm ISW}_{\ell m}, g^1_{\ell m}, \ldots, g^n_{\ell m}, a^{\rm obs}_{\ell m}),$$

with $a_{\ell m}$ indicating spherical components of ISW and CMB temperature fluctuations and $g_{\ell m}$ indicating components of LSS overdensity. The covariance matrix is similarly expanded to account for the cross-correlation of $T^{\rm obs}$ with the ISW and LSS tracers

$$D_\ell \rightarrow \begin{pmatrix} C_\ell^{\rm ISW,ISW} & C_\ell^{\rm ISW,1} & \cdots & C_\ell^{\rm ISW,n} & C_\ell^{\rm ISW,obs} \\ C_\ell^{\rm 1,ISW} & C_\ell^{\rm 1,1} & \cdots & C_\ell^{\rm 1,n} & C_\ell^{\rm 1,obs} \\ \vdots & \vdots & \ddots & \vdots & \\ C_\ell^{\rm n,ISW} & C_\ell^{\rm n,1} & \cdots & C_\ell^{\rm n,n} & C_\ell^{\rm n,obs} \\ C_\ell^{\rm obs,ISW} & C_\ell^{\rm 1,obs} & \cdots & C_\ell^{\rm n,obs} & C_\ell^{\rm obs,obs} \end{pmatrix}. \tag{A.3}$$

Assuming that at the scales we consider the observed CMB is cross-correlated with other LSS tracers only through the ISW, we have

$$\begin{aligned} C_\ell^{\rm obs,ISW} &= C_\ell^{\rm ISW,ISW} \\ C_\ell^{\rm obs,LSS_i} &= C_\ell^{\rm ISW,LSS_i}, \\ C_\ell^{\rm obs,obs} &= C_\ell^{p,p} + C_\ell^{\rm ISW,ISW}, \end{aligned} \tag{A.4}$$

assuming there is no residual cross-correlation between the primordial and late-time CMB. Maximizing the resulting likelihood

$$\mathcal{L}(T^{\rm ISW}) \propto \frac{1}{\sqrt{\det(D)}} \exp\left\{-\frac{1}{2}\boldsymbol{d}^T D^{-1} \boldsymbol{d}\right\}, \tag{A.5}$$

gives the optimal estimator given in Sec. 2.2.2.

To show that this is equivalent to the estimator derived from Eq. (A.1), we focus on the case of using CMB temperature and a single LSS tracer as input maps. For compactness, and to make the connections with other ISW estimators in the literature more apparent, we adopt notation from B08, where s, g, and T indicate the ISW, LSS tracer, and observed CMB temperature signals, respectively. We then have

$$\mathbf{d}_{\ell m} = (a^s_{\ell m}, g^g_{\ell m}, a^T_{\ell m}), \tag{A.6}$$

and

$$D_\ell = \begin{pmatrix} C_\ell^{ss} & C_\ell^{sg} & C_\ell^{ss} \\ C_\ell^{sg} & C_\ell^{gg} & C_\ell^{sg} \\ C_\ell^{ss} & C_\ell^{sg} & C_\ell^{tt} \end{pmatrix}. \tag{A.7}$$

From Eqs. (2.7) and (2.8) our estimator gives

$$\begin{aligned} \hat{a}^s_{\ell m} &= \frac{-1}{[D_\ell^{-1}]_{11}} \left([D_\ell^{-1}]_{12}\ g^g_{\ell m} + [D_\ell^{-1}]_{13}\ a^T_{\ell m} \right) \\ &= \left(\frac{C_\ell^{sg}(C_\ell^{tt} - C_\ell^{ss})}{C_\ell^{gg} C_\ell^{tt} - (C_\ell^{sg})^2} \right) g^g_{\ell m} + \left(\frac{C_\ell^{ss} C_\ell^{gg} - (C_\ell^{sg})^2}{C_\ell^{gg} C_\ell^{tt} - (C_\ell^{sg})^2} \right) a^T_{\ell m}. \end{aligned} \tag{A.8}$$

We now calculate the estimator of MD14. Denoting their covariance matrix without the CMB as D'_ℓ, we use their Eq. (9)

$$\begin{aligned} \hat{a}^{s,\mathrm{MD}}_{\ell m} &= \left([C_\ell^{pp}]^{-1} + [D'^{-1}]_{11} \right)^{-1} \left(-[D'^{-1}_\ell]_{12}\ g^g_{\ell m} + [C_\ell^{pp}]^{-1} a^T_{\ell m} \right) \\ &= \left(\frac{1}{C_\ell^{pp}} + \frac{C_\ell^{gg}}{\det|D'_\ell|} \right)^{-1} \left[\frac{C_\ell^{sg}}{\det|D'_\ell|} g^g_{\ell m} + \frac{1}{C_\ell^{pp}} a^T_{\ell m} \right] \\ &= \left(\frac{1}{\det|D'_\ell| + C_\ell^{gg} C_\ell^{pp}} \right) \left[C_\ell^{pp} C_\ell^{sg} g^g_{\ell m} + (\det|D'_\ell|) a^T_{\ell m} \right] \end{aligned}$$

Expanding the determinant, we find

$$\hat{a}^{s,\mathrm{MD}}_{\ell m} = \left(\frac{C_\ell^{sg} C_\ell^{pp}}{C_\ell^{gg}(C_\ell^{ss} + C_\ell^{pp}) - (C_\ell^{sg})^2} \right) g^g_{\ell m} + \left(\frac{C_\ell^{ss} C_\ell^{gg} - (C_\ell^{sg})^2}{C_\ell^{gg}(C_\ell^{ss} + C_\ell^{pp}) - (C_\ell^{sg})^2} \right) a^T_{\ell m}, \tag{A.9}$$

which, using the relation $C_\ell^{tt} = C_\ell^{pp} + C_\ell^{ss}$, is equivalent to the estimator given by Eq. (A.8).

This is also equivalent to the estimator proposed in B08, which uses the Cholesky decomposition (L) of the covariance matrix D''_ℓ (denoted C_ℓ in Ref. [75]).

$$D''_\ell = \begin{bmatrix} C_\ell^{gg} & C_\ell^{sg} \\ C_\ell^{sg} & C_\ell^{ss} \end{bmatrix} = L_\ell L_\ell^T. \tag{A.10}$$

Note that here the ISW index is last instead of first, in contrast to the covariance matrices

D_ℓ and D'_ℓ used previously. The Cholesky decomposition is written as

$$L_\ell = \begin{bmatrix} \sqrt{C_\ell^{gg}} & 0 \\ \dfrac{C_\ell^{sg}}{\sqrt{C_\ell^{gg}}} & \sqrt{C_\ell^{ss} - \dfrac{(C_\ell^{sg})^2}{C_\ell^{gg}}} \end{bmatrix}. \tag{A.11}$$

Eqs. (8) and (9) from Ref. [75] give the ISW estimate (their $\hat{s}_{\ell m}$) as

$$\hat{a}_{\ell m}^{s,\mathrm{B08}} = \frac{L_{12}}{L_{11}} \left(1 - \frac{L_{22}^2}{L_{22}^2 + C_\ell^{pp}}\right) g_{\ell m}^g + \frac{L_{22}^2}{L_{22}^2 + C_\ell^{pp}} a_{\ell m}^T$$

where we have suppressed the ℓ-dependence of L and combined their observed ISW signal ($s_{\ell m}$) and noise ($n_{\ell m}$) terms into the single term $a_{\ell m}^T$. We use C_ℓ^{pp} to denote the combined power of noise and the primordial CMB, in keeping with the notation above. Plugging this into Eq. (A.11), we obtain

$$\begin{aligned} \hat{a}_{\ell m}^{s,\mathrm{B08}} &= \frac{C_\ell^{sg}}{C_\ell^{gg}} \left(\frac{C_\ell^{pp}}{(C_\ell^{ss} + C_\ell^{pp}) - (C_\ell^{sg})^2/C_\ell^{gg}}\right) g_{\ell m}^g + \frac{C_\ell^{ss} - (C_\ell^{sg})^2/C_\ell^{gg}}{(C_\ell^{ss} + C_\ell^{pp}) - (C_\ell^{sg})^2/C_\ell^{gg}} a_{\ell m}^T \\ &= \left(\frac{C_\ell^{sg}(C_\ell^{tt} - C_\ell^{ss})}{C_\ell^{gg} C_\ell^{tt} - (C_\ell^{sg})^2}\right) g_{\ell m}^g + \left(\frac{C_\ell^{ss} C_\ell^{gg} - (C_\ell^{sg})^2}{C_\ell^{gg} C_\ell^{tt} - (C_\ell^{sg})^2}\right) a_{\ell m}^T, \end{aligned}$$

which is the same as Eq. (A.8).

A.2 Estimating ρ with $R_\ell(\tilde{C}_\ell)$

Here we show why using raw pseudo-C_ℓ's ($\tilde{C}_\ell$) in the estimator results in a degraded reconstruction, for which $\bar{\rho}$ is not well approximated by $\hat{\rho}$ (Eq. (2.17)).

For a given realization, ρ is constructed from the covariance between the true and reconstructed ISW maps ($\mathrm{Cov}(T^{\mathrm{ISW}}, T^{\mathrm{rec}})$, i.e. the numerator in Eq. (2.13)) normalized by the square root of the individual variances of the true and reconstructed ISW maps (σ^2_{True} and σ^2_{rec}, respectively). We therefore focus on how using realization-specific $\tilde{C}_\ell$'s in the estimator filter R_ℓ affects the individual C_ℓ contributions to σ^2_{rec} and $\mathrm{Cov}(T^{\mathrm{ISW}}, T^{\mathrm{rec}})$ (σ^2_{ISW} is unaffected by our choice of R_ℓ). For simplicity, we work with a single input map.

If the ISW estimator filter R_ℓ is constructed from analytically computed model C_ℓ's, the

power spectrum of the ISW map for a given realization will be

$$\begin{aligned}\tilde{C}_\ell^{\,\mathrm{rec-rec,th}} &= R_\ell^2(C_\ell^{\mathrm{th}})\tilde{C}_\ell^{\,\mathrm{gal-gal}}, \\ &= \left(\frac{C_\ell^{\mathrm{ISW-gal}}}{C_\ell^{\mathrm{gal-gal}}}\right)^2 \tilde{C}_\ell^{\,\mathrm{gal-gal}}.\end{aligned} \tag{A.12}$$

We add the superscript "th" to distinguish this reconstructed ISW power spectrum from the one where the filter R_ℓ is built from $\tilde{C}_\ell$'s, which will be discussed shortly. The expectation value for this over many realizations is

$$\begin{aligned}\left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec,th}} \right\rangle &= \left(\frac{C_\ell^{\mathrm{ISW-gal}}}{C_\ell^{\mathrm{gal-gal}}}\right)^2 \left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle, \\ &= \frac{\left(C_\ell^{\mathrm{ISW-gal}}\right)^2}{C_\ell^{\mathrm{gal-gal}}}.\end{aligned} \tag{A.13}$$

Now let us look at the behavior of the reconstructed ISW power when the galaxy autopower spectra in the estimator filter are extracted from the observed maps. Denoting this version of the filter by

$$\tilde{R}_\ell \equiv \frac{C_\ell^{\mathrm{ISW-gal}}}{\tilde{C}_\ell^{\,\mathrm{gal-gal}}}, \tag{A.14}$$

we write

$$\begin{aligned}\tilde{C}_\ell^{\,\mathrm{rec-rec}} &= \tilde{R}_\ell^{\,2}\tilde{C}_\ell^{\,\mathrm{gal-gal}}, \\ &= \left(C_\ell^{\mathrm{ISW-gal}}\right)^2 \left(\frac{1}{\tilde{C}_\ell^{\,\mathrm{gal-gal}}}\right).\end{aligned} \tag{A.15}$$

Because the measured $\tilde{C}_\ell^{\,\mathrm{gal-gal}}$ appears in the denominator of this expression, taking its expectation value over many realizations is somewhat more complicated. To do so we use the fact that $(2\ell+1)\tilde{C}_\ell^{\,\mathrm{gal-gal}}$ is χ^2-distributed with $2\ell+1$ degrees of freedom. This means $\tilde{C}_\ell^{\,\mathrm{rec-rec}}/(2\ell+1)$ follows an inverse-χ^2 distribution, with an expectation value[1]

$$\left\langle \frac{\tilde{C}_\ell^{\,\mathrm{rec-rec}}}{2\ell+1} \right\rangle = \frac{1}{2\ell-1}\frac{\left(C_\ell^{\mathrm{ISW-gal}}\right)^2}{C_\ell^{\mathrm{gal-gal}}}. \tag{A.16}$$

[1]We refer the reader to Refs. [305] and [306] for discussions of the bias introduced when inverting an estimator, with implications specifically for estimating the inverse covariance matrix.

Therefore the average reconstructed power is

$$\begin{aligned}\left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec}} \right\rangle &= \frac{2\ell+1}{2\ell-1} \left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec,th}} \right\rangle \\ &= \left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec,th}} \right\rangle \left(1 + \frac{2}{2\ell-1}\right).\end{aligned} \tag{A.17}$$

Because $\langle \tilde{C}_\ell^{\,\mathrm{rec-rec}} \rangle$ is strictly positive, this increased power results in an increase in the total variance of the reconstruction map $\tilde{\sigma}^2_{\mathrm{rec}}$ compared to that from the theory-only filter reconstruction $\sigma^2_{\mathrm{rec,th}}$

$$\begin{aligned}\langle \tilde{\sigma}^2_{\mathrm{rec}} \rangle &= \frac{1}{4\pi} \sum_\ell (2\ell+1) \left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec}} \right\rangle \\ &= \sigma^2_{\mathrm{rec,th}} + \frac{1}{4\pi} \sum_\ell (2\ell+1) \frac{\left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec,th}} \right\rangle}{\ell - 1/2},\end{aligned} \tag{A.18}$$

In contrast, we find the average cross-power $\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec}} \rangle$ between reconstructed and true ISW maps remains unchanged. The increased power of the reconstruction thus results in a net decrease in $\langle \rho \rangle$, per Eq. (2.13), and hence is not well approximated by simply substituting the theory C_ℓ, as is done to compute $\hat{\rho}$. Additionally, this suggests that a simple scaling of R_ℓ in order to "debias" the reconstruction will not improve ρ.

To understand why the cross-power does not increase, we again use the observed galaxy autopower in the estimator and approximate its expectation value. The cross-power is given by

$$\begin{aligned}\tilde{C}_\ell^{\,\mathrm{ISW-rec}} &= \tilde{R}_\ell \tilde{C}_\ell^{\,\mathrm{ISW-gal}} \\ &= \left(\frac{C_\ell^{\mathrm{ISW-gal}}}{\tilde{C}_\ell^{\,\mathrm{gal-gal}}} \right) \tilde{C}_\ell^{\,\mathrm{ISW-gal}}.\end{aligned} \tag{A.19}$$

Here we have a quotient of two *non*-independent χ^2 random variables. Generically, we can approximate the average of a function of two random variables X and Y through a second-order Taylor expansion about the mean of each (μ_X, μ_Y):

$$\begin{aligned}\langle f(X,Y) \rangle \approx\ & f(\mu_X, \mu_Y) + \frac{1}{2} f''_{XX}(\mu_X, \mu_Y) \left\langle (X-\mu_X)^2 \right\rangle \\ & + f''_{XY}(\mu_X, \mu_Y) \left\langle (X-\mu_X)(Y-\mu_Y) \right\rangle \\ & + \frac{1}{2} f''_{YY}(\mu_X, \mu_Y) \left\langle (Y-\mu_Y)^2 \right\rangle,\end{aligned}$$

where a prime indicates a derivative with respect to the respective subscripted variable. By

taking $f(X,Y)$ to be $\tilde{C}_\ell^{\,\mathrm{ISW-rec}}$, of the form X/Y, then from Eq. (A.19) we can approximate the mean cross-power to be

$$\begin{aligned}\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec}} \right\rangle &\approx C_\ell^{\mathrm{ISW-gal}} \frac{\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-gal}} \right\rangle}{\left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle} \left(1 - \frac{\mathrm{Cov}(\tilde{C}_\ell^{\,\mathrm{ISW-gal}}, \tilde{C}_\ell^{\,\mathrm{gal-gal}})}{\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-gal}} \right\rangle \left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle} + \frac{\mathrm{Var}(\tilde{C}_\ell^{\,\mathrm{gal-gal}})}{\left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle^2} \right) \\ &= C_\ell^{\mathrm{ISW-gal}} \frac{\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-gal}} \right\rangle}{\left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle} \left(1 - \frac{2}{2\ell+1} + \frac{2}{2\ell+1} \right),\end{aligned}$$

where we used

$$\begin{aligned}\mathrm{Var}\left(\tilde{C}_\ell^{\,\mathrm{gal-gal}}\right) &= \frac{2}{2\ell+1} C_\ell^{\mathrm{gal-gal}}, \\ \mathrm{Cov}\left(\tilde{C}_\ell^{\,\mathrm{gal-gal}}, \tilde{C}_\ell^{\,\mathrm{ISW-gal}}\right) &= \frac{2}{2\ell+1} C_\ell^{\mathrm{gal-gal}} C_\ell^{\mathrm{ISW-gal}}.\end{aligned}$$

The corrective terms vanish and we find

$$\begin{aligned}\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec}} \right\rangle &\approx C_\ell^{\mathrm{ISW-gal}} \frac{\left\langle \tilde{C}_\ell^{\,\mathrm{ISW-gal}} \right\rangle}{\left\langle \tilde{C}_\ell^{\,\mathrm{gal-gal}} \right\rangle} \\ &= \left\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec,th}} \right\rangle.\end{aligned} \tag{A.20}$$

Then on average, the cross-power between the true and reconstructed ISW maps is unchanged from the theory case. Since the multipoles are independent, this means the total covariance between the true and reconstructed ISW maps is unchanged as well:

$$\begin{aligned}\left\langle \tilde{\mathrm{Cov}}\left(T^{\mathrm{ISW}}, T^{\mathrm{rec}}\right) \right\rangle &= \frac{1}{4\pi} \sum_\ell (2\ell+1) \left\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec}} \right\rangle \\ &= \frac{1}{4\pi} \sum_\ell (2\ell+1) \left\langle \tilde{C}_\ell^{\,\mathrm{ISW-rec,th}} \right\rangle\end{aligned}$$

While for the autopower $\tilde{C}_\ell^{\,\mathrm{rec-rec}}$ we were able to derive an analytical result, a similar Taylor expansion treatment to the same order as the cross-power results in an additive correction of $2/(2\ell+1)$, or

$$\left\langle \tilde{C}_\ell^{\,\mathrm{rec-rec}} \right\rangle \approx \left(\frac{2\ell+3}{2\ell+1} \right) \left\langle C_\ell^{\mathrm{rec-rec,th}} \right\rangle, \tag{A.21}$$

which is a good approximation to the analytical result found in Eq. (A.17).

Appendix B

LSS Systematics Mitigation Appendices

B.1 Lognormal vs. Gaussian Signal Maps

While the methods presented in Chapter 3 are quite general for any case where systematic contamination can be traced using a template, we have specifically worked in the context of galaxy clustering. In this case, the signal map s that we are attempting to model is the galaxy overdensity δ, which is subject to the constraint $\delta > -1$ (as is the case for any overdensity statistic). Thus our assumption that s is Gaussian breaks down at low redshift and at small scales, when $|\delta|$ can be large.

It is well known that galaxy and shear overdensities are better approximated by a lognormal distribution [see e.g. 307–310]), so we run the methods on a series of lognormal maps to see if the relative performance of the methods changes.

We generate 100 Gaussian signal realizations $s_G(\hat{\mathbf{n}})$ of the lowest redshift bin of our fiducial DES survey, for which the cosmological signal will be most non-Gaussian. We generate lognormal versions of these maps by first computing the transformation that achieves zero-mean lognormal overdensity field *in the ensemble* [309], then centering and scaling so that each realization of the lognormal field has the same mean and variance as its Gaussian counterpart. The two steps correspond to the mathematical operations:

1. $s'_{LN}(\hat{\mathbf{n}}) = e^{s_G(\hat{\mathbf{n}})} - e^{\mathrm{Var}[s_G(\hat{\mathbf{n}})]/2}$

2. $s_{LN}(\hat{\mathbf{n}}) = \sqrt{\dfrac{\mathrm{Var}[s_G(\hat{\mathbf{n}})]}{\mathrm{Var}[s'_{LN}(\hat{\mathbf{n}})]}}\,(s'_{LN}(\hat{\mathbf{n}}) - \bar{s}'_{LN}(\hat{\mathbf{n}}))$.

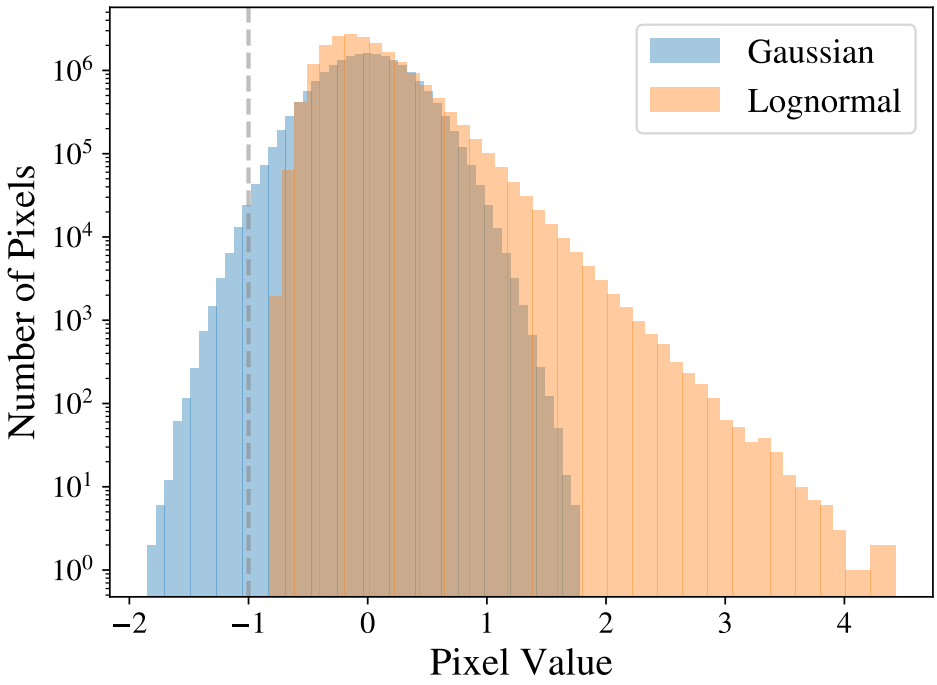

Figure B.1: Distribution of pixel overdensities across all 100 realizations of the lognormal (orange) and Gaussian (blue) maps of the galaxy overdensity in the lowest redshift bin of our fiducial DES-like survey. The Gaussian maps contain pixels with $s < -1$, which is nonphysical in cases like this where s corresponds to an overdensity.

The resulting lognormal realizations are then of the form

$$s_{LN}(\hat{\mathbf{n}}) = \lambda_1 e^{s_G}(\hat{\mathbf{n}}) - \lambda_0, \tag{B.1}$$

with scale and shift parameters of $\lambda_1 = 0.9123 \pm 0.0017$ and $\lambda_0 = 0.9697 \pm 0.0017$, respectively for our lowest redshift bin, which is the most non-Gaussian.

Fig. B.1 shows the distribution of pixel overdensities across all realizations of the lognormal and Gaussian signal maps. It is clear that the Gaussian maps contain many pixels with $s < -1$, which is nonphysical for our case, where s corresponds to an overdensity. The lognormal maps avoid this problem and are a better approximation of the true overdensity distribution. As we have shown, most of the cleaning methods can be viewed under a regression framework wherein the signal distribution is assumed to be Gaussian, so we investigate whether our comparison of methods changes when using a more realistic lognormal distribution.

Fig. B.2 shows the error in the power spectrum reconstruction, given by the $\Delta\chi^2_{C_\ell}$ statistic, for the different methods. We find that while there is some overall shift, using the

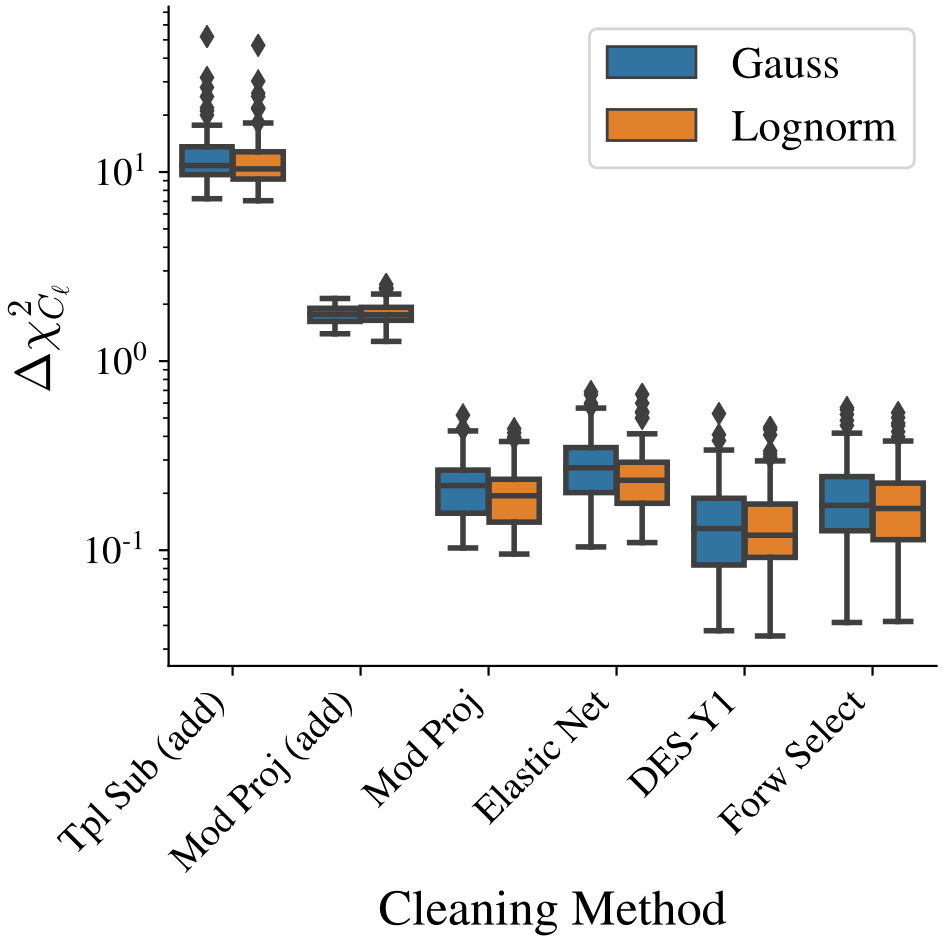

Figure B.2: Box plot showing the performance of each cleaning method when using Gaussian (blue, left) versus lognormal (orange, right) signal maps, as measured by $\Delta\chi^2_{C_\ell}$ of the power spectrum. Filled boxes show the 25-50-75% quartiles, with whiskers encompassing the rest of the distribution out to 1.5× the inter-quartile range. Points beyond this range are indicated by diamonds. Regardless of whether lognormal or Gaussian maps are used, the relative performance of the methods to one another is largely unchanged, and the Gaussian approximation is negligible compared to neglecting the multiplicative correction of Sec. 3.4.2.

lognormal signal maps does not change the relative behavior of the methods; none of them display a unique susceptibility to the assumption of Gaussianity in the signal maps.

B.2 Effect of Prewhitening

In their derivation of the bias on the estimated power spectrum after (pseudo-C_ℓ) Mode Projection, Elsner et al [99] assume that the map d has been decorrelated ("prewhitened") before projecting out the templates. This is quite difficult to do in practice, as it requires

the inversion of an $N_{\rm pix} \times N_{\rm pix}$ matrix, the same problem with QML estimators for the power spectrum. Indeed, one of the assumptions of pseudo-C_ℓ estimation is that pixels are uncorrelated (though individual pixels are weighted by an estimate of their inverse noise variance and by the mask, see e.g. [13].)

As shown in Sec. 3.4.1, however, the dominant 'noise' in our observations is actually our true clustering signal, so a true 'prewhitening' step should more appropriately inverse weight the data by the expected clustering variance. This can be done efficiently in harmonic space when there is no mask, as the clustering signal is diagonal, circumventing the need to invert a large covariance matrix.

We can define prewhitened data vectors for our observed overdensity field and templates as

$$(d_{\rm obs})'_{\ell m} = (d_{\rm obs})_{\ell m} / \sqrt{C_\ell^{ss}}. \tag{B.2}$$

$$(t_i)'_{\ell m} = (t_i)_{\ell m} / \sqrt{C_\ell^{ss}}, \tag{B.3}$$

which results in coefficient estimates of

$$\hat{\alpha} = ({\mathbf{T}'}^\dagger \mathbf{T}')^{-1} {\mathbf{T}'}^\dagger \mathbf{d}'_{\rm obs}, \tag{B.4}$$

where T' is a $N_{\ell m} \times N_{\rm tpl}$ matrix with complex entries defined in Eq. (B.3). We can compute the amplitudes directly with

$$\hat{\alpha} = \frac{\sum_{\ell=0}^{\ell_{\rm max}} (2\ell+1) \tilde{C}_\ell^{\,td} / C_\ell^{ss}}{\sum_{\ell=0}^{\ell_{\rm max}} (2\ell+1) \tilde{C}_\ell^{\,tt} / C_\ell^{ss}}. \tag{B.5}$$

We found that prewhitening improved $\Delta\chi^2_{\rm C_\ell}$ by a mean of ~ 0.05 with dispersion 0.08 across the mocks, with similar shifts regardless of whether the multiplicative correction was applied or not. Fig. B.3 shows the improvement from the standard case (blue) to the prewhitened case (orange) for both additive and multiplicative mode projection. While we do not show it, we found that the benefit of prewhitening increased for realizations that had worse power spectrum estimates (higher $\Delta\chi^2_{\rm C_\ell}$), in effect catching and mitigating particularly bad realizations.

In practice, one would either assume a prior power spectrum for prewhitening or compute it iteratively, just as one does for the Mode Projection debiasing step, so this could easily be incorporated into existing Mode Projection routines such as `NaMaster`. As noted in Sec. 3.4.1, since Mode Projection is equivalent to regression, this improvement also quantifies the expected level of improvement that would come from accounting for the covariance

between pixels in pixel-based regression methods.

Analyses on real data will of course be complicated by the mask, which correlates different multipoles, but this can be addressed by suitable binning of the multipoles. Indeed, the standard pseudo-C_ℓ Mode Projection assumes a flat power spectrum and so can be thought of as the limiting case of using only a single bin across multipoles with equal weighting, such that even a rough estimate of the signal power spectrum should offer improvement.

The other methods tested here should benefit similarly from prewhitening, with the possible exception of the DES-Y1 method, which already incorporates an estimate of the covariance of s (which accounts for much of the methods' complexity). The Forward Selection method we presented may be particularly impacted, since the estimated covariance of the fit parameters is underestimated when the pixel covariance is neglected, and this is used for the significance criterion for selecting a template. This could be one reason why the Forward Selection method sometimes failed to reduce all templates to below a significance of $\Delta\chi^2/\Delta\chi^2_0 = 2$ — such a threshold was artificially low compared to what would be expected from random variation.

As noted in Eq. 3.51, the prewhitening step in Eq. (B.3) should optimally include contributions from the systematics as well. However as this represents minor perturbations to the major prewhitening correction above and is hence a small 'error on the error', the effects should be small. This is consistent with Elvin-Poole et al [212], who found negligible impact on their method from neglecting the additional systematics contribution to their estimated covariance matrices.

B.3 Comparison with `NaMaster`

We have used our own implementation of the Mode Projection method and have tested it against that of `NaMaster`, finding good agreement. `NaMaster` computes the power spectra given a set of templates and observations, but does not produce map estimates, so we compare the two implementations using the cleaned power spectrum only. The left panel of Fig. B.4 shows the relative error of the estimated power spectrum when cleaned using NaMaster vs. our own implementation, using the exact same contaminated map and templates and we find good agreement (this held true for all realizations tested). There is very slight disagreement at larger scales (low ℓ), which may be numerical artifacts from the `Master` [221] algorithm implemented to account for mode coupling on a cut sky being applied to full-sky input maps. Regardless, the deviations between the two are small for $\ell > 2$.

B.4 Accounting for the Monopole

It is worth saying a few words about the monopole term, both as in terms of prediction and as it relates to regression.

Firstly, the *overdensity* residuals do not correspond to the *number* density residuals. Even with a perfect reconstruction $\hat{s} = s$, the true number density will be unknown up to a factor of γ,

$$N_{\rm true} = \gamma \langle N_{\rm obs} \rangle_{\rm pix} (s + 1), \tag{B.6}$$

and as such the estimated number density could be quite different from the truth. Fig. 3.5 shows a somewhat unintuitive consequence of this. The single sytematic that contaminates the field has the form $f_{\rm sys} \propto -t$, so that it only obscures galaxies from view ($f_{\rm sys} \leq 0$). At $t = 0$, there is no contamination and so $N_{\rm obs} = N_{\rm true}$, however as the figure shows the *over*-density residuals are quite large. This is because the *mean* number density is significantly underestimated, so pixels with no obscuration are preferentially (and wrongly) estimated to reside in overdense regions.[1]

Secondly, a net monopole in $f_{\rm sys}$ corresponds to the intercept in the regression methods (a column of ones in T). In OLS regression (or pseudo-C_ℓ Mode Projection), the fit is guaranteed to go through the center of mass of the points, $(\bar{t}, \bar{d}_{\rm obs})$, such that including a monopole is unnecessary with such methods if working with overdensities and zero-centered templates. In such cases, the 'projection' of the monopole has already been done by subtracting the mean from the density and template maps (consider Eq. (3.17) with a template of all 1s). We showed in Eqs. (3.38)-(3.44) how how this also holds in the multiplicative case.

In realistic situations, there is high susceptibility to human error if a monopole term is not included — previously zero-centered maps can easily shift through template transformations, mask adjustments, and the application of a mask to mocks, resulting in wildly biased contamination estimates that may be difficult to detect. For example, it is easy to pass templates that are not zero-centered to current pseudo-C_ℓ Mode Projection methods such as implemented in `NaMaster` and receive highly biased spectra without warning (see right panel of Fig. B.4).

The DES-Y1 and Forward Selection methods both already include an intercept term, in keeping with the original formulation of the DES-Y1 method, though in practice it should be very close to zero.

We therefore opt to include a monopole term in our Elastic Net method, as this ensures the method is robust and generalizes the process beyond overdensities to non-zero mean fields, and it will naturally be ignored as a template if it does not contribute information.

[1] In other words, $\gamma > 1$, so from Eq. (3.38), $\langle d_{\rm obs}|_{f_{\rm sys}=0} \rangle_{\rm pix} > 0$.

B.5 Map Error when Varying Number of Templates

Figure B.5 shows the root mean squared error (RMSE) of the estimated overdensity map when varying the number of templates used for cleaning (see Fig. 3.11 for details). Typical map errors are small, with $\mathrm{RMSE}_s \lesssim 0.03$ for most of the cases studied and trends similar to Fig. 3.11. Template Subtraction shows particularly bad map reconstruction due to the fact that it fully removes the largest scale modes which have a small contribution to $\Delta\chi^2_{\mathrm{C}_\ell}$ but a large contribution to the RMSE because of inducing a bias on a large number of pixels.

B.6 Impact of $\Delta\chi^2/\Delta\chi^2_0$ on DES-Y1 Analysis

Here we investigate the effect of $\Delta\chi^2/\Delta\chi^2_0$ and σ^2_{sys} on the efficacy of the DES-Y1 method, as described in Sec. 3.3.1. We describe the reconstruction quality with the residual chi squared between the cleaned and true model, $\Delta\chi^2_{\mathrm{C}_\ell}$.

Fig. B.6 shows how $\Delta\chi^2/\Delta\chi^2_0$ affects the reconstruction quality for the DES-Y1 method, as a function of the level of contamination parameterized by the systematic-error variance σ^2_{sys}. We find little reduction in error by lowering the significance threshold below $\Delta\chi^2_{\mathrm{threshold}} = 4$.

At our fiducial level of contamination ($\sigma^2_{\mathrm{sys}} = 10^{-2}$), almost all contaminating templates exceed the highest threshold displayed of $\Delta\chi^2/\Delta\chi^2_0 = 32$ and so are corrected for. The larger the contamination, the more precisely its form can be determined, so as the level of contamination decreases, some contaminated templates are left uncorrected for. This results in the somewhat counter-intuitive turnover in the error for a given threshold level. We found that the lowest threshold of $\Delta\chi^2/\Delta\chi^2_0 = 1$ consistently outperformed higher thresholds, despite the risk of overfitting, in agreement with our results in Sec. 3.7.2, which showed that the extra power from residual contamination is likely more pernicious than the excess removal of power due to overfitting.

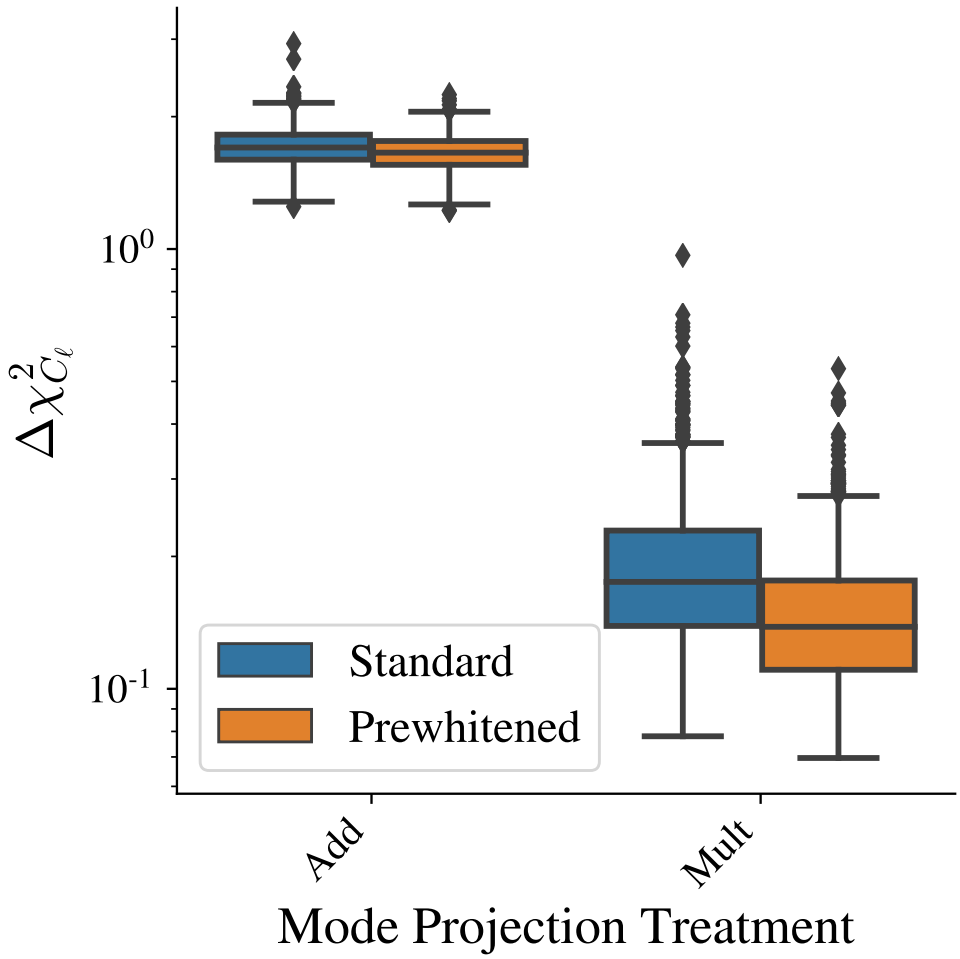

Figure B.3: Impact of prewhitening before cleaning with the multiplicative and additive versions of the Mode Projection method on 1000 realizations for our fiducial contamination model. The standard Mode Projection method assumes a flat power spectrum for the target signal, resulting in a suboptimal estimate of contamination. This can be improved through 'prewhitening' the data vector and templates using a prior power spectrum, which can be shown to be equivalent to a standard weighted regression procedure in harmonic space. There is clear but modest improvement from the standard case (blue) to the nearly-optimal, prewhitened case (orange), with the most improvement seen for realizations that have large error. This can be seen by the preferential reduction of extreme points at the high end of the box plots in the prewhitened case (note the log scale).

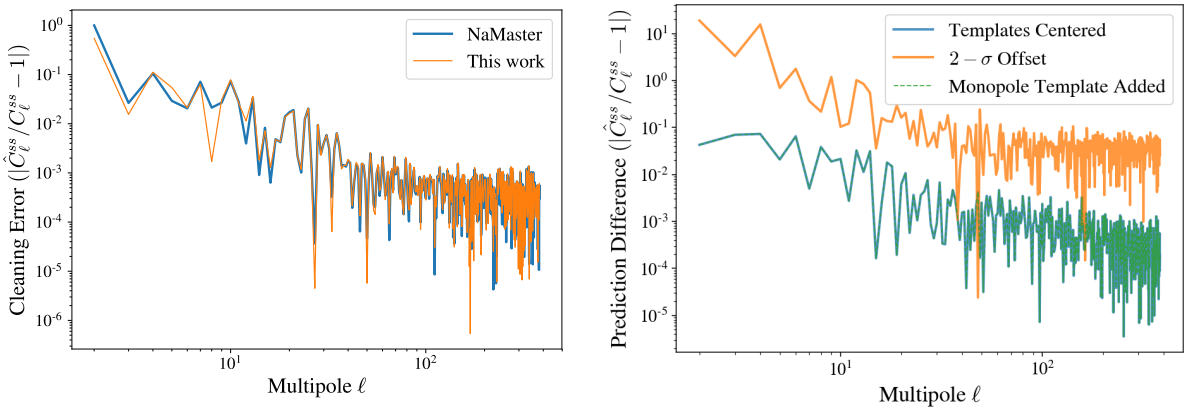

Figure B.4: Validation tests of the Mode Projection map-cleaning procedure. *Left panel:* Comparison of Mode Projection performance on an additive-only contaminated map (as assumed by the Mode Projection method), using `NaMaster` (blue) and our own implementation (orange). The agreement between the two is very good. *Right panel:* Impact of not pre-centering cleaning templates in `NaMaster`. The blue curve indicates the standard use case, where contamination is additive and completely described by the templates, which have been individually centered at zero. If templates are instead centered at another value (here we add a constant $2\sigma_{\mathrm{tpl},i}$ offset to each template, where $\sigma_{\mathrm{tpl},i}$ is the standard deviation of values in template map i. Adding a monopole template completely mitigates the bias from non-centered templates.

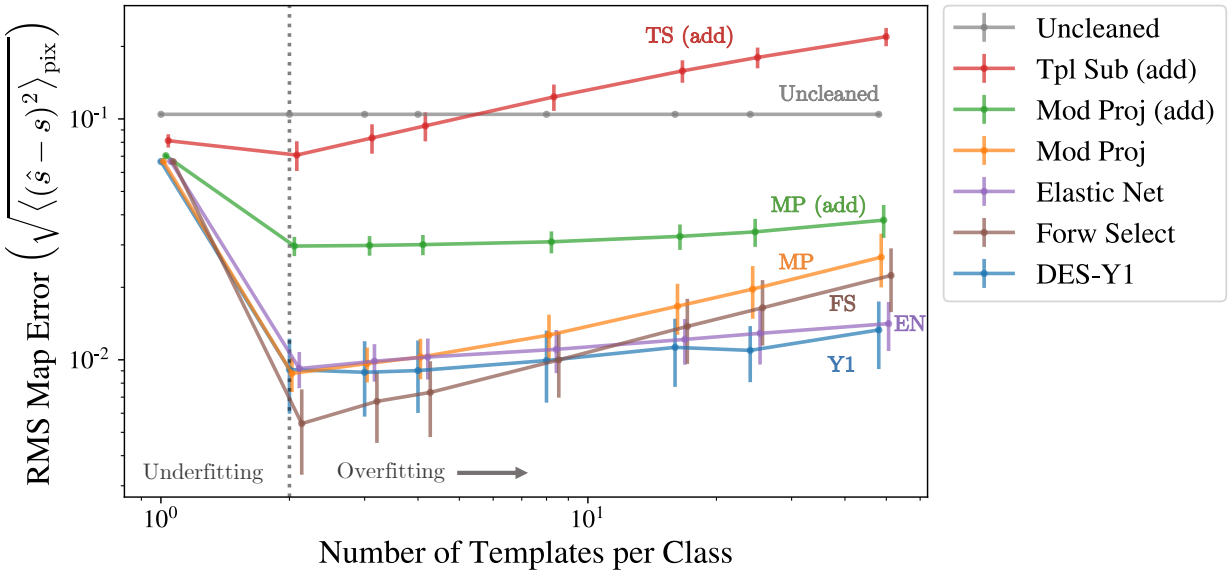

Figure B.5: Same as Fig. 3.11 but showing the RMSE in the estimated overdensity map for each method, rather than error in the power spectrum. Trends are very similar. See Sec. 3.7.2 for details.

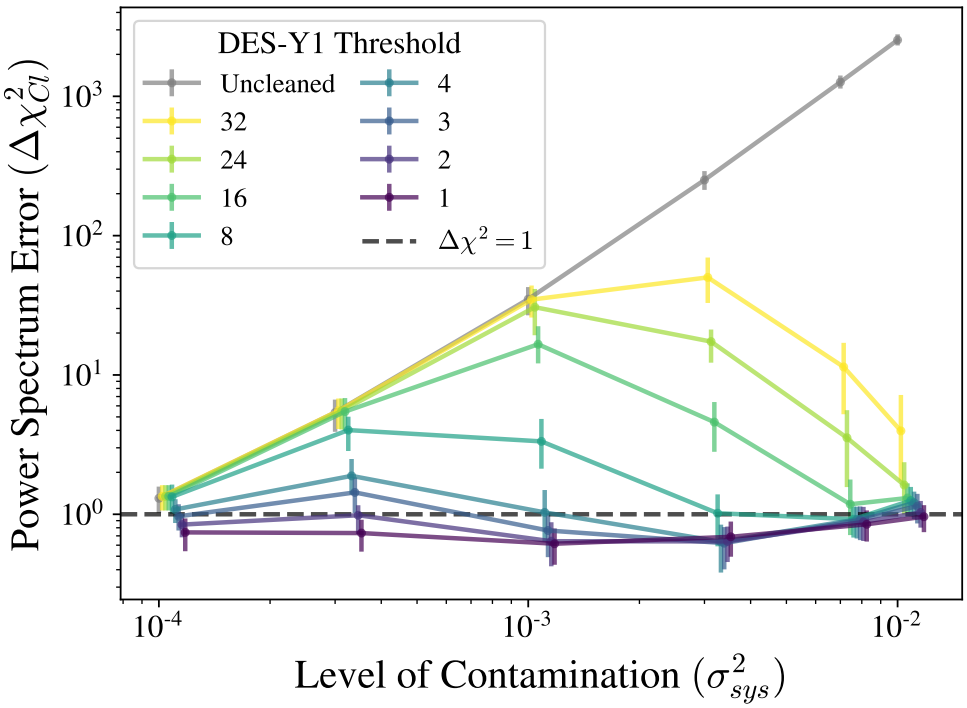

Figure B.6: Dependence of the power spectrum error ($\Delta\chi^2_{C_\ell}$) on the level of contamination $\sigma^2_{\rm sys}$ (x-axis), and on the stopping criterion $\Delta\chi^2/\Delta\chi^2_0$ used for the DES-Y1 method (colors). Points are offset for clarity. For comparison, the variance across pixels from the true overdensity in each bin ranges from $\sigma^2_{\rm sig} \in [0.075, 0.122]$ for the 5 redshift bins of our fiducial survey, corresponding to factors of 7.5 — 1220× larger than $\sigma^2_{\rm sys}$ for the points shown.

Appendix C

Chasing the Spectral Runnings Appendices

C.1 Increasing the Number Density $n(z)$

As shown in Figures 4.3 and 4.6, a Euclid-like survey will be unable to constrain the spectral running to $\sigma_\alpha < 10^{-3}$, which would be necessary to be able to detect the value predicted by single-field, slow-roll inflation. To better understand the limiting factors of these forecasted constraints we consider constraints for a survey similar to the Euclid-like one studied above, but with the number density increased dramatically to $n(z) \to 1000 \times n(z)$. We show forecasts for its statistical errors and systematic biases in Fig. C.1. We find that for this high-source-density survey, the LSS information tightens constraints at lower $k_{\rm max}$, reaching $\sigma_\alpha \lesssim 10^{-3}$ at $k_{\rm max} \approx \{0.5, 0.7, 2\} h\,{\rm Mpc}^{-1}$ for the no nuisance, Mead, and MFP models, respectively. We also find that the increased density makes parameter estimation for α_s more sensitive to systematic biases: if $P(k)$ is mismodeled, then only the MFP model improves constraints over CMB-S4 before introducing unacceptable levels of bias.

This hypothetical $1000 \times n(z)$ survey is also useful to gauge the effect of the trispectrum-induced covariance on cosmological parameter constraints from modes in the strongly nonlinear regime. Unlike our main results in Figure 4.3, where the covariance term was shot-noise dominated at small scales, the trispectrum term becomes important when the number density is very high. The result, as can be seen in Figure C.1, is that there is little improvement in σ_α—especially when combining with CMB-S4—from wavenumbers $k \gtrsim 2h\,{\rm Mpc}^{-1}$. Note that we have not included the additional "super-sample covariance" term [311] that could further degrade the contribution from modes in the nonlinear regime.

Therefore we conclude that, once the realistic systematics are accounted for, even a

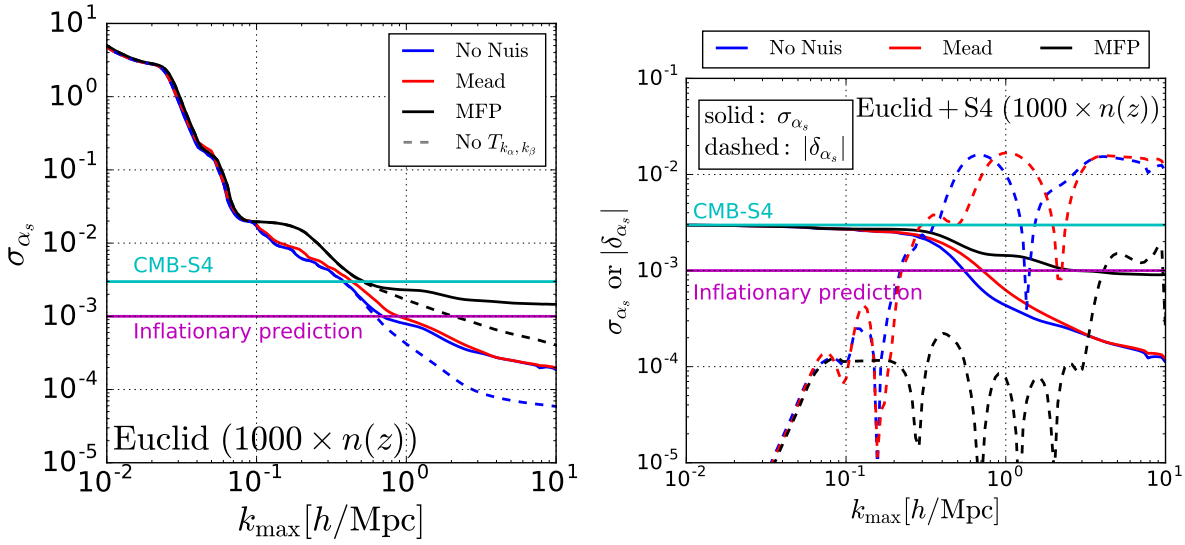

Figure C.1: Constraints on the spectral running α_s for a hypothetical survey with $n(z) \to 1000 \times [n(z)_{\mathrm{Euclid}}]$ alone (left, compare to Fig. 4.3) and with a CMB-S4 experiment (right, compare to Fig. 4.6). While the LSS constraints improve with the increased number density, the trispectrum still limits the information that can be gained from nonlinear scales of $k \gtrsim 0.6h\,\mathrm{Mpc}^{-1}$ (*left*, dashed vs. solid). If $P(k,z)$ is mismodeled, then only the MFP prescription (black) improves constraints over CMB-S4 before significantly biasing the results.

Euclid-like survey with an artificially high number density of sources is unable to reach the precision required to detect the spectral runnings predicted by single-field, slow-roll inflationary models.

C.2 Robustness of Results to Choice of $\delta P(k,\mu,z)$

As noted in Sec. 4.4, here we consider the robustness of our parameter bias results against changes to the shape of $\delta P(k,\mu,z)$. We do this by computing the differences between various prescriptions for the nonlinear power spectrum available in `CAMB`. Because we want to test sensitivity to the shape of $\delta P(k,\mu,z)$, we normalize each curve so that its RMS over all z and $0.005 < k \leq 10h\,\mathrm{Mpc}^{-1}$ is equal to that of our fiducial "takahashi-mead" $\delta P(k,\mu,z)$.

Fig. C.2 shows the resulting ensemble of $\delta P(k,\mu,z)$ considered, for the shallowest redshift bin, $z = 0.65$. When looking at this Figure, there are a couple of things worth noting. First, because we are primarily interested in how constraints on the runnings become biased as we push to higher scales, i.e. $k_{\max}$ at which $|\delta_{p_i}|/\sigma_{p_i} = 1$, the results are insensitive to the sign of $\delta P(k,\mu,z)$. Second, the relatively small magnitude of the bird−peacock (orange) curve is due to its large magnitude at higher redshifts compared to the other curves. Thus the low-z range shown contributes less to its normalized RMS is less than it does for the other curves.

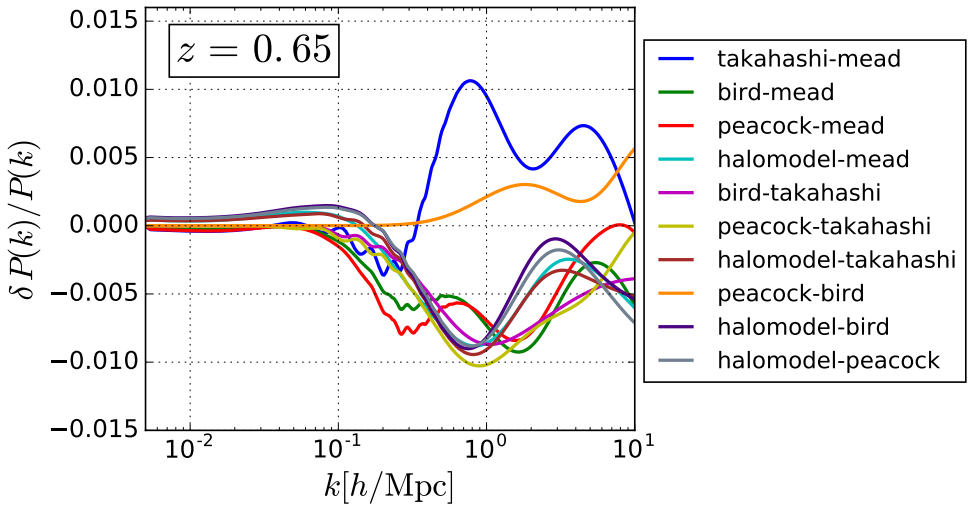

Figure C.2: Other systematic shifts in $P(k)$ tested to verify that the results of Sec. 4.4 are robust to choice of $\delta P(k)$. Note that because Eq. (4.17) is linear in $\delta P(k)$ and we are interested in where $|\delta|/\sigma = 1$, the overall sign of $\delta P(k)$ is inconsequential.

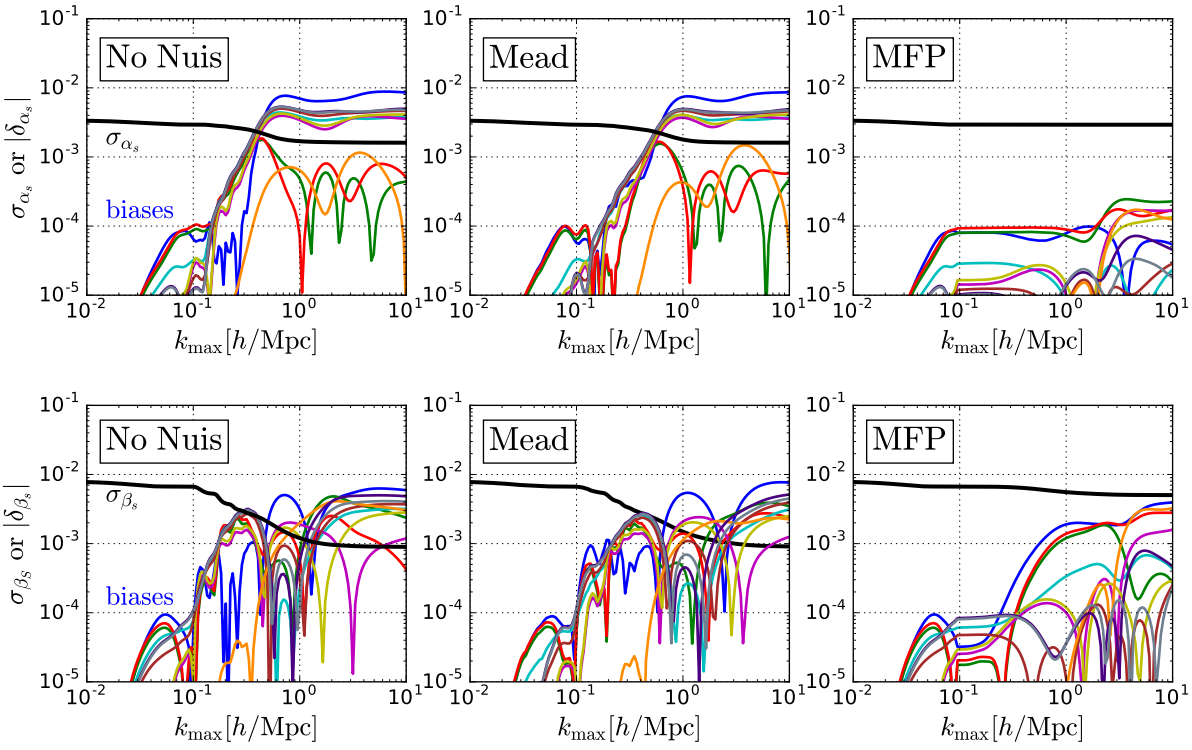

Figure C.3: Parameter bias from different $\delta P(k, z)$ for ΛCDM+α_s + β_s using Euclid + CMB-S4 for α_s (top) and β_s (bottom). The 1σ uncertainty is in black and columns correspond to different nonlinear prescriptions from Sec. 4.2.

The parameter biases in α_s and β_s resulting from these $\delta P(k, \mu, z)$ curves are shown in Fig. C.3 for the combined analysis of Euclid and CMB-S4. These biases are analogous to those shown in Fig. 4.7. Though there is certainly variation in the shape of the curves, we see that the results for $\delta_{\alpha_s}(k_{\rm max})$ and $\delta_{\beta_s}(k_{\rm max})$ for our fiducial $\delta P(k, \mu, z)$ (blue solid curves) are fairly typical. Therefore, we conclude that our fiducial choice of the uncorrected bias in $P(k, z)$ at small scales, given in Eq. (4.16), is fairly typical of such choices.

Bibliography

[1] G. Van Rossum and F. L. Drake Jr, Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

[2] Anaconda Software Distribution, "Anaconda."

[3] J. D. Hunter, *Matplotlib: A 2d graphics environment*, Computing in Science & Engineering **9** (2007), no. 3 90–95.

[4] A. Lewis, A. Challinor, and A. Lasenby, *Efficient computation of CMB anisotropies in closed FRW models*, Astrophys. J. **538** (2000) 473–476, [astro-ph/9911177].

[5] J. Lesgourgues, *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview*, arXiv:1104.2932.

[6] R. Takahashi, M. Sato, T. Nishimichi, A. Taruya, and M. Oguri, *Revising the halofit model for the nonlinear matter power spectrum*, The Astrophysical Journal **761** (2012), no. 2 152.

[7] A. Mead, C. Heymans, L. Lombriser, J. Peacock, O. Steele, and H. Winther, *Accurate halo-model matter power spectra with dark energy, massive neutrinos and modified gravitational forces*, Mon. Not. Roy. Astron. Soc. **459** (2016), no. 2 1468–1488, [arXiv:1602.02154].

[8] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. Fernández del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, *Array programming with NumPy*, Nature **585** (2020) 357–362.

[9] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python*, Nature Methods **17** (2020) 261–272.

[10] M. L. Waskom, *seaborn: statistical data visualization*, Journal of Open Source Software **6** (2021), no. 60 3021.

[11] A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski, *healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in python*, Journal of Open Source Software **4** (Mar., 2019) 1298.

[12] K. M. Gorski, E. 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** (2005) 759–771, [astro-ph/0409513].

[13] **LSST Dark Energy Science** Collaboration, D. Alonso, J. Sanchez, and A. Slosar, *A unified pseudo-C_ℓ framework*, Mon. Not. Roy. Astron. Soc. **484** (2019), no. 3 4127–4151, [arXiv:1809.09603].

[14] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, *Scikit-learn: Machine learning in Python*, Journal of Machine Learning Research **12** (2011) 2825–2830.

[15] S. Seabold and J. Perktold, *statsmodels: Econometric and statistical modeling with python*, in 9th Python in Science Conference, 2010.

[16] N. Weaverdyck, J. Muir, and D. Huterer, *Integrated Sachs-Wolfe map reconstruction in the presence of systematic errors*, Phys. Rev. **D97** (2018), no. 4 043515, [arXiv:1709.08661].

[17] J. Muir and D. Huterer, *Reconstructing the integrated Sachs-Wolfe map with galaxy surveys*, Phys. Rev. **D94** (2016), no. 4 043503, [arXiv:1603.06586].

[18] N. Weaverdyck and D. Huterer, *Mitigating contamination in lss surveys: a comparison of methods*, Monthly Notices of the Royal Astronomical Society **503** (Mar, 2021) 5061–5084.

[19] X. Li, N. Weaverdyck, S. Adhikari, D. Huterer, J. Muir, and H.-Y. Wu, *The Quest for the Inflationary Spectral Runnings in the Presence of Systematic Errors*, Astrophys. J. **862** (2018), no. 2 137, [arXiv:1806.02515].

[20] T. Abbott, F. Abdalla, S. Avila, M. Banerji, E. Baxter, K. Bechtol, M. Becker, E. Bertin, J. Blazek, S. Bridle, and et al., *Dark energy survey year 1 results: Constraints on extended cosmological models from galaxy clustering and weak lensing*, Physical Review D **99** (Jun, 2019).

[21] P. Lemos, M. Raveri, A. Campos, Y. Park, C. Chang, N. Weaverdyck, D. Huterer, A. R. Liddle, J. Blazek, R. Cawthon, A. Choi, J. DeRose, S. Dodelson, C. Doux, M. Gatti, D. Gruen, I. Harrison, E. Krause, O. Lahav, N. MacCrann, J. Muir, J. Prat, M. M. Rau, R. P. Rollins, S. Samuroff, J. Zuntz, W. G. Hartley, B. Hoyle,

I. Sevilla-Noarbe, M. A. Troxel, M. Aguena, S. Allam, J. Annis, S. Avila, D. Bacon, G. M. Bernstein, E. Bertin, D. Brooks, D. L. Burke, A. C. Rosell, M. C. Kind, J. Carretero, C. Conselice, M. Costanzi, M. Crocce, M. E. S. Pereira, J. D. Vicente, S. Desai, H. T. Diehl, P. Doel, K. Eckert, T. F. Eifler, J. Elvin-Poole, S. Everett, A. E. Evrard, I. Ferrero, A. Ferté, B. Flaugher, P. Fosalba, J. Frieman, J. García-Bellido, E. Gaztanaga, D. W. Gerdes, T. Giannantonio, R. A. Gruendl, J. Gschwend, G. Gutierrez, S. R. Hinton, D. L. Hollowood, K. Honscheid, E. M. Huff, D. J. James, M. Jarvis, M. Lima, M. A. G. Maia, M. March, J. L. Marshall, P. Martini, P. Melchior, F. Menanteau, R. Miquel, J. J. Mohr, R. Morgan, J. Myles, R. L. C. Ogando, A. Palmese, S. Pandey, F. Paz-Chinchón, A. A. Plazas, M. Rodriguez-Monroy, A. Roodman, E. Sanchez, V. Scarpine, M. Schubnell, L. F. Secco, S. Serrano, M. Smith, M. Soares-Santos, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, C. To, T. N. Varga, J. Weller, and W. Wester, *Assessing tension metrics with dark energy survey and planck data*, 2020.

[22] A. Chen, D. Huterer, S. Lee, A. Ferté, N. Weaverdyck, O. A. Alves, C. D. Leonard, N. MacCrann, M. Raveri, A. Porredon, E. D. Valentino, J. Muir, P. Lemos, A. Liddle, J. Blazek, A. Campos, R. Cawthon, A. Choi, S. Dodelson, J. Elvin-Poole, D. Gruen, A. Ross, L. F. Secco, I. Sevilla, E. Sheldon, M. A. Troxel, J. Zuntz, T. Abbott, M. Aguena, S. Allam, J. Annis, S. Avila, E. Bertin, S. Bhargava, S. Bridle, D. Brooks, A. C. Rosell, M. C. Kind, J. Carretero, M. Costanzi, M. Crocce, L. da Costa, M. E. da Silva Pereira, T. Davis, P. Doel, T. Eifler, I. Ferrero, P. Fosalba, J. Frieman, J. Garcia-Bellido, E. Gaztanaga, D. Gerdes, R. Gruendl, J. Gschwend, G. Gutierrez, S. Hinton, D. L. Hollowood, K. Honscheid, B. Hoyle, D. James, M. Jarvis, K. Kuehn, O. Lahav, M. Maia, J. Marshall, F. Menanteau, R. Miquel, R. Morgan, A. Palmese, F. Paz-Chinchon, A. P. Malagón, A. Roodman, E. Sanchez, V. Scarpine, M. Schubnell, S. Serrano, M. Smith, E. Suchyta, G. Tarle, D. Thomas, C.-H. To, T. N. Varga, J. Weller, and R. Wilkinson, *Constraints on dark matter to dark radiation conversion in the late universe with des-y1 and external data*, 2021.

[23] M. Rodríguez-Monroy, N. Weaverdyck, J. Elvin-Poole, M. Crocce, A. C. Rosell, F. Andrade-Oliveira, S. Avila, K. Bechtol, G. M. Bernstein, J. Blazek, H. Camacho, R. Cawthon, J. D. Vicente, J. DeRose, S. Dodelson, S. Everett, X. Fang, I. Ferrero, A. Ferté, O. Friedrich, E. Gaztanaga, G. Giannini, R. A. Gruendl, W. G. Hartley, K. Herner, E. M. Huff, M. Jarvis, E. Krause, N. MacCrann, J. Mena-Fernández, J. Muir, S. Pandey, Y. Park, A. Porredon, J. Prat, R. Rosenfeld, A. J. Ross, E. Rozo, E. S. Rykoff, E. Sanchez, D. S. Cid, I. Sevilla-Noarbe, M. Tabbutt, C. To, E. L. Wagoner, R. H. Wechsler, M. Aguena, S. Allam, A. Amon, J. Annis, D. Bacon, E. Baxter, E. Bertin, S. Bhargava, D. Brooks, D. L. Burke, M. C. Kind, J. Carretero, F. J. Castander, A. Choi, C. Conselice, M. Costanzi, L. N. da Costa, M. E. S. Pereira, S. Desai, H. T. Diehl, B. Flaugher, P. Fosalba, J. Frieman, J. García-Bellido, T. Giannantonio, D. Gruen, J. Gschwend, G. Gutierrez, S. R. Hinton, D. L. Hollowood, K. Honscheid, D. Huterer, B. Jain, D. J. James, K. Kuehn, N. Kuropatkin, M. Lima, M. A. G. Maia, M. March, J. L. Marshall, P. Melchior, F. Menanteau, C. J. Miller, R. Miquel, J. J. Mohr, R. Morgan, A. Palmese,

F. Paz-Chinchón, A. Pieres, A. A. P. Malagón, A. Roodman, V. Scarpine, S. Serrano, M. Smith, M. Soares-Santos, E. Suchyta, G. Tarle, D. Thomas, and T. N. Varga, *Dark energy survey year 3 results: Galaxy clustering and systematics treatment for lens galaxy samples*, 2021.

[24] N. Suzuki, D. Rubin, C. Lidman, G. Aldering, R. Amanullah, K. Barbary, L. F. Barrientos, J. Botyanszki, M. Brodwin, N. Connolly, and et al., *Thehubble space telescopecluster supernova survey. v. improving the dark-energy constraints abovez¿ 1 and building an early-type-hosted supernova sample*, The Astrophysical Journal **746** (Jan, 2012) 85.

[25] **Planck Collaboration** Collaboration, P. Ade et al., *Planck 2013 results. XIX. The integrated Sachs-Wolfe effect*, Astron.Astrophys. **571** (2014) A19, [arXiv:1303.5079].

[26] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XXI. The integrated Sachs-Wolfe effect*, Astron. Astrophys. **594** (2016) A21, [arXiv:1502.01595].

[27] J. B. Muñoz, E. D. Kovetz, A. Raccanelli, M. Kamionkowski, and J. Silk, *Towards a measurement of the spectral runnings*, Journal of Cosmology and Astroparticle Physics **2017** (2017), no. 05 032.

[28] A. Mead, J. Peacock, C. Heymans, S. Joudaki, and A. Heavens, *An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models*, Monthly Notices of the Royal Astronomical Society **454** (2015), no. 2 1958–1975.

[29] H. Mo, F. C. van den Bosch, and S. White, Galaxy Formation and Evolution. 2010.

[30] D. Baumann, *Tasi lectures on inflation*, 2012.

[31] B. Ryden, Introduction to cosmology. Cambridge University Press, 1970.

[32] D. Huterer, *Physics 526: Cosmology ii (lecture notes)*, 2014.

[33] S. Dodelson, Modern Cosmology. Academic Press, Amsterdam, 2003.

[34] A. Einstein, *Kosmologische Betrachtungen zur allgemeinen Relativitätstheorie*, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin (Jan., 1917) 142–152.

[35] D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris, and J. D. McEwen, *How isotropic is the universe?*, Physical Review Letters **117** (Sep, 2016).

[36] S. Weinberg, *The cosmological constant problem*, Rev. Mod. Phys. **61** (Jan, 1989) 1–23.

[37] M. Chevallier and D. Polarski, *Accelerating universes with scaling dark matter*, International Journal of Modern Physics D **10** (2001) 213–224.

[38] E. Linder, *Biased cosmology: Pivots, parameters, and figures of merit*, Lawrence Berkeley National Laboratory (2006).

[39] A. Friedmann, *Über die Krümmung des Raumes*, Zeitschrift fur Physik **10** (Jan., 1922) 377–386.

[40] S. M. Feeney, D. J. Mortlock, and N. Dalmasso, *Clarifying the hubble constant tension with a bayesian hierarchical model of the local distance ladder*, Monthly Notices of the Royal Astronomical Society **476** (Feb, 2018) 3861–3882.

[41] E. P. Hubble, *A spiral nebula as a stellar system, Messier 31.*, Astrophys.J. **69** (Mar., 1929) 103–158.

[42] V. M. Slipher, *Spectrographic Observations of Nebulae*, Popular Astronomy **23** (Jan., 1915) 21–24.

[43] F. Hoyle, *A New Model for the Expanding Universe*, Mon. Not. R. Astron. Soc. **108** (Jan., 1948) 372.

[44] H. Bondi and T. Gold, *The Steady-State Theory of the Expanding Universe*, Mon. Not. R. Astron. Soc. **108** (Jan., 1948) 252.

[45] R. Cyburt, B. Fields, K. Olive, and T. Yeh, *Big bang nucleosynthesis: Present status*, Reviews of Modern Physics **88** (2016) 015004.

[46] **Planck** Collaboration, N. Aghanim et al., *Planck 2018 results. VI. Cosmological parameters*, Astron. Astrophys. **641** (2020) A6, [arXiv:1807.06209].

[47] J. C. Jackson, *A critique of rees's theory of primordial gravitational radiation*, Monthly Notices of the Royal Astronomical Society **156** (Feb, 1972) 1P–5P.

[48] N. Kaiser, *Clustering in real space and in redshift space*, Mon. Not. R. Astron. Soc. **227** (July, 1987) 1–21.

[49] M. Kilbinger, *Cosmology with cosmic shear observations: a review*, Reports on Progress in Physics **78** (Jul, 2015) 086901.

[50] R. Mandelbaum, *Weak lensing for precision cosmology*, Annual Review of Astronomy and Astrophysics **56** (Sep, 2018) 393–433.

[51] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley,

F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca, *Planck 2018 results. VI. Cosmological parameters*, A&A **641** (Sept., 2020) A6, [arXiv:1807.06209].

[52] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry, *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, Astro. Journal **116** (Sept., 1998) 1009–1038, [astro-ph/9805201].

[53] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and T. S. C. Project, *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, Astrophys.J. **517** (June, 1999) 565–586, [astro-ph/9812133].

[54] C. Sagan, A. Druyan, and S. Soter, *Cosmos: A Personal Voyage*, 1980.

[55] O. Friedrich, F. Andrade-Oliveira, H. Camacho, O. Alves, R. Rosenfeld, J. Sanchez, X. Fang, T. F. Eifler, E. Krause, C. Chang, Y. Omori, A. Amon, E. Baxter, J. Elvin-Poole, D. Huterer, A. Porredon, J. Prat, V. Terra, A. Troja, A. Alarcon, K. Bechtol, G. M. Bernstein, R. Buchs, A. Campos, A. C. Rosell, M. C. Kind, R. Cawthon, A. Choi, J. Cordero, M. Crocce, C. Davis, J. DeRose, H. T. Diehl, S. Dodelson, C. Doux, A. Drlica-Wagner, F. Elsner, S. Everett, P. Fosalba, M. Gatti, G. Giannini, D. Gruen, R. A. Gruendl, I. Harrison, W. G. Hartley, B. Jain, M. Jarvis, N. MacCrann, J. McCullough, J. Muir, J. Myles, S. Pandey, M. Raveri,

A. Roodman, M. Rodriguez-Monroy, E. S. Rykoff, S. Samuroff, C. Sánchez, L. F. Secco, I. Sevilla-Noarbe, E. Sheldon, M. A. Troxel, N. Weaverdyck, B. Yanny, M. Aguena, S. Avila, D. Bacon, E. Bertin, S. Bhargava, D. Brooks, D. L. Burke, J. Carretero, M. Costanzi, L. N. da Costa, M. E. S. Pereira, J. D. Vicente, S. Desai, A. E. Evrard, I. Ferrero, J. Frieman, J. García-Bellido, E. Gaztanaga, D. W. Gerdes, T. Giannantonio, J. Gschwend, G. Gutierrez, S. R. Hinton, D. L. Hollowood, K. Honscheid, D. J. James, K. Kuehn, O. Lahav, M. Lima, M. A. G. Maia, F. Menanteau, R. Miquel, R. Morgan, A. Palmese, F. Paz-Chinchón, A. A. Plazas, E. Sanchez, V. Scarpine, S. Serrano, M. Soares-Santos, M. Smith, E. Suchyta, G. Tarle, D. Thomas, C. To, T. N. Varga, J. Weller, and R. D. Wilkinson, *Dark energy survey year 3 results: Covariance modelling and its impact on parameter estimation and quality of fit*, 2020.

[56] R. Sachs and A. Wolfe, *Perturbations of a cosmological model and angular variations of the microwave background*, Astrophys.J. **147** (1967) 73–90.

[57] W. Hu and N. Sugiyama, *The Small scale integrated Sachs-Wolfe effect*, Phys. Rev. **D50** (1994) 627–631, [astro-ph/9310046].

[58] S. Boughn and R. Crittenden, *A Correlation of the cosmic microwave sky with large scale structure*, Nature **427** (2004) 45–47, [astro-ph/0305001].

[59] P. Fosalba, E. Gaztanaga, and F. Castander, *Detection of the ISW and SZ effects from the CMB-galaxy correlation*, Astrophys. J. **597** (2003) L89–92, [astro-ph/0307249].

[60] **WMAP** Collaboration, M. R. Nolta et al., *First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Dark energy induced correlation with radio sources*, Astrophys. J. **608** (2004) 10–15, [astro-ph/0305097].

[61] P.-S. Corasaniti, T. Giannantonio, and A. Melchiorri, *Constraining dark energy with cross-correlated CMB and large scale structure data*, Phys. Rev. **D71** (2005) 123521, [astro-ph/0504115].

[62] N. Padmanabhan, C. M. Hirata, U. Seljak, D. J. Schlegel, J. Brinkmann, and D. P. Schneider, *Correlating the CMB with luminous red galaxies: The Integrated Sachs-Wolfe effect*, Phys. Rev. **D72** (2005) 043525, [astro-ph/0410360].

[63] P. Vielva, E. Martinez-Gonzalez, and M. Tucci, *WMAP and NVSS cross-correlation in wavelet space: ISW detection and dark energy constraints*, Mon. Not. Roy. Astron. Soc. **365** (2006) 891, [astro-ph/0408252].

[64] J. D. McEwen, P. Vielva, M. P. Hobson, E. Martinez-Gonzalez, and A. N. Lasenby, *Detection of the integrated Sachs–Wolfe effect and corresponding dark energy constraints made with directional spherical wavelets*, Mon. Not. Roy. Astron. Soc. **376** (2007) 1211–1226, [astro-ph/0602398].

[65] T. Giannantonio, R. G. Crittenden, R. C. Nichol, R. Scranton, G. T. Richards, A. D. Myers, R. J. Brunner, A. G. Gray, A. J. Connolly, and D. P. Schneider, *A high redshift detection of the integrated Sachs-Wolfe effect*, Phys. Rev. **D74** (2006) 063520, [astro-ph/0607572].

[66] A. Cabre, P. Fosalba, E. Gaztanaga, and M. Manera, *Error analysis in cross-correlation of sky maps: Application to the ISW detection*, Mon.Not.Roy.Astron.Soc. **381** (2007) 1347, [astro-ph/0701393].

[67] A. Rassat, K. Land, O. Lahav, and F. B. Abdalla, *Cross-correlation of 2MASS and WMAP3: Implications for the Integrated Sachs-Wolfe effect*, Mon. Not. Roy. Astron. Soc. **377** (2007) 1085–1094, [astro-ph/0610911].

[68] T. Giannantonio, R. Scranton, R. G. Crittenden, R. C. Nichol, S. P. Boughn, et al., *Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications*, Phys.Rev. **D77** (2008) 123520, [arXiv:0801.4380].

[69] S. Ho, C. Hirata, N. Padmanabhan, U. Seljak, and N. Bahcall, *Correlation of CMB with large-scale structure: I. ISW Tomography and Cosmological Implications*, Phys. Rev. **D78** (2008) 043519, [arXiv:0801.0642].

[70] J.-Q. Xia, M. Viel, C. Baccigalupi, and S. Matarrese, *The High Redshift Integrated Sachs-Wolfe Effect*, JCAP **0909** (2009) 003, [arXiv:0907.4753].

[71] T. Giannantonio, R. Crittenden, R. Nichol, and A. J. Ross, *The significance of the integrated Sachs-Wolfe effect revisited*, Mon.Not.Roy.Astron.Soc. **426** (2012) 2581–2599, [arXiv:1209.2125].

[72] D. J. Schwarz, C. J. Copi, D. Huterer, and G. D. Starkman, *CMB Anomalies after Planck*, vol. 33, p. 184001, 2016. arXiv:1510.07929.

[73] J. Kim, A. Rotti, and E. Komatsu, *Removing the ISW-lensing bias from the local-form primordial non-Gaussianity estimation*, JCAP **1304** (2013) 021, [arXiv:1302.5799].

[74] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XVII. Constraints on primordial non-Gaussianity*, Astron. Astrophys. **594** (2016) A17, [arXiv:1502.01592].

[75] R. Barreiro, P. Vielva, C. Hernandez-Monteagudo, and E. Martinez-Gonzalez, *A linear filter to reconstruct the ISW effect from CMB and LSS observations*, Radio Sci. **44** (2009) 5004, [arXiv:0809.2557].

[76] B. R. Granett, M. C. Neyrinck, and I. Szapudi, *A Map of the Integrated Sachs-Wolfe Signal from Luminous Red Galaxies*, Astrophys. J. **701** (2009) 414–422, [arXiv:0812.1025].

[77] C. Francis and J. Peacock, *An estimate of the local ISW signal, and its impact on CMB anomalies*, Mon.Not.Roy.Astron.Soc. **406** (2010) 14, [arXiv:0909.2495].

[78] R. B. Barreiro, P. Vielva, A. Marcos-Caballero, and E. Martinez-Gonzalez, *Integrated Sachs-Wolfe map recovery from NVSS and WMAP 7yr data*, Mon. Not. Roy. Astron. Soc. **430** (2013) 259, [arXiv:1208.2350].

[79] A. Rassat, J. L. Starck, and F. X. Dupe, *Removal of two large scale Cosmic Microwave Background anomalies after subtraction of the Integrated Sachs Wolfe effect*, Astron.Astrophys. **557** (2013) A32, [arXiv:1303.4727].

[80] A. Manzotti and S. Dodelson, *Mapping the Integrated Sachs-Wolfe Effect*, Phys.Rev. **D90** (2014), no. 12 123009, [arXiv:1407.5623].

[81] B. Granett, A. Kovács, and A. Hawken, *The integrated sachs–wolfe signal from boss superstructures*, Monthly Notices of the Royal Astronomical Society **454** (2015) 2804–2814.

[82] L. Bonavera, R. B. Barreiro, A. Marcos-Caballero, and P. Vielva, *On the recovery of ISW fluctuations using large-scale structure tracers and CMB temperature and polarization anisotropies*, Mon. Not. Roy. Astron. Soc. **459** (2016), no. 1 657–672, [arXiv:1602.05893].

[83] A. R. Pullen and C. M. Hirata, *Systematic effects in large-scale angular power spectra of photometric quasars and implications for constraining primordial nongaussianity*, Publications of the Astronomical Society of the Pacific, Volume 125, Issue 928, pp. **705-718** (2013) [arXiv:1212.4500].

[84] S. Ho, A. Cuesta, H.-J. Seo, R. de Putter, A. J. Ross, et al., *Clustering of Sloan Digital Sky Survey III Photometric Luminous Galaxies: The Measurement, Systematics and Cosmological Implications*, Astrophys.J. **761** (2012) 14, [arXiv:1201.2137].

[85] S. Ho, N. Agarwal, A. D. Myers, R. Lyons, A. Disbrow, et al., *Sloan Digital Sky Survey III Photometric Quasar Clustering: Probing the Initial Conditions of the Universe using the Largest Volume*, 1311.2597 (2013) [arXiv:1311.2597].

[86] N. Agarwal, S. Ho, and S. Shandera, *Constraining the initial conditions of the Universe using large scale structure*, JCAP **1402** (2014) 038, [arXiv:1311.2606].

[87] T. Giannantonio, A. J. Ross, W. J. Percival, R. Crittenden, D. Bacher, et al., *Improved Primordial Non-Gaussianity Constraints from Measurements of Galaxy Clustering and the Integrated Sachs-Wolfe Effect*, Phys.Rev. **D89** (2014) 023511, [arXiv:1303.1349].

[88] N. Agarwal, S. Ho, A. D. Myers, H.-J. Seo, A. J. Ross, et al., *Characterizing unknown systematics in large scale structure surveys*, JCAP **1404** (2014) 007, [arXiv:1309.2954].

[89] C. Hernández-Monteagudo et al., *The SDSS-III Baryonic Oscillation Spectroscopic Survey: constraints on the integrated Sachs–Wolfe effect*, Mon. Not. Roy. Astron. Soc. **438** (2014), no. 2 1724–1740, [arXiv:1303.4302].

[90] S. A. Thomas, F. B. Abdalla, and O. Lahav, *Excess Clustering on Large Scales in the MegaZ DR7 Photometric Redshift Survey*, Phys. Rev. Lett. **106** (2011) 241301, [arXiv:1012.2272].

[91] C. Blake and J. Wall, *Measurement of the angular correlation function of radio galaxies from the NRAO VLA Sky Survey*, Mon. Not. Roy. Astron. Soc. **329** (2002) L37–L41, [astro-ph/0111328].

[92] C. Hernandez-Monteagudo, *Revisiting the WMAP - NVSS angular cross correlation. A skeptic view*, Astron. Astrophys. **520** (2010) A101, [arXiv:0909.4294].

[93] **SDSS** Collaboration, R. Scranton et al., *Analysis of systematic effects and statistical uncertainties in angular clustering of galaxies from Early SDSS Data*, Astrophys. J. **579** (2002) 48–75, [astro-ph/0107416].

[94] A. J. Ross et al., *Ameliorating Systematic Uncertainties in the Angular Clustering of Galaxies: A Study using SDSS-III*, Mon. Not. Roy. Astron. Soc. **417** (2011) 1350, [arXiv:1105.2320].

[95] B. Leistedt, H. V. Peiris, D. J. Mortlock, A. Benoit-Lévy, and A. Pontzen, *Estimating the large-scale angular power spectrum in the presence of systematics: a case study of Sloan Digital Sky Survey quasars*, Mon.Not.Roy.Astron.Soc. **435** (2013) 1857–1873, [arXiv:1306.0005].

[96] B. Leistedt and H. V. Peiris, *Exploiting the full potential of photometric quasar surveys: Optimal power spectra through blind mitigation of systematics*, Mon.Not.Roy.Astron.Soc. **444** (2014) 2, [arXiv:1404.6530].

[97] F. Elsner, B. Leistedt, and H. V. Peiris, *Unbiased methods for removing systematics from galaxy clustering measurements*, Mon. Not. Roy. Astron. Soc. **456** (2016), no. 2 2095–2104, [arXiv:1509.08933].

[98] H. Awan, E. Gawiser, P. Kurczynski, R. L. Jones, H. Zhan, N. D. Padilla, A. M. Muñnoz Arancibia, A. Orsi, S. A. Cora, and P. Yoachim, *Testing LSST Dither Strategies for Survey Uniformity and Large-Scale Structure Systematics*, Astrophys. J. **829** (2016), no. 1 50, [arXiv:1605.00555].

[99] F. Elsner, B. Leistedt, and H. V. Peiris, *Unbiased pseudo-C? power spectrum estimation with mode projection*, Mon. Not. Roy. Astron. Soc. **465** (2017), no. 2 1847–1855, [arXiv:1609.03577].

[100] **DES** Collaboration, E. Suchyta et al., *No galaxy left behind: accurate measurements with the faintest objects in the Dark Energy Survey*, Mon. Not. Roy. Astron. Soc. **457** (2016), no. 1 786–808, [arXiv:1507.08336].

[101] D. Huterer, C. E. Cunha, and W. Fang, *Calibration errors unleashed: effects on cosmological parameters and requirements for large-scale structure surveys*, Mon.Not.Roy.Astron.Soc. **432** (2013) 2945, [arXiv:1211.1015].

[102] D. L. Shafer and D. Huterer, *Multiplicative errors in the galaxy power spectrum: self-calibration of unknown photometric systematics for precision cosmology*, Mon. Not. Roy. Astron. Soc. **447** (2015) 2961, [arXiv:1410.0035].

[103] N. Afshordi, *Integrated Sachs-Wolfe effect in cross - correlation: The Observer's manual*, Phys. Rev. **D70** (2004) 083536, [astro-ph/0401166].

[104] **EUCLID** Collaboration, R. Laureijs et al., *Euclid Definition Study Report*, arXiv:1110.3193.

[105] M. Douspis, P. G. Castro, C. Caprini, and N. Aghanim, *Optimising large galaxy surveys for ISW detection*, Astron. Astrophys. **485** (2008) 395, [arXiv:0802.0983].

[106] N. Martinet, J. G. Bartlett, A. Kiessling, and B. Sartoris, *Constraining cosmology with shear peak statistics: tomographic analysis*, Astron. Astrophys. **581** (2015) A101, [arXiv:1506.02192].

[107] O. Doré et al., *Cosmology with the SPHEREX All-Sky Spectral Survey*, arXiv:1412.4872.

[108] **Planck** Collaboration, N. Aghanim et al., *Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters*, Astron. Astrophys. **594** (2016) A11, [arXiv:1507.02704].

[109] **Planck** Collaboration, P. A. R. Ade et al., *Planck 2015 results. XV. Gravitational lensing*, Astron. Astrophys. **594** (2016) A15, [arXiv:1502.01591].

[110] E. S. Rykoff, E. Rozo, and R. Keisler, *Assessing Galaxy Limiting Magnitudes in Large Optical Surveys*, Submitted to: Astrophysical Journal (2015) [arXiv:1509.00870].

[111] N. Padmanabhan et al., *An Improved Photometric Calibration of the Sloan Digital Sky Survey Imaging Data*, Astrophys. J. **674** (2008) 1217–1233, [astro-ph/0703454].

[112] F.-X. Dupe, A. Rassat, J.-L. Starck, and M. Fadili, *Measuring the Integrated Sachs-Wolfe Effect*, Astron.Astrophys. **534** (2011) 51, [arXiv:1010.2192].

[113] W. Hu and R. Scranton, *Measuring Dark Energy Clustering with CMB-Galaxy Correlations*, Phys. Rev. **D70** (2004) 123002, [astro-ph/0408456].

[114] M. Frommert, T. A. Ensslin, and F. S. Kitaura, *Optimal ISW detection and joint likelihood for cosmological parameter estimation*, Mon. Not. Roy. Astron. Soc. **391** (2008) 1315, [arXiv:0807.0464].

[115] G.-C. Liu, K.-W. Ng, and U.-L. Pen, *Improved Dark Energy Detection through the Polarization-assisted WMAP-NVSS ISW Correlation*, Phys. Rev. **D83** (2011) 063001, [arXiv:1010.0578].

[116] V. de Lapparent, M. J. Geller, and J. P. Huchra, *A Slice of the universe*, Astrophys. J. **302** (1986) L1–L5.

[117] S. L. Lumsden, R. C. Nichol, C. A. Collins, and L. Guzzo, *The Edinburgh-Durham southern galaxy catalogue. IV. Thecluster catalogue.*, Mon. Not. R. Astron. Soc. **258** (Sept., 1992) 1–22.

[118] **2DFGRS Collaboration** Collaboration, M. Colless et al., *The 2dF Galaxy Redshift Survey: Spectra and redshifts*, Mon.Not.Roy.Astron.Soc. **328** (2001) 1039, [astro-ph/0106498].

[119] M. J. Drinkwater, R. J. Jurek, C. Blake, D. Woods, K. A. Pimbblet, et al., *The WiggleZ Dark Energy Survey: Survey Design and First Data Release*, Mon.Not.Roy.Astron.Soc. **401** (2010) 1429–1452, [arXiv:0911.4246].

[120] **SDSS Collaboration** Collaboration, D. G. York et al., *The Sloan Digital Sky Survey: Technical Summary*, Astron.J. **120** (2000) 1579–1587, [astro-ph/0006396].

[121] K. S. Dawson, D. J. Schlegel, C. P. Ahn, S. F. Anderson, É. Aubourg, S. Bailey, R. H. Barkhouser, J. E. Bautista, A. r. Beifiori, A. A. Berlind, V. Bhardwaj, D. Bizyaev, C. H. Blake, M. R. Blanton, M. Blomqvist, A. S. Bolton, A. Borde, J. Bovy, W. N. Brandt, H. Brewington, J. Brinkmann, P. J. Brown, J. R. Brownstein, K. Bundy, N. G. Busca, W. Carithers, A. R. Carnero, M. A. Carr, Y. Chen, J. Comparat, N. Connolly, F. Cope, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, T. Delubac, R. de Putter, S. Dhital, A. Ealet, G. L. Ebelke, D. J. Eisenstein, S. Escoffier, X. Fan, N. Filiz Ak, H. Finley, A. Font-Ribera, R. Génova-Santos, J. E. Gunn, H. Guo, D. Haggard, P. B. Hall, J.-C. Hamilton, B. Harris, D. W. Harris, S. Ho, D. W. Hogg, D. Holder, K. Honscheid, J. Huehnerhoff, B. Jordan, W. P. Jordan, G. Kauffmann, E. A. Kazin, D. Kirkby, M. A. Klaene, J.-P. Kneib, J.-M. Le Goff, K.-G. Lee, D. C. Long, C. P. Loomis, B. Lundgren, R. H. Lupton, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, M. Manera, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. McDonald, I. D. McGreer, R. G. McMahon, O. Mena, J. Miralda-Escudé, A. D. Montero-Dorta, F. Montesano, D. Muna, A. D. Myers, T. Naugle, R. C. Nichol, P. Noterdaeme, S. E. Nuza, M. D. Olmstead, A. Oravetz, D. J. Oravetz, R. Owen, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, J. K. Parejko, I. Pâris, W. J. Percival, I. Pérez-Fournon, I. Pérez-Ràfols, P. Petitjean, R. Pfaffenberger, J. Pforr, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, R. Rebolo, J. Rich, G. T. Richards, C. M. Rockosi, N. A. Roe, A. J. Ross, N. P. Ross, G. Rossi, J. A. Rubiño-Martin, L. Samushia, A. G. Sánchez, C. Sayres, S. J. Schmidt, D. P. Schneider, C. G. Scóccola, H.-J. Seo, A. Shelden, E. Sheldon, Y. Shen, Y. Shu, A. Slosar, S. A. Smee, S. A. Snedden, F. Stauffer, O. Steele, M. A. Strauss, A. Streblyanska, N. Suzuki, M. E. C. Swanson, T. Tal, M. Tanaka, D. Thomas, J. L. Tinker, R. Tojeiro, C. A. Tremonti, M. Vargas Magaña, L. Verde, M. Viel, D. A. Wake, M. Watson, B. A. Weaver, D. H. Weinberg, B. J. Weiner, A. A. West, M. White, W. M. Wood-Vasey, C. Yeche, I. Zehavi, G.-B. Zhao, and Z. Zheng, *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, Astro. Journal **145** (Jan., 2013) 10, [arXiv:1208.0022].

[122] H. Aihara, N. Arimoto, R. Armstrong, S. Arnouts, N. A. Bahcall, S. Bickerton, J. Bosch, K. Bundy, P. L. Capak, J. H. H. Chan, M. Chiba, J. Coupon, E. Egami, M. Enoki, F. Finet, H. Fujimori, S. Fujimoto, H. Furusawa, J. Furusawa, T. Goto, A. Goulding, J. P. Greco, J. E. Greene, J. E. Gunn, T. Hamana, Y. Harikane, Y. Hashimoto, T. Hattori, M. Hayashi, Y. Hayashi, K. G. Hełminiak, R. Higuchi, C. Hikage, P. T. P. Ho, B.-C. Hsieh, K. Huang, S. Huang, H. Ikeda, M. Imanishi, A. K. Inoue, K. Iwasawa, I. Iwata, A. T. Jaelani, H.-Y. Jian, Y. Kamata, H. Karoji, N. Kashikawa, N. Katayama, S. Kawanomoto, I. Kayo, J. Koda, M. Koike, T. Kojima, Y. Komiyama, A. Konno, S. Koshida, Y. Koyama, H. Kusakabe, A. Leauthaud, C.-H. Lee, L. Lin, Y.-T. Lin, R. H. Lupton, R. Mand elbaum, Y. Matsuoka, E. Medezinski, S. Mineo, S. Miyama, H. Miyatake, S. Miyazaki, R. Momose, A. More, S. More, Y. Moritani, T. J. Moriya, T. Morokuma, S. Mukae, R. Murata, H. Murayama, T. Nagao, F. Nakata, M. Niida, H. Niikura, A. J. Nishizawa, Y. Obuchi, M. Oguri, Y. Oishi, N. Okabe, S. Okamoto, Y. Okura, Y. Ono, M. Onodera, M. Onoue, K. Osato, M. Ouchi, P. A. Price, T.-S. Pyo, M. Sako, M. Sawicki, T. Shibuya, K. Shimasaku, A. Shimono, M. Shirasaki, J. D. Silverman, M. Simet, J. Speagle, D. N. Spergel, M. A. Strauss, Y. Sugahara, N. Sugiyama, Y. Suto, S. H. Suyu, N. Suzuki, P. J. Tait, M. Takada, T. Takata, N. Tamura, M. M. Tanaka, M. Tanaka, M. Tanaka, Y. Tanaka, T. Terai, Y. Terashima, Y. Toba, N. Tominaga, J. Toshikawa, E. L. Turner, T. Uchida, H. Uchiyama, K. Umetsu, F. Uraguchi, Y. Urata, T. Usuda, Y. Utsumi, S.-Y. Wang, W.-H. Wang, K. C. Wong, K. Yabe, Y. Yamada, H. Yamanoi, N. Yasuda, S. Yeh, A. Yonehara, and S. Yuma, *The Hyper Suprime-Cam SSP Survey: Overview and survey design*, Publ. Astr. Soc. Japan **70** (Jan., 2018) S4, [arXiv:1704.05858].

[123] J. T. A. de Jong, G. A. Verdoes Kleijn, D. R. Boxhoorn, H. Buddelmeijer, M. Capaccioli, F. Getman, A. Grado, E. Helmich, Z. Huang, N. Irisarri, K. Kuijken, F. La Barbera, J. P. McFarland, N. R. Napolitano, M. Radovich, G. Sikkema, E. A. Valentijn, K. G. Begeman, M. Brescia, S. Cavuoti, A. Choi, O.-M. Cordes, G. Covone, M. Dall'Ora, H. Hildebrandt, G. Longo, R. Nakajima, M. Paolillo, E. Puddu, A. Rifatto, C. Tortora, E. van Uitert, A. Buddendiek, J. Harnois-Déraps, T. Erben, M. B. Eriksen, C. Heymans, H. Hoekstra, B. Joachimi, T. D. Kitching, D. Klaes, L. V. E. Koopmans, F. Köhlinger, N. Roy, C. Sifón, P. Schneider, W. J. Sutherland, M. Viola, and W.-J. Vriend, *The first and second data releases of the Kilo-Degree Survey*, A&A **582** (Oct., 2015) A62, [arXiv:1507.00742].

[124] **DES** Collaboration, T. M. C. Abbott et al., *Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing*, Phys. Rev. **D98** (2018), no. 4 043526, [arXiv:1708.01530].

[125] **BOSS** Collaboration, L. Anderson et al., *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples*, Mon. Not. Roy. Astron. Soc. **441** (2014), no. 1 24–62, [arXiv:1312.4877].

[126] **BOSS** Collaboration, S. 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** (2017), no. 3 2617–2652, [arXiv:1607.03155].

[127] **DES** Collaboration, T. Abbott et al., *Cosmological Constraints from Multiple Probes in the Dark Energy Survey*, Phys. Rev. Lett. **122** (2019), no. 17 171301, [arXiv:1811.02375].

[128] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. Allende Prieto, J. Annis, S. Bailey, C. Balland, O. Ballester, C. Baltay, L. Beaufore, C. Bebek, T. C. Beers, E. F. Bell, J. L. Bernal, R. Besuner, F. Beutler, C. Blake, H. Bleuler, M. Blomqvist, R. Blum, A. S. Bolton, C. Briceno, D. Brooks, J. R. Brownstein, E. Buckley-Geer, A. Burden, E. Burtin, N. G. Busca, R. N. Cahn, Y.-C. Cai, L. Cardiel-Sas, R. G. Carlberg, P.-H. Carton, R. Casas, F. J. Castand er, J. L. Cervantes-Cota, T. M. Claybaugh, M. Close, C. T. Coker, S. Cole, J. Comparat, A. P. Cooper, M. C. Cousinou, M. Crocce, J.-G. Cuby, D. P. Cunningham, T. M. Davis, K. S. Dawson, A. de la Macorra, J. De Vicente, T. Delubac, M. Derwent, A. Dey, G. Dhungana, Z. Ding, P. Doel, Y. T. Duan, A. Ealet, J. Edelstein, S. Eftekharzadeh, D. J. Eisenstein, A. Elliott, S. Escoffier, M. Evatt, P. Fagrelius, X. Fan, K. Fanning, A. Farahi, J. Farihi, G. Favole, Y. Feng, E. Fernandez, J. R. Findlay, D. P. Finkbeiner, M. J. Fitzpatrick, B. Flaugher, S. Flender, A. Font-Ribera, J. E. Forero-Romero, P. Fosalba, C. S. Frenk, M. Fumagalli, B. T. Gaensicke, G. Gallo, J. Garcia-Bellido, E. Gaztanaga, N. Pietro Gentile Fusillo, T. Gerard, I. Gershkovich, T. Giannantonio, D. Gillet, G. Gonzalez-de-Rivera, V. Gonzalez-Perez, S. Gott, O. Graur, G. Gutierrez, J. Guy, S. Habib, H. Heetderks, I. Heetderks, K. Heitmann, W. A. Hellwing, D. A. Herrera, S. Ho, S. Holland, K. Honscheid, E. Huff, T. A. Hutchinson, D. Huterer, H. S. Hwang, J. M. Illa Laguna, Y. Ishikawa, D. Jacobs, N. Jeffrey, P. Jelinsky, E. Jennings, L. Jiang, J. Jimenez, J. Johnson, R. Joyce, E. Jullo, S. Juneau, S. Kama, A. Karcher, S. Karkar, R. Kehoe, N. Kennamer, S. Kent, M. Kilbinger, A. G. Kim, D. Kirkby, T. Kisner, E. Kitanidis, J.-P. Kneib, S. Koposov, E. Kovacs, K. Koyama, A. Kremin, R. Kron, L. Kronig, A. Kueter-Young, C. G. Lacey, R. Lafever, O. Lahav, A. Lambert, M. Lampton, M. Land riau, D. Lang, T. R. Lauer, J.-M. Le Goff, L. Le Guillou, A. Le Van Suu, J. H. Lee, S.-J. Lee, D. Leitner, M. Lesser, M. E. Levi, B. L'Huillier, B. Li, M. Liang, H. Lin, E. Linder, S. R. Loebman, Z. Lukić, J. Ma, N. MacCrann, C. Magneville, L. Makarem, M. Manera, C. J. Manser, R. Marshall, P. Martini, R. Massey, T. Matheson, J. McCauley, P. McDonald, I. D. McGreer, A. Meisner, N. Metcalfe, T. N. Miller, R. Miquel, J. Moustakas, A. Myers, M. Naik, J. A. Newman, R. C. Nichol, A. Nicola, L. Nicolati da Costa, J. Nie, G. Niz, P. Norberg, B. Nord, D. Norman, P. Nugent, T. O'Brien, M. Oh, K. A. G. Olsen, C. Padilla, H. Padmanabhan, N. Padmanabhan, N. Palanque-Delabrouille, A. Palmese, D. Pappalardo, I. Pâris, C. Park, A. Patej, J. A. Peacock, H. V. Peiris, X. Peng, W. J. Percival, S. Perruchot, M. M. Pieri, R. Pogge, J. E. Pollack, C. Poppett, F. Prada, A. Prakash, R. G. Probst, D. Rabinowitz, A. Raichoor, C. H. Ree, A. Refregier, X. Regal, B. Reid, K. Reil, M. Rezaie, C. M. Rockosi, N. Roe,

S. Ronayette, A. Roodman, A. J. Ross, N. P. Ross, G. Rossi, E. Rozo, V. Ruhlmann-Kleider, E. S. Rykoff, C. Sabiu, L. Samushia, E. Sanchez, J. Sanchez, D. J. Schlegel, M. Schneider, M. Schubnell, A. Secroun, U. Seljak, H.-J. Seo, S. Serrano, A. Shafieloo, H. Shan, R. Sharples, M. J. Sholl, W. V. Shourt, J. H. Silber, D. R. Silva, M. M. Sirk, A. Slosar, A. Smith, G. F. Smoot, D. Som, Y.-S. Song, D. Sprayberry, R. Staten, A. Stefanik, G. Tarle, S. Sien Tie, J. L. Tinker, R. Tojeiro, F. Valdes, O. Valenzuela, M. Valluri, M. Vargas-Magana, L. Verde, A. R. Walker, J. Wang, Y. Wang, B. A. Weaver, C. Weaverdyck, R. H. Wechsler, D. H. Weinberg, M. White, Q. Yang, C. Yeche, T. Zhang, G.-B. Zhao, Y. Zheng, X. Zhou, Z. Zhou, Y. Zhu, H. Zou, and Y. Zu, *The DESI Experiment Part I: Science,Targeting, and Survey Design*, arXiv e-prints (Oct., 2016) arXiv:1611.00036, [arXiv:1611.00036].

[129] L. D. E. S. Collaboration, *Large synoptic survey telescope: Dark energy science collaboration*, 2012.

[130] O. Doré, J. Bock, M. Ashby, P. Capak, A. Cooray, R. de Putter, T. Eifler, N. Flagey, Y. Gong, S. Habib, K. Heitmann, C. Hirata, W.-S. Jeong, R. Katti, P. Korngut, E. Krause, D.-H. Lee, D. Masters, P. Mauskopf, G. Melnick, B. Mennesson, H. Nguyen, K. Öberg, A. Pullen, A. Raccanelli, R. Smith, Y.-S. Song, V. Tolls, S. Unwin, T. Venumadhav, M. Viero, M. Werner, and M. Zemcov, *Cosmology with the SPHEREX All-Sky Spectral Survey*, arXiv e-prints (Dec., 2014) arXiv:1412.4872, [arXiv:1412.4872].

[131] L. Amendola, S. Appleby, A. Avgoustidis, D. Bacon, T. Baker, M. Baldi, N. Bartolo, A. Blanchard, C. Bonvin, and et al., *Cosmology and fundamental physics with the euclid satellite*, Living Reviews in Relativity **21** (Apr, 2018).

[132] D. Spergel, N. Gehrels, J. Breckinridge, M. Donahue, A. Dressler, B. S. Gaudi, T. Greene, O. Guyon, C. Hirata, J. Kalirai, N. J. Kasdin, W. Moos, S. Perlmutter, M. Postman, B. Rauscher, J. Rhodes, Y. Wang, D. Weinberg, J. Centrella, W. Traub, C. Baltay, J. Colbert, D. Bennett, A. Kiessling, B. Macintosh, J. Merten, M. Mortonson, M. Penny, E. Rozo, D. Savransky, K. Stapelfeldt, Y. Zu, C. Baker, E. Cheng, D. Content, J. Dooley, M. Foote, R. Goullioud, K. Grady, C. Jackson, J. Kruk, M. Levine, M. Melton, C. Peddie, J. Ruffa, and S. Shaklan, *Wide-field infrared survey telescope-astrophysics focused telescope assets wfirst-afta final report*, 2013.

[133] M. G. Hauser and P. J. E. Peebles, *Statistical Analysis of Catalogs of Extragalactic Objects. II. the Abell Catalog of Rich Clusters*, Astrophys.J. **185** (Nov., 1973) 757–786.

[134] P. J. E. Peebles and M. G. Hauser, *Statistical Analysis of Catalogs of Extragalactic Objects. III. The Shane-Wirtanen and Zwicky Catalogs*, Astrophys. J. Supp. **28** (Nov., 1974) 19.

[135] M. Davis and P. J. E. Peebles, *A survey of galaxy redshifts. V. The two-point position and velocity correlations.*, Astrophys.J. **267** (Apr., 1983) 465–482.

[136] W. Saunders, M. Rowan-Robinson, and A. Lawrence, *The spatial correlation function of IRAS galaxies on small and intermediate scales*, Mon. Not. R. Astron. Soc. **258** (Sept., 1992) 134–146.

[137] K. B. Fisher, M. Davis, M. A. Strauss, A. Yahil, and J. P. Huchra, *The Power Spectrum of IRAS Galaxies*, Astrophys.J. **402** (Jan., 1993) 42.

[138] K. B. Fisher, M. Davis, M. A. Strauss, A. Yahil, and J. P. Huchra, *The Power Spectrum of IRAS Galaxies*, Astrophys.J. **402** (Jan., 1993) 42.

[139] J. A. Peacock and D. Nicholson, *The large-scale clustering of radio galaxies.*, Mon. Not. R. Astron. Soc. **253** (Nov., 1991) 307–319.

[140] H. A. Feldman, N. Kaiser, and J. A. Peacock, *Power-Spectrum Analysis of Three-dimensional Redshift Surveys*, Astrophys.J. **426** (May, 1994) 23, [astro-ph/9304022].

[141] C. M. Baugh and G. Efstathiou, *The three-dimensional power spectrum measured from the APM galaxy survey - I. Use of the angular correlation function.*, Mon. Not. R. Astron. Soc. **265** (Nov., 1993) 145–156.

[142] C. M. Baugh, *The real-space correlation function measured from the APM Galaxy Survey*, Mon. Not. R. Astron. Soc. **280** (May, 1996) 267–275, [astro-ph/9512011].

[143] C. J. Miller and D. J. Batuski, *The Power Spectrum of Rich Clusters on Near-Gigaparsec Scales*, Astrophys.J. **551** (Apr., 2001) 635–642, [astro-ph/0002295].

[144] S. Dodelson and E. Gaztañaga, *Inverting the angular correlation function*, Mon. Not. R. Astron. Soc. **312** (Mar., 2000) 774–780, [astro-ph/9906289].

[145] D. Huterer, L. Knox, and R. C. Nichol, *The Angular Power Spectrum of Edinburgh/Durham Southern Galaxy Catalogue Galaxies*, Astrophys.J. **555** (July, 2001) 547–557, [astro-ph/0011069].

[146] D. J. Eisenstein and M. Zaldarriaga, *Correlations in the Spatial Power Spectra Inferred from Angular Clustering: Methods and Application to the Automated Plate Measuring Survey*, Astrophys.J. **546** (Jan., 2001) 2–19, [astro-ph/9912149].

[147] J. A. Peacock, S. Cole, P. Norberg, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. D. Cannon, M. Colless, C. Collins, W. Couch, G. Dalton, K. Deeley, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, W. J. Percival, B. A. Peterson, I. Price, W. Sutherland, and K. Taylor, *A measurement of the cosmological mass density from clustering in the 2dF Galaxy Redshift Survey*, Nature **410** (Mar., 2001) 169–173, [astro-ph/0103143].

[148] S. Dodelson, V. K. Narayanan, M. Tegmark, R. Scranton, T. Budavári, A. Connolly, I. Csabai, D. Eisenstein, J. A. Frieman, J. E. Gunn, L. Hui, B. Jain, D. Johnston, S. Kent, J. Loveday, R. C. Nichol, L. O'Connell, R. Scoccimarro, R. K. Sheth,

A. Stebbins, M. A. Strauss, A. S. Szalay, I. Szapudi, M. S. Vogeley, I. Zehavi, J. Annis, N. A. Bahcall, J. Brinkman, M. Doi, M. Fukugita, G. Hennessy, Ž. Ivezić, G. R. Knapp, P. Kunszt, D. Q. Lamb, B. C. Lee, R. H. Lupton, J. A. Munn, J. Peoples, J. R. Pier, C. Rockosi, D. Schlegel, C. Stoughton, D. L. Tucker, B. Yanny, and D. G. York, *The Three-dimensional Power Spectrum from Angular Clustering of Galaxies in Early Sloan Digital Sky Survey Data*, Astrophys.J. **572** (June, 2002) 140–156, [astro-ph/0107421].

[149] A. J. Connolly, R. Scranton, D. Johnston, S. Dodelson, D. J. Eisenstein, J. A. Frieman, J. E. Gunn, L. Hui, B. Jain, S. Kent, J. Loveday, R. C. Nichol, L. O'Connell, M. Postman, R. Scoccimarro, R. K. Sheth, A. Stebbins, M. A. Strauss, A. S. Szalay, I. Szapudi, M. Tegmark, M. S. Vogeley, I. Zehavi, J. Annis, N. Bahcall, J. Brinkmann, I. Csabai, M. Doi, M. Fukugita, G. S. Hennessy, R. Hindsley, T. Ichikawa, Ž. Ivezić, R. S. J. Kim, G. R. Knapp, P. Kunszt, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. A. McKay, J. Munn, J. Peoples, J. Pier, C. Rockosi, D. Schlegel, C. Stoughton, D. L. Tucker, B. Yanny, and D. G. York, *The Angular Correlation Function of Galaxies from Early Sloan Digital Sky Survey Data*, Astrophys.J. **579** (Nov., 2002) 42–47, [astro-ph/0107417].

[150] W. J. Percival, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, S. Moody, P. Norberg, J. A. Peacock, B. A. Peterson, W. Sutherland, and K. Taylor, *The 2dF Galaxy Redshift Survey: the power spectrum and the matter content of the Universe*, Mon. Not. R. Astron. Soc. **327** (Nov., 2001) 1297–1306, [astro-ph/0105252].

[151] M. Tegmark, M. R. Blanton, M. A. Strauss, F. Hoyle, D. Schlegel, R. Scoccimarro, M. S. Vogeley, D. H. Weinberg, I. Zehavi, A. Berlind, T. Budavari, A. Connolly, D. J. Eisenstein, D. Finkbeiner, J. A. Frieman, J. E. Gunn, A. J. S. Hamilton, L. Hui, B. Jain, D. Johnston, S. Kent, H. Lin, R. Nakajima, R. C. Nichol, J. P. Ostriker, A. Pope, R. Scranton, U. Seljak, R. K. Sheth, A. Stebbins, A. S. Szalay, I. Szapudi, L. Verde, Y. Xu, J. Annis, N. A. Bahcall, J. Brinkmann, S. Burles, F. J. Castander, I. Csabai, J. Loveday, M. Doi, M. Fukugita, I. Gott, J. Richard, G. Hennessy, D. W. Hogg, Ž. Ivezić, G. R. Knapp, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. A. McKay, P. Kunszt, J. A. Munn, L. O'Connell, J. Peoples, J. R. Pier, M. Richmond, C. Rockosi, D. P. Schneider, C. Stoughton, D. L. Tucker, D. E. Vand en Berk, B. Yanny, D. G. York, and SDSS Collaboration, *The Three-Dimensional Power Spectrum of Galaxies from the Sloan Digital Sky Survey*, Astrophys.J. **606** (May, 2004) 702–740, [astro-ph/0310725].

[152] C. Blake, *Power spectrum modelling of galaxy and radio intensity maps including observational effects*, Mon. Not. R. Astron. Soc. **489** (Oct., 2019) 153–167, [arXiv:1902.07439].

[153] P. J. E. Peebles and E. J. Groth, *Statistical analysis of catalogs of extragalactic*

objects. V. Three-point correlation function for the galaxy distribution in the Zwicky catalog., Astrophys.J. **196** (Feb., 1975) 1–11.

[154] H. A. Feldman, J. A. Frieman, J. N. Fry, and R. Scoccimarro, *Constraints on Galaxy Bias, Matter Density, and Primordial Non-Gaussianity from the PSCz Galaxy Redshift Survey*, Phys. Rev. Lett. **86** (Feb., 2001) 1434–1437, [astro-ph/0010205].

[155] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, *The Bispectrum of IRAS Redshift Catalogs*, Astrophys.J. **546** (Jan., 2001) 652–664, [astro-ph/0004087].

[156] A. Cooray and W. Hu, *Weak Gravitational Lensing Bispectrum*, Astrophys.J. **548** (Feb., 2001) 7–18, [astro-ph/0004151].

[157] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, *Cosmology and the bispectrum*, Phys. Rev. D **74** (July, 2006) 023522, [astro-ph/0604505].

[158] L. Verde, A. F. Heavens, W. J. Percival, S. Matarrese, C. M. Baugh, J. Bland-Hawthorn, T. Bridges, R. Cannon, S. Cole, M. Colless, C. Collins, W. Couch, G. Dalton, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, C. S. Frenk, K. Glazebrook, C. Jackson, O. Lahav, I. Lewis, S. Lumsden, S. Maddox, D. Madgwick, P. Norberg, J. A. Peacock, B. A. Peterson, W. Sutherland, and K. Taylor, *The 2dF Galaxy Redshift Survey: the bias of galaxies and the density of the Universe*, Mon. Not. R. Astron. Soc. **335** (Sept., 2002) 432–440, [astro-ph/0112161].

[159] F. A. Marín, C. Blake, G. B. Poole, C. K. McBride, S. Brough, M. Colless, C. Contreras, W. Couch, D. J. Croton, S. Croom, T. Davis, M. J. Drinkwater, K. Forster, D. Gilbank, M. Gladders, K. Glazebrook, B. Jelliffe, R. J. Jurek, I. h. Li, B. Madore, D. C. Martin, K. Pimbblet, M. Pracy, R. Sharp, E. Wisnioski, D. Woods, T. K. Wyder, and H. K. C. Yee, *The WiggleZ Dark Energy Survey: constraining galaxy bias and cosmic growth with three-point correlation functions*, Mon. Not. R. Astron. Soc. **432** (July, 2013) 2654–2668, [arXiv:1303.6644].

[160] H. Gil-Marín, J. Noreña, L. Verde, W. J. Percival, C. Wagner, M. Manera, and D. P. Schneider, *The power spectrum and bispectrum of SDSS DR11 BOSS galaxies - I. Bias and gravity*, Mon. Not. R. Astron. Soc. **451** (July, 2015) 539–580, [arXiv:1407.5668].

[161] Z. Slepian and D. J. Eisenstein, *Computing the three-point correlation function of galaxies in O(N$\hat{2}$) time*, Mon. Not. R. Astron. Soc. **454** (Dec., 2015) 4142–4158, [arXiv:1506.02040].

[162] Z. Slepian, D. J. Eisenstein, F. Beutler, A. J. Cuesta, J. Ge, H. Gil-Marín, S. Ho, F.-S. Kitaura, C. K. McBride, R. C. Nichol, W. J. Percival, S. Rodríguez-Torres, A. J. Ross, R. Scoccimarro, H.-J. Seo, J. Tinker, R. Tojeiro, and M. Vargas-Magaña, *The large-scale 3-point correlation function of the SDSS BOSS DR12 CMASS galaxies*, arXiv e-prints (Dec., 2015) arXiv:1512.02231, [arXiv:1512.02231].

[163] B. Jain and L. Van Waerbeke, *Statistics of Dark Matter Halos from Gravitational Lensing*, Astrophys. J. Lett. **530** (Feb., 2000) L1–L4, [astro-ph/9910459].

[164] L. Marian, S. Hilbert, R. E. Smith, P. Schneider, and V. Desjacques, *Measuring Primordial Non-gaussianity Through Weak-lensing Peak Counts*, Astrophys. J. Lett. **728** (Feb., 2011) L13, [arXiv:1010.5242].

[165] J. Liu, A. Petri, Z. Haiman, L. Hui, J. M. Kratochvil, and M. May, *Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data*, Phys. Rev. D **91** (Mar., 2015) 063507, [arXiv:1412.0757].

[166] S. D. M. White, *The hierarchy of correlation functions and its relation to other measures of galaxy clustering.*, Mon. Not. R. Astron. Soc. **186** (Jan., 1979) 145–154.

[167] J. N. Fry, *Statistics of Voids in Hierarchical Universes*, Astrophys.J. **306** (July, 1986) 358.

[168] R. Biswas, E. Alizadeh, and B. D. Wandelt, *Voids as a precision probe of dark energy*, Phys. Rev. D **82** (July, 2010) 023002, [arXiv:1002.0014].

[169] E. G. P. Bos, R. van de Weygaert, K. Dolag, and V. Pettorino, *The darkness that shaped the void: dark energy and cosmic voids*, Mon. Not. R. Astron. Soc. **426** (Oct., 2012) 440–461, [arXiv:1205.4238].

[170] A. Pisani, P. M. Sutter, N. Hamaus, E. Alizadeh, R. Biswas, B. D. Wandelt, and C. M. Hirata, *Counting voids to probe dark energy*, Phys. Rev. D **92** (Oct., 2015) 083531, [arXiv:1503.07690].

[171] S. Nadathur and S. Hotchkiss, *The nature of voids - I. Watershed void finders and their connection with theoretical models*, Mon. Not. R. Astron. Soc. **454** (Dec., 2015) 2228–2241, [arXiv:1504.06510].

[172] F. Leclercq, J. Jasche, P. M. Sutter, N. Hamaus, and B. Wandelt, *Dark matter voids in the SDSS galaxy survey*, JCAP **2015** (Mar., 2015) 047, [arXiv:1410.0355].

[173] O. Friedrich, D. Gruen, J. DeRose, D. Kirk, E. Krause, T. McClintock, E. Rykoff, S. Seitz, R. Wechsler, G. Bernstein, and et al., *Density split statistics: Joint model of counts and lensing in cells*, Physical Review D **98** (Jul, 2018).

[174] A. Repp and I. Szapudi, *Precision prediction of the log power spectrum*, Monthly Notices of the Royal Astronomical Society: Letters **464** (Sep, 2016) L21–L25.

[175] R. K. Sheth, *The halo-model description of marked statistics*, Mon. Not. R. Astron. Soc. **364** (Dec., 2005) 796–806, [astro-ph/0511772].

[176] M. White, *A marked correlation function for constraining modified gravity models*, Journal of Cosmology and Astroparticle Physics **2016** (Nov, 2016) 057–057.

[177] M. White and N. Padmanabhan, *Breaking halo occupation degeneracies with marked statistics*, Mon. Not. R. Astron. Soc. **395** (June, 2009) 2381–2384, [arXiv:0812.4288].

[178] O. H. E. Philcox, E. Massara, and D. N. Spergel, *What does the Marked Power Spectrum Measure? Insights from Perturbation Theory*, arXiv e-prints (June, 2020) arXiv:2006.10055, [arXiv:2006.10055].

[179] I. Gott, J. Richard, D. H. Weinberg, and A. L. Melott, *A Quantitative Approach to the Topology of Large-Scale Structure*, Astrophys.J. **319** (Aug., 1987) 1.

[180] J. M. Kratochvil, E. A. Lim, S. Wang, Z. Haiman, M. May, and K. Huffenberger, *Probing cosmology with weak lensing Minkowski functionals*, Phys. Rev. D **85** (May, 2012) 103513, [arXiv:1109.6334].

[181] C. Hikage, E. Komatsu, and T. Matsubara, *Primordial Non-Gaussianity and Analytical Formula for Minkowski Functionals of the Cosmic Microwave Background and Large-Scale Structure*, Astrophys.J. **653** (Dec., 2006) 11–26, [astro-ph/0607284].

[182] D. Munshi, L. van Waerbeke, J. Smidt, and P. Coles, *From weak lensing to non-Gaussianity via Minkowski functionals*, Mon. Not. R. Astron. Soc. **419** (Jan., 2012) 536–555, [arXiv:1103.1876].

[183] A. Petri, J. Liu, Z. Haiman, M. May, L. Hui, and J. M. Kratochvil, *Emulating the CFHTLenS weak lensing data: Cosmological constraints from moments and Minkowski functionals*, Phys. Rev. D **91** (May, 2015) 103511, [arXiv:1503.06214].

[184] B. Mawdsley, D. Bacon, C. Chang, P. Melchior, E. Rozo, S. Seitz, N. Jeffrey, M. Gatti, E. Gaztanaga, D. Gruen, W. G. Hartley, B. Hoyle, S. Samuroff, E. Sheldon, M. A. Troxel, J. Zuntz, T. M. C. Abbott, J. Annis, E. Bertin, S. L. Bridle, D. Brooks, E. Buckley-Geer, D. L. Burke, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, L. N. da Costa, J. De Vicente, S. Desai, H. T. Diehl, P. Doel, A. E. Evrard, B. Flaugher, P. Fosalba, J. Frieman, J. García-Bellido, D. W. Gerdes, R. A. Gruendl, J. Gschwend, G. Gutierrez, D. L. Hollowood, K. Honscheid, D. J. James, M. Jarvis, T. Jeltema, K. Kuehn, N. Kuropatkin, M. Lima, M. A. G. Maia, J. L. Marshall, R. Miquel, A. A. Plazas, A. Roodman, E. Sanchez, V. Scarpine, S. Serrano, I. Sevilla-Noarbe, M. Smith, R. C. Smith, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, D. L. Tucker, V. Vikram, A. R. Walker, and DES Collaboration, *Dark Energy Survey Year 1 Results: Wide-field mass maps via forward fitting in harmonic space*, Mon. Not. R. Astron. Soc. **493** (Mar., 2020) 5662–5679, [arXiv:1905.12682].

[185] E. Allys, T. Marchand, J. F. Cardoso, F. Villaescusa-Navarro, S. Ho, and S. Mallat, *New interpretable statistics for large scale structure analysis and generation*, 2020.

[186] S. Cheng, Y.-S. Ting, B. Ménard, and J. Bruna, *A new approach to observational cosmology using the scattering transform*, 2020.

[187] M. S. Vogeley, Toward High-Precision Measures of Large-Scale Structure, vol. 231 of Astrophysics and Space Science Library, p. 395. 1998.

[188] R. Scranton, D. Johnston, S. Dodelson, J. A. Frieman, A. Connolly, D. J. Eisenstein, J. E. Gunn, L. Hui, B. Jain, S. Kent, J. Loveday, V. Narayanan, R. C. Nichol, L. O'Connell, R. Scoccimarro, R. K. Sheth, A. Stebbins, M. A. Strauss, A. S. Szalay, I. Szapudi, M. Tegmark, M. Vogeley, I. Zehavi, J. Annis, N. A. Bahcall, J. Brinkman, I. Csabai, R. Hindsley, Z. Ivezic, R. S. J. Kim, G. R. Knapp, D. Q. Lamb, B. C. Lee, R. H. Lupton, T. McKay, J. Munn, J. Peoples, J. Pier, G. T. Richards, C. Rockosi, D. Schlegel, D. P. Schneider, C. Stoughton, D. L. Tucker, B. Yanny, and D. G. York, *Analysis of Systematic Effects and Statistical Uncertainties in Angular Clustering of Galaxies from Early Sloan Digital Sky Survey Data*, Astrophys.J. **579** (Nov., 2002) 48–75, [astro-ph/0107416].

[189] T. Goto, I. Szapudi, and B. R. Granett, *Cross-correlation of WISE Galaxies with the Cosmic Microwave Background*, Mon.Not.Roy.Astron.Soc. **422** (2012) L77–L81, [arXiv:1202.5306].

[190] S. Ho, A. Cuesta, H.-J. Seo, R. de Putter, A. J. Ross, M. White, N. Padmanabhan, S. Saito, D. J. Schlegel, E. Schlafly, U. Seljak, C. Hernández-Monteagudo, A. G. Sánchez, W. J. Percival, M. Blanton, R. Skibba, D. Schneider, B. Reid, O. Mena, M. Viel, D. J. Eisenstein, F. Prada, B. A. Weaver, N. Bahcall, D. Bizyaev, H. Brewinton, J. Brinkman, L. N. da Costa, J. R. Gott, E. Malanushenko, V. Malanushenko, B. Nichol, D. Oravetz, K. Pan, N. Palanque-Delabrouille, N. P. Ross, A. Simmons, F. de Simoni, S. Snedden, and C. Yeche, *CLUSTERING OF SLOAN DIGITAL SKY SURVEY III PHOTOMETRIC LUMINOUS GALAXIES: THE MEASUREMENT, SYSTEMATICS, AND COSMOLOGICAL IMPLICATIONS*, The Astrophysical Journal **761** (nov, 2012) 14.

[191] **DES** Collaboration, B. Leistedt et al., *Mapping and simulating systematics due to spatially-varying observing conditions in DES Science Verification data*, Astrophys. J. Suppl. **226** (2016), no. 2 24, [arXiv:1507.05647].

[192] T. Delubac, A. Raichoor, J. Comparat, S. Jouvel, J.-P. Kneib, C. Yèche, H. Zou, J. R. Brownstein, F. B. Abdalla, K. Dawson, and et al., *The sdss-iv eboss: emission line galaxy catalogues at $z \approx 0.8$ and study of systematic errors in the angular clustering*, Monthly Notices of the Royal Astronomical Society **465** (Oct, 2016) 1831–1846.

[193] C. B. Morrison and H. Hildebrandt, *Mitigating systematic errors in angular correlation function measurements from wide field surveys*, Monthly Notices of the Royal Astronomical Society **454** (Oct, 2015) 3121–3133.

[194] A. Prakash, T. C. Licquia, J. A. Newman, A. J. Ross, A. D. Myers, K. S. Dawson, J.-P. Kneib, W. J. Percival, J. E. Bautista, J. Comparat, J. L. Tinker, D. J. Schlegel, R. Tojeiro, S. Ho, D. Lang, S. M. Rao, C. K. McBride, G. B. Zhu, J. R. Brownstein, S. Bailey, A. S. Bolton, T. Delubac, V. Mariappan, M. R. Blanton, B. Reid, D. P. Schneider, H.-J. Seo, A. C. Rosell, and F. Prada, *THE SDSS-IV EXTENDED*

BARYON OSCILLATION SPECTROSCOPIC SURVEY: LUMINOUS RED GALAXY TARGET SELECTION, The Astrophysical Journal Supplement Series **224** (jun, 2016) 34.

[195] B. Kalus, W. J. Percival, D. J. Bacon, and L. Samushia, *Unbiased contaminant removal for 3D galaxy power spectrum measurements*, Monthly Notices of the Royal Astronomical Society **463** (08, 2016) 467–476, [http://oup.prod.sis.lan/mnras/article-pdf/463/1/467/18472203/stw2008.pdf].

[196] J. E. Bautista, M. Vargas-Magaña, K. S. Dawson, W. J. Percival, J. Brinkmann, J. Brownstein, B. Camacho, J. Comparat, H. Gil-Marín, E.-M. Mueller, and et al., *The sdss-iv extended baryon oscillation spectroscopic survey: Baryon acoustic oscillations at redshift of 0.72 with the dr14 luminous red galaxy sample*, The Astrophysical Journal **863** (Aug, 2018) 110.

[197] B. Kalus, W. J. Percival, D. J. Bacon, E. M. Mueller, L. Samushia, L. Verde, A. J. Ross, and J. L. Bernal, *A map-based method for eliminating systematic modes from galaxy clustering power spectra with application to BOSS*, Mon. Not. Roy. Astron. Soc. **482** (2019), no. 1 453–470, [arXiv:1806.02789].

[198] H. Awan and E. Gawiser, *Angular correlation function estimators accounting for contamination from probabilistic distance measurements*, 2019.

[199] E. Kitanidis, M. White, Y. Feng, D. Schlegel, J. Guy, A. Dey, M. Landriau, D. Brooks, M. Levi, J. Moustakas, F. Prada, G. Tarle, and B. A. Weaver, *Imaging systematics and clustering of desi main targets*, Monthly Notices of the Royal Astronomical Society **496** (2020) 2262–2291.

[200] M. Rezaie, H. Seo, A. Ross, and R. C. Bunescu, *Improving galaxy clustering measurements with deep learning: analysis of the decals dr7 data*, Monthly Notices of the Royal Astronomical Society **495** (2020) 1613–1640.

[201] E. L. Wagoner, E. Rozo, and X. Fang, *Linear systematics mitigation in galaxy clustering in the dark energy survey year 1 data*, 2020. In Preparation.

[202] H. Kong, K. J. Burleigh, A. Ross, J. Moustakas, C.-H. Chuang, J. Comparat, A. de Mattia, H. du Mas des Bourboux, K. Honscheid, S. Lin, A. Raichoor, G. Rossi, and C. Zhao, *Removing imaging systematics from galaxy clustering measurements with obiwan : Application to the sdss-iv extended baryon oscillation spectroscopic survey emission line galaxy sample*, 2020.

[203] A. J. Ross, J. Bautista, R. Tojeiro, S. Alam, S. Bailey, E. Burtin, J. Comparat, K. S. Dawson, A. de Mattia, H. du Mas des Bourboux, H. Gil-Marín, J. Hou, H. Kong, B. W. Lyke, F. G. Mohammad, J. Moustakas, E.-M. Mueller, A. D. Myers, W. J. Percival, A. Raichoor, M. Rezaie, H.-J. Seo, A. Smith, J. L. Tinker, P. Zarrouk, C. Zhao, G.-B. Zhao, D. Bizyaev, J. Brinkmann, J. R. Brownstein, A. C. Rosell, S. Chabanier, P. D. Choi, C.-H. Chuang, I. Cruz-Gonzalez, A. de la Macorra, S. de la Torre, S. Escoffier, S. Fromenteau, A. Higley, E. Jullo, J.-P. Kneib, J. N. McLane,

A. Muñoz-Gutiérrez, R. Neveux, J. A. Newman, C. Nitschelm, N. Palanque-Delabrouille, R. Paviot, A. R. Pullen, G. Rossi, V. Ruhlmann-Kleider, D. P. Schneider, M. V. Magaña, M. Vivek, and Y. Zhang, *The completed sdss-iv extended baryon oscillation spectroscopic survey: Large-scale structure catalogs for cosmological analysis*, 2020.

[204] J. Jasche and F. S. Kitaura, *Fast Hamiltonian sampling for large scale structure inference*, Mon. Not. Roy. Astron. Soc. **407** (2010) 29, [arXiv:0911.2496].

[205] J. Jasche and B. D. Wandelt, *Bayesian physical reconstruction of initial conditions from large scale structure surveys*, Mon. Not. Roy. Astron. Soc. **432** (2013) 894, [arXiv:1203.3639].

[206] F.-S. Kitaura, *The Initial Conditions of the Universe from Constrained Simulations*, Mon. Not. Roy. Astron. Soc. **429** (2013) 84, [arXiv:1203.4184].

[207] H. Wang, H. J. Mo, X. Yang, Y. P. Jing, and W. P. Lin, *ELUCID - Exploring the Local Universe with reConstructed Initial Density field I: Hamiltonian Markov Chain Monte Carlo Method with Particle Mesh Dynamics*, Astrophys. J. **794** (2014), no. 1 94, [arXiv:1407.3451].

[208] J. Jasche, F. Leclercq, and B. D. Wandelt, *Past and present cosmic structure in the SDSS DR7 main sample*, JCAP **1501** (2015), no. 01 036, [arXiv:1409.6308].

[209] H. Wang, H. J. Mo, X. Yang, Y. Zhang, J. Shi, Y. P. Jing, C. Liu, S. Li, X. Kang, and Y. Gao, *ELUCID - Exploring the Local Universe with reConstructed Initial Density field III: Constrained Simulation in the SDSS Volume*, Astrophys. J. **831** (2016), no. 2 164, [arXiv:1608.01763].

[210] C. Modi, M. White, A. Slosar, and E. Castorina, *Reconstructing large-scale structure with neutral hydrogen surveys*, JCAP **1911** (2019), no. 11 023, [arXiv:1907.02330].

[211] N. Porqueres, D. Kodi Ramanah, J. Jasche, and G. Lavaux, *Explicit bayesian treatment of unknown foreground contaminations in galaxy surveys*, Astronomy & Astrophysics **624** (Apr, 2019) A115.

[212] **DES** Collaboration, J. Elvin-Poole et al., *Dark Energy Survey year 1 results: Galaxy clustering for combined probes*, Phys. Rev. **D98** (2018), no. 4 042006, [arXiv:1708.01536].

[213] **BOSS** Collaboration, A. J. Ross et al., *The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: Observational systematics and baryon acoustic oscillations in the correlation function*, Mon. Not. Roy. Astron. Soc. **464** (2017), no. 1 1168–1191, [arXiv:1607.03145].

[214] P. Laurent et al., *Clustering of quasars in SDSS-IV eBOSS : study of potential systematics and bias determination*, JCAP **1707** (2017), no. 07 017, [arXiv:1705.04718].

[215] M. Ata et al., *The clustering of the SDSS-IV extended Baryon Oscillation Spectroscopic Survey DR14 quasar sample: first measurement of baryon acoustic oscillations between redshift 0.8 and 2.2*, Mon. Not. Roy. Astron. Soc. **473** (2018), no. 4 4773–4794, [arXiv:1705.06373].

[216] A. Nicola, A. Refregier, and A. Amara, *Integrated approach to cosmology: Combining CMB, large-scale structure and weak lensing*, Phys. Rev. **D94** (2016), no. 8 083517, [arXiv:1607.01014].

[217] **DES** Collaboration, M. Crocce et al., *Galaxy clustering, photometric redshifts and diagnosis of systematics in the DES Science Verification data*, Mon. Not. Roy. Astron. Soc. **455** (2016), no. 4 4301–4324, [arXiv:1507.05360].

[218] W. J. Percival, *Recent Developments in the Analysis of Galaxy Surveys*, 2018. arXiv:1810.04263.

[219] A. Nicola, D. Alonso, J. Sánchez, A. Slosar, H. Awan, A. Broussard, J. Dunkley, Z. Gomes, E. Gawiser, R. Mandelbaum, H. Miyatake, J. A. Newman, I. Sevilla, S. Skinner, and E. Wagoner, *Tomographic galaxy clustering with the subaru hyper suprime-cam first year public data release*, 2019.

[220] G. B. Rybicki and W. H. Press, *Interpolation, realization, and reconstruction of noisy, irregularly sampled data*, Astrophys. J. **398** (1992) 169–176.

[221] E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen, *Master of the cosmic microwave background anisotropy power spectrum: a fast method for statistical analysis of large and complex cosmic microwave background data sets*, Astrophys. J. **567** (2002) 2, [astro-ph/0105302].

[222] D. P. Foster and E. I. George, *The risk inflation criterion for multiple regression*, Ann. Statist. **22** (12, 1994) 1947–1975.

[223] J.-L. Starck, D. L. Donoho, M. J. Fadili, and A. Rassat, *Sparsity and the bayesian perspective*, Astronomy & Astrophysics **552** (Apr, 2013) A133.

[224] J. Bobin, J. Starck, J. Fadili, and Y. Moudden, *Sparsity and morphological diversity in blind source separation*, IEEE Transactions on Image Processing **16** (Nov, 2007) 2662–2674.

[225] J. Bobin, J.-L. Starck, F. Sureau, and S. Basak, *Sparse component separation for accurate cosmic microwave background estimation*, Astronomy & Astrophysics **550** (Jan, 2013) A73.

[226] S. Wagner-Carena, M. Hopkins, A. D. Rivero, and C. Dvorkin, *A novel cmb component separation method: Hierarchical generalized morphological component analysis*, 2019.

[227] A. Leonard, F. Lanusse, and J.-L. Starck, *Glimpse: accurate 3d weak lensing reconstructions using sparsity*, Monthly Notices of the Royal Astronomical Society **440** (Mar, 2014) 1281–1294.

[228] Lanusse, F., Starck, J.-L., Leonard, A., and Pires, S., *High resolution weak lensing mass mapping combining shear and flexion*, A&A **591** (2016) A2.

[229] N. Jeffrey, F. B. Abdalla, O. Lahav, F. Lanusse, J. L. Starck, A. Leonard, D. Kirk, C. Chang, E. Baxter, T. Kacprzak, S. Seitz, V. Vikram, L. Whiteway, T. M. C. Abbott, S. Allam, S. Avila, E. Bertin, D. Brooks, A. Carnero Rosell, M. Carrasco Kind, J. Carretero, F. J. Castander, M. Crocce, C. E. Cunha, C. B. D'Andrea, L. N. da Costa, C. Davis, J. De Vicente, S. Desai, P. Doel, T. F. Eifler, A. E. Evrard, B. Flaugher, P. Fosalba, J. Frieman, J. García-Bellido, D. W. Gerdes, D. Gruen, R. A. Gruendl, J. Gschwend, G. Gutierrez, W. G. Hartley, K. Honscheid, B. Hoyle, D. J. James, M. Jarvis, K. Kuehn, M. Lima, H. Lin, M. March, P. Melchior, F. Menanteau, R. Miquel, A. A. Plazas, K. Reil, A. Roodman, E. Sanchez, V. Scarpine, M. Schubnell, I. Sevilla-Noarbe, M. Smith, M. Soares-Santos, F. Sobreira, E. Suchyta, M. E. C. Swanson, G. Tarle, D. Thomas, A. R. Walker, and DES Collaboration, *Improving weak lensing mass map reconstructions using Gaussian and sparsity priors: application to DES SV*, Mon. Not. R. Astron. Soc. **479** (Sep, 2018) 2871–2888, [arXiv:1801.08945].

[230] H. Zou and T. Hastie, *Regularization and variable selection via the elastic net*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **67** (2005), no. 2 301–320.

[231] R. D. Cook, *Detection of influential observation in linear regression*, Technometrics **19** (1977), no. 1 15–18.

[232] R. D. Cook, *Influential observations in linear regression*, Journal of the American Statistical Association **74** (1979), no. 365 169–174.

[233] **Planck** Collaboration, A. Abergel et al., *Planck 2013 results. XI. All-sky model of thermal dust emission*, Astron. Astrophys. **571** (2014) A11, [arXiv:1312.1300].

[234] J. Lesgourgues, *The cosmic linear anisotropy solving system (class) i: Overview*, arXiv: Instrumentation and Methods for Astrophysics (2011).

[235] Y. Benjamini and Y. Hochberg, *Controlling the false discovery rate: A practical and powerful approach to multiple testing*, Journal of the Royal Statistical Society. Series B (Methodological) **57** (1995), no. 1 289–300.

[236] A. H. Guth, *The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems*, Phys. Rev. **D23** (1981) 347–356.

[237] A. D. Linde, *A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems*, Phys. Lett. **B108** (1982) 389–393.

[238] A. Albrecht and P. J. Steinhardt, *Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking*, Phys. Rev. Lett. **48** (1982) 1220–1223.

[239] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, *Spontaneous Creation of Almost Scale - Free Density Perturbations in an Inflationary Universe*, Phys. Rev. **D28** (1983) 679.

[240] S. Dodelson, *Coherent Phase Argument for Inflation*, in Neutrinos, Flavor Physics, and Precision Cosmology (J. F. Nieves and R. R. Volkas, eds.), vol. 689 of American Institute of Physics Conference Series, pp. 184–196, Oct., 2003. hep-ph/0309057.

[241] P. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, et al., *Planck 2015 results-xx. constraints on inflation*, Astronomy & Astrophysics **594** (2016) A20.

[242] G. Cabass, E. Di Valentino, A. Melchiorri, E. Pajer, and J. Silk, *Constraints on the running of the running of the scalar tilt from CMB anisotropies and spectral distortions*, Phys. Rev. **D94** (2016), no. 2 023523, [arXiv:1605.00209].

[243] A. Kosowsky and M. S. Turner, *CBR anisotropy and the running of the scalar spectral index*, Phys. Rev. **D52** (1995) R1739–R1743, [astro-ph/9504071].

[244] J. Garcia-Bellido and D. Roest, *Large-N running of the spectral index of inflation*, Phys. Rev. **D89** (2014), no. 10 103527, [arXiv:1402.2059].

[245] M. Escudero, H. Ramírez, L. Boubekeur, E. Giusarma, and O. Mena, *The present and future of the most favoured inflationary models after Planck 2015*, JCAP **1602** (2016), no. 02 020, [arXiv:1509.05419].

[246] R. Easther and H. Peiris, *Implications of a Running Spectral Index for Slow Roll Inflation*, JCAP **0609** (2006) 010, [astro-ph/0604214].

[247] J. P. Vieira, C. T. Byrnes, and A. Lewis, *Can power spectrum observations rule out slow-roll inflation?*, Journal of Cosmology and Astroparticle Physics **2018** (2018), no. 01 019.

[248] T. Kobayashi and F. Takahashi, *Running Spectral Index from Inflation with Modulations*, JCAP **1101** (2011) 026, [arXiv:1011.3988].

[249] M. Czerny, T. Kobayashi, and F. Takahashi, *Running Spectral Index from Large-field Inflation with Modulations Revisited*, Phys. Lett. **B735** (2014) 176–180, [arXiv:1403.4589].

[250] M. Drees and E. Erfani, *Running Spectral Index and Formation of Primordial Black Hole in Single Field Inflation Models*, JCAP **1201** (2012) 035, [arXiv:1110.6052].

[251] B. Carr, F. Kuhnel, and M. Sandstad, *Primordial Black Holes as Dark Matter*, Phys. Rev. **D94** (2016), no. 8 083504, [arXiv:1607.06077].

[252] Z. Huang, L. Verde, and F. Vernizzi, *Constraining inflation with future galaxy redshift surveys*, JCAP **1204** (2012) 005, [arXiv:1201.5955].

[253] L. Amendola, S. Appleby, D. Bacon, T. Baker, M. Baldi, N. Bartolo, A. Blanchard, C. Bonvin, S. Borgani, E. Branchini, et al., *Cosmology and fundamental physics with the euclid satellite*, Living Reviews in Relativity **16** (2013), no. 1 6.

[254] T. Basse, J. Hamann, S. Hannestad, and Y. Y. Y. Wong, *Getting leverage on inflation with a large photometric redshift survey*, JCAP **1506** (2015), no. 06 042, [arXiv:1409.3469].

[255] F. Bernardeau, S. Colombi, E. Gaztanaga, and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, Phys. Rept. **367** (2002) 1–248, [astro-ph/0112551].

[256] R. E. Smith, J. A. Peacock, A. Jenkins, S. White, C. Frenk, F. Pearce, P. A. Thomas, G. Efstathiou, and H. Couchman, *Stable clustering, the halo model and non-linear cosmological power spectra*, Monthly Notices of the Royal Astronomical Society **341** (2003), no. 4 1311–1332.

[257] K. Heitmann, D. Higdon, M. White, S. Habib, B. J. Williams, E. Lawrence, and C. Wagner, *The coyote universe. ii. cosmological models and precision emulation of the nonlinear matter power spectrum*, The Astrophysical Journal **705** (2009), no. 1 156.

[258] K. Heitmann, M. White, C. Wagner, S. Habib, and D. Higdon, *The coyote universe. i. precision determination of the nonlinear matter power spectrum*, The Astrophysical Journal **715** (2010), no. 1 104.

[259] E. Lawrence, K. Heitmann, M. White, D. Higdon, C. Wagner, S. Habib, and B. Williams, *The coyote universe. iii. simulation suite and precision emulator for the nonlinear matter power spectrum*, The Astrophysical Journal **713** (2010), no. 2 1322.

[260] D. H. Rudd, A. R. Zentner, and A. V. Kravtsov, *Effects of Baryons and Dissipation on the Matter Power Spectrum*, Astrophys. J. **672** (2008) 19–32, [astro-ph/0703741].

[261] M. P. van Daalen, J. Schaye, C. M. Booth, and C. D. Vecchia, *The effects of galaxy formation on the matter power spectrum: A challenge for precision cosmology*, Mon. Not. Roy. Astron. Soc. **415** (2011) 3649–3665, [arXiv:1104.1174].

[262] N. E. Chisari, M. L. A. Richardson, J. Devriendt, Y. Dubois, A. Schneider, M. C. Brun, Amandine Le, R. S. Beckmann, S. Peirani, A. Slyz, and C. Pichon, *The impact of baryons on the matter power spectrum from the Horizon-AGN cosmological hydrodynamical simulation*, arXiv:1801.08559.

[263] R. Scoccimarro, M. Zaldarriaga, and L. Hui, *Power spectrum correlations induced by nonlinear clustering*, Astrophys. J. **527** (1999) 1, [astro-ph/9901099].

[264] P. A. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. Banday, R. Barreiro, J. Bartlett, N. Bartolo, et al., *Planck 2015 results-xiii. cosmological parameters*, Astronomy & Astrophysics **594** (2016) A13.

[265] K. N. Abazajian, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, K. S. Arnold, C. Baccigalupi, J. G. Bartlett, N. Battaglia, B. A. Benson, et al., *Cmb-s4 science book*, arXiv preprint arXiv:1610.02743 (2016).

[266] N. Aghanim, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. Banday, R. Barreiro, N. Bartolo, S. Basak, R. Battye, et al., *Planck intermediate results-xlvi. reduction of large-scale systematic effects in hfi polarization maps and estimation of the reionization optical depth*, Astronomy & Astrophysics **596** (2016) A107.

[267] V. Böhm, M. Schmittfull, and B. D. Sherwin, *Bias to CMB lensing measurements from the bispectrum of large-scale structure*, Phys. Rev. **D94** (2016), no. 4 043519, [arXiv:1605.01392].

[268] A. Lewis and G. Pratten, *Effect of lensing non-Gaussianity on the CMB power spectra*, JCAP **1612** (2016), no. 12 003, [arXiv:1608.01263].

[269] M. Tegmark, M. A. Strauss, M. R. Blanton, K. Abazajian, S. Dodelson, H. Sandvik, X. Wang, D. H. Weinberg, I. Zehavi, N. A. Bahcall, et al., *Cosmological parameters from sdss and wmap*, Physical Review D **69** (2004), no. 10 103501.

[270] L. Samushia, W. J. Percival, and A. Raccanelli, *Interpreting large-scale redshift-space distortion measurements*, Monthly Notices of the Royal Astronomical Society **420** (2012), no. 3 2102–2119.

[271] S. Alam, M. Ata, S. Bailey, F. Beutler, D. Bizyaev, J. A. Blazek, A. S. Bolton, J. R. Brownstein, A. Burden, C.-H. Chuang, et al., *The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample*, Monthly Notices of the Royal Astronomical Society **470** (2017), no. 3 2617–2652.

[272] A. Lewis, A. Challinor, and A. Lasenby, *Efficient computation of cmb anisotropies in closed frw models, 2000*, Astrophys. J **538** (2000) 473.

[273] D. Blas, J. Lesgourgues, and T. Tram, *The cosmic linear anisotropy solving system (class). part ii: approximation schemes*, Journal of Cosmology and Astroparticle Physics **2011** (2011), no. 07 034.

[274] U. Seljak, *Analytic model for galaxy and dark matter clustering*, Mon. Not. Roy. Astron. Soc. **318** (2000) 203, [astro-ph/0001493].

[275] J. Peacock and R. Smith, *Halo occupation numbers and galaxy bias*, Mon.Not.Roy.Astron.Soc. **318** (2000) 1144, [astro-ph/0005010].

[276] A. Cooray and R. K. Sheth, *Halo models of large scale structure*, Phys. Rept. **372** (2002) 1–129, [astro-ph/0206508].

[277] H.-J. Seo and D. J. Eisenstein, *Improved forecasts for the baryon acoustic oscillations and cosmological distance scale*, The Astrophysical Journal **665** (2007), no. 1 14.

[278] A. E. Evrard, J. Bialek, M. Busha, et al., *Virial Scaling of Massive Dark Matter Halos: Why Clusters Prefer a High Normalization Cosmology*, Astrophys.J. **672** (Jan., 2008) 122–137, [astro-ph/0].

[279] J. Bielefeld, D. Huterer, and E. V. Linder, *Cosmological leverage from the matter power spectrum in the presence of baryon and nonlinear effects*, Journal of Cosmology and Astroparticle Physics **2015** (2015), no. 05 023.

[280] H.-Y. Wu and D. Huterer, *The impact of systematic uncertainties in n-body simulations on the precision cosmology from galaxy clustering: a halo model approach*, Monthly Notices of the Royal Astronomical Society **434** (2013), no. 3 2556–2571.

[281] A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. A. Prieto, J. Annis, S. Bailey, C. Balland, O. Ballester, et al., *The desi experiment part i: Science, targeting, and survey design*, arXiv preprint arXiv:1611.00036 (2016).

[282] D. Spergel, N. Gehrels, C. Baltay, D. Bennett, J. Breckinridge, M. Donahue, A. Dressler, B. Gaudi, T. Greene, O. Guyon, et al., *Wide-field infrarred survey telescope-astrophysics focused telescope assets wfirst-afta 2015 report*, arXiv preprint arXiv:1503.03757 (2015).

[283] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Augueres, J. Brinchmann, R. Cole, M. Cropper, C. Dabin, L. Duvet, A. Ealet, et al., *Euclid definition study report*, arXiv preprint arXiv:1110.3193 (2011).

[284] A. Font-Ribera, P. McDonald, N. Mostek, B. A. Reid, H.-J. Seo, and A. Slosar, *Desi and other dark energy experiments in the era of neutrino mass measurements*, Journal of Cosmology and Astroparticle Physics **2014** (2014), no. 05 023.

[285] A. Albrecht, L. Amendola, G. Bernstein, D. Clowe, D. Eisenstein, L. Guzzo, C. Hirata, D. Huterer, R. Kirshner, E. Kolb, et al., *Findings of the joint dark energy mission figure of merit science working group*, arXiv preprint arXiv:0901.0721 (2009).

[286] M. Tegmark, A. N. Taylor, and A. F. Heavens, *Karhunen-Loève Eigenvalue Problems in Cosmology: How Should We Tackle Large Data Sets?*, Astrophys.J. **480** (May, 1997) 22–35, [astro-ph/9603021].

[287] U. Seljak, *Measuring polarization in the cosmic microwave background*, The Astrophysical Journal **482** (1997), no. 1 6.

[288] M. Zaldarriaga and U. Seljak, *All-sky analysis of polarization in the microwave background*, Physical Review D **55** (1997), no. 4 1830.

[289] D. J. Eisenstein, W. Hu, and M. Tegmark, *Cosmic complementarity: joint parameter estimation from cosmic microwave background experiments and redshift surveys*, The Astrophysical Journal **518** (1999), no. 1 2.

[290] L. Knox, R. Scoccimarro, and S. Dodelson, *The Impact of inhomogeneous reionization on cosmic microwave background anisotropy*, Phys. Rev. Lett. **81** (1998) 2004–2007, [astro-ph/9805012].

[291] D. Huterer, *Weak lensing and dark energy*, Phys. Rev. **D65** (2002) 063001, [astro-ph/0106399].

[292] S. Bird, M. Viel, and M. G. Haehnelt, *Massive Neutrinos and the Non-linear Matter Power Spectrum*, Mon. Not. Roy. Astron. Soc. **420** (2012) 2551–2561, [arXiv:1109.4416].

[293] D. Huterer and M. Takada, *Calibrating the nonlinear matter power spectrum: Requirements for future weak lensing surveys*, Astropart. Phys. **23** (2005) 369–376, [astro-ph/0412142].

[294] K. Kohri and T. Terada, *Primordial Black Hole Dark Matter and LIGO/Virgo Merger Rate from Inflation with Running Spectral Indices*, arXiv:1802.06785.

[295] H.-J. Huang, T. Eifler, R. Mandelbaum, and S. Dodelson, *Modelling baryonic physics in future weak lensing surveys*, Monthly Notices of the Royal Astronomical Society **488** (Jun, 2019) 1652–1678.

[296] L. Knox and M. Millea, *Hubble constant hunter's guide*, Physical Review D **101** (Feb, 2020).

[297] E. Di Valentino, O. Mena, S. Pan, L. Visinelli, W. Yang, A. Melchiorri, D. F. Mota, A. G. Riess, and J. Silk, *In the realm of the hubble tension—a review of solutions **, Classical and Quantum Gravity **38** (Jul, 2021) 153001.

[298] **DES** Collaboration, A. Kovács et al., *More out of less: an excess integrated Sachs-Wolfe signal from supervoids mapped out by the Dark Energy Survey*, Mon. Not. Roy. Astron. Soc. **484** (2019) 5267–5277, [arXiv:1811.07812].

[299] A. Kovács, R. Beck, I. Szapudi, I. Csabai, G. Rácz, and L. Dobos, *A common explanation of the hubble tension and anomalous cold spots in the cmb*, Monthly Notices of the Royal Astronomical Society **499** (Sep, 2020) 320–333.

[300] B. Stölzner, A. Cuoco, J. Lesgourgues, and M. Bilicki, *Updated tomographic analysis of the integrated Sachs-Wolfe effect and implications for dark energy*, Phys. Rev. D **97** (2018), no. 6 063506, [arXiv:1710.03238].

[301] Q. Hang, S. Alam, J. A. Peacock, and Y.-C. Cai, *Galaxy clustering in the DESI Legacy Survey and its imprint on the CMB*, Mon. Not. Roy. Astron. Soc. **501** (2021), no. 1 1481–1498, [arXiv:2010.00466].

[302] G. Rácz, L. Dobos, R. Beck, I. Szapudi, and I. Csabai, *Concordance cosmology without dark energy*, Mon. Not. Roy. Astron. Soc. **469** (2017), no. 1 L1–L5, [arXiv:1607.08797].

[303] R. Beck, I. Csabai, G. Rácz, and I. Szapudi, *The integrated sachs–wolfe effect in the avera cosmology*, Monthly Notices of the Royal Astronomical Society **479** (Jun, 2018) 3582–3591.

[304] A. Leauthaud, S. Saito, S. Hilbert, A. Barreira, S. More, M. White, S. Alam, P. Behroozi, K. Bundy, J. Coupon, and et al., *Lensing is low: cosmology, galaxy formation or new physics?*, Monthly Notices of the Royal Astronomical Society **467** (Feb, 2017) 3024–3047.

[305] J. Hartlap, P. Simon, and P. Schneider, *Why your model parameter confidences might be too optimistic: Unbiased estimation of the inverse covariance matrix*, Astron. Astrophys. (2006) [astro-ph/0608064]. [Astron. Astrophys.464,399(2007)].

[306] A. Taylor, B. Joachimi, and T. Kitching, *Putting the Precision in Precision Cosmology: How accurate should your data covariance matrix be?*, Mon. Not. Roy. Astron. Soc. **432** (2013) 1928, [arXiv:1212.4359].

[307] P. Coles and B. Jones, *A lognormal model for the cosmological mass distribution.*, Mon. Not. R. Astron. Soc. **248** (Jan., 1991) 1–13.

[308] A. Taruya, M. Takada, T. Hamana, I. Kayo, and T. Futamase, *Lognormal property of weak-lensing fields*, The Astrophysical Journal **571** (Jun, 2002) 638–653.

[309] S. Hilbert, J. Hartlap, and P. Schneider, *Cosmic shear covariance: the log-normal approximation*, Astronomy & Astrophysics **536** (Dec, 2011) A85.

[310] H. S. Xavier, F. B. Abdalla, and B. Joachimi, *Improving lognormal models for cosmological fields*, Monthly Notices of the Royal Astronomical Society **459** (Apr, 2016) 3693–3710.

[311] M. Takada and W. Hu, *Power Spectrum Super-Sample Covariance*, Phys. Rev. **D87** (2013), no. 12 123504, [arXiv:1302.6994].