Investigating the Formation of Star Clusters with Simulations and Observations

by

Gillen B. Brown

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Astronomy and Astrophysics)
in the University of Michigan
2022

Doctoral Committee:

Professor Oleg Y. Gnedin, Chair
Professor Eric Bell
Professor August Evrard
Associate Professor Mateusz Ruszkowski

Gillen Brown

gillenb@umich.edu

ORCID iD: 0000-0002-9114-5197

ACKNOWLEDGMENTS

I would like to start by thanking my advisor, Oleg Gnedin, for his guidance over the last six years. I came to the University of Michigan wanting to run simulations of galaxy formation, but having no experience and little idea what that truly entailed. He gave me the guidance, encouragement, and feedback to help me run the suite of simulations described in this dissertation. He also suggested that I take on a side project to measure star cluster radii. While the simulations have been a struggle at times (I wouldn't have written Chapter IV if they were easy), this project was always fun. It broadened my horizons, and I suspect it will be the most impactful part of this dissertation. I'm grateful that he suggested it. Finally, I greatly appreciate the guidance he gave on all the other things that are part of being an astronomer, such as writing papers, giving talks, and reading papers. He has a particular knack for taking my verbose clunky sentences and turning them into something half as long but twice as clear.

Next, I'd like to thank the other members of my thesis committee: Eric Bell, Gus Evrard, and Mateusz Ruszkowski. Every year I came to the committee meeting with a plan for the next several years, not all of which were realistic. In particular, they were gracious when in my 4th year, difficulties with the simulation led to me proposing the same timeline as I'd presented in my 3rd year, only pushed back one year. They gave me guidance on both the big picture of my dissertation and the small details. I'm grateful for their feedback throughout my time here.

I'd also like to thank my other collaborators, including Hui Li, Vadim Semenov,

Nick Gnedin, Arya Farahi, and Oliver Hahn. Hui got me up to speed on the simulations when I arrived at the University of Michigan, and without him I never would have gotten started. He also helped greatly later in the process when I encountered difficulty in my simulations. Vadim also helped me greatly when I was struggling to figure out what was wrong with my simulations. Without him, Chapter V would not be possible.

I started doing astronomy at the University of Missouri - Kansas City (UMKC), and would not have done so without the people there. In particular, Mark Brodwin and Dan McIntosh got me started doing research and taught all my astronomy classes. I also would like to thank my classmates, including Ben Sunnquist, Kameswara Bharadwaj Mantha, Ripon Saha, and Colton Luttrell.

The Stellar Halos group has been a place where I've presented ideas and gotten fantastic feedback. I'd particularly like to thank Ian Roederer, Eric Bell, Monica Valluri, and Adam Smercina. Beyond getting feedback on my work, hearing their thoughts about recent papers or their own results helped me grow as a scientist.

I have greatly enjoyed my time at Michigan, and the other grad students are a big part of that. From the first time I set foot on campus, I felt welcome here. I met Tyler Gardner and Juan Remolina (both fellow Missourians!) during the visit weekend, and we've remained friends throughout grad school, even after Juan moved to Washington. The rest of my cohort – Paco Holguin, Chaoran Wang, Thanawuth Thanathibodee, Benjamin Setterholm, and Chris Merchantz – all made the first two years of this program much more fun. We did all kinds of things together, including studying for the prelim.

I also must thank the other grad students in Oleg's research group: Xi Meng and Yingtian Chen. They have been excellent collaborators. Xi has helped me figure out all kinds of issues with my simulations with her physical intuition and great analysis skills. While Yingtian and I haven't overlapped long, he always lets me know when

I try to get too fancy with colors on plots and he can't understand. I've enjoyed working with both of them, and my work is better because of them.

I've greatly enjoyed spending time with the grad students outside of the office, too. There are a few groups in particular I'd like to thank. The board game group of Jesse Golden-Marx, Juliette Becker, and Juan Remolina made the first two years of my time here very fun. The intramural sports teams with Paco Holguin, Cam Pratt, Tyler Gardner, Matthew DeFurio, Jesse Golden-Marx, Rory Bowens, and several others were always great. Win or lose (mostly lose), we always had a great time. I've had so many good adventures with the camping crew, including Jenny Calahan, Kate Napier, Annie Blackwell, Irene Vargas-Salazar, and Lena Komarova. And I really should thank all the grad students, even if not specifically mentioned here. My time at Michigan has been better because of you. I'd also like to thank some other friends I've leaned on during my time here, including Shelby Bell, Andrea Cruz, and Anna Anders.

Finally, I would like to thank my family for their support. They encouraged me to read a lot from a young age, and that dramatically shaped my future. While I have gone my own way a bit, they have always cared about me and supported me throughout my journey. My success would not be possible without them.

Funding Acknowledgements

I acknowledge support from NASA through grant HST-AR-16614.001-A and the US National Science Foundation through grants 1412144, 1909063 and PHY-1748958.

TABLE OF CONTENTS

LIST OF TABLES

Table

LIST OF FIGURES

Figure

LIST OF APPENDICES

Appendix

ABSTRACT

I present my work to measure the radii of observed star clusters and investigate the formation and evolution of star clusters with cosmological simulations. By measuring the radii of over 6000 young star clusters in 31 nearby galaxies, I find that galaxies share a common cluster radius distribution and that a clear mass-radius relation exists. These results provide observational constraints for models of cluster formation and provide the initial conditions for detailed studies of cluster evolution. Next, I examine the properties of nuclear star clusters (NSCs) in an existing suite of cosmological simulations. I find that NSCs have spreads in age and iron abundance similar to those of the iron-complex globular clusters of the Milky Way. This lends support to the hypothesis that these globular clusters are the remnant NSCs of disrupted dwarf galaxies.

To investigate the formation of star clusters at high redshift and their evolution over cosmic time, I ran a series of cosmological simulations that directly modeled star clusters. I first create a new set of initial conditions (ICs) for these simulations. I developed a method that optimizes the grid structure of the IC to make simulations using it more computationally efficient and nearly a factor of two faster. I then use these ICs for the new simulations which directly model star cluster formation and evolution. After implementing several updates to the existing prescriptions, I investigate how variations in the subgrid models change cluster properties. I find that increasing the momentum boost from supernovae lowers the galactic star formation rate by decreasing the number of low-mass clusters. The star formation efficiency per freefall time dramatically changes cluster properties, with higher values leading to

more massive clusters, shorter timescales for cluster formation, and higher integrated star formation efficiencies. I find that most massive clusters form at high redshift $z > 4$, and extrapolation of cluster disruption to $z = 0$ produces good agreement with both the Galactic globular cluster mass function and age-metallicity relation.

CHAPTER I

Introduction

Star clusters are some of the most beautiful objects in the night sky. With a backyard telescope, one can see these dense collections of stars in a single field of view. The view only gets better when using the Hubble Space Telescope (Figure 1.1). Thousands, if not millions, of stars are all together, bound to each other and acting like a single object. Star clusters are more than just jewels in the night sky, though. They are important probes for understanding how stars form within our home galaxy, the Milky Way. Where did the stars within a cluster form? How old are these clusters? How big are they? How have these stars remained together for so long? Why are some stars in clusters but not others? All these questions have motivated scientific investigations over the last several hundred years. I am one in a long line of astronomers asking these questions about star clusters and developing increasingly sophisticated tools for answering them. In this dissertation, I'll describe my work to contribute to this field of knowledge.

To start, this introduction describes how galaxies form. Clusters are but one component of galaxies, so we must first understand the bigger picture. I'll then dive into some properties of star clusters, setting up the questions I'm asking in this dissertation. I close this introduction with some recent approaches to solving these questions and laying out where my work fits in.

Figure 1.1 Image of star cluster M15 from the Hubble Space Telescope (Credit NASA/ESA).

1.1 Galaxy Formation

1.1.1 Galaxy Formation in a Cosmological Context

The Universe started from an initial singularity in a dramatically quick expansion known as the Big Bang. While the Universe was nearly homogenous in the instant after the Big Bang, quantum fluctuations result in very small primordial density perturbations. Only 10^{-36} seconds after the beginning of the Universe, a process known as inflation began (Guth, 1981). The size of the Universe increased dramatically, stretching the initial quantum fluctuations to much larger scales. As the Universe continues to expand even after inflation ends, eventually matter becomes the dominant component of the Universe, with radiation and dark energy having lower energy densities. However, the photon density is still high enough that baryonic matter and radiation are coupled together. Radiation pressure acts on the gas in addition to the regular gas pressure, greatly increasing the gravitational force required to overcome this pressure and begin collapse. However, dark matter, which does not interact with photons, lacks any pressure support and is able to freely collapse, starting the process of structure formation. Eventually, as the Universe expands further and cools, electrons recombine with protons, forming the first atoms. As Thomson scattering from free electrons is the primary way in which photons interact with ionized baryonic matter, this decrease in the number of free electrons lets photons decouple from the baryonic gas. These photons are able to stream away freely and continue all the way to $z = 0$ to be seem by us now as the Cosmic Microwave Background (CMB). When decoupling happens, it removes the radiation pressure support on the gas. This makes it much easier for baryonic gas to begin gravitational collapse. With the dark matter structures growing for quite a while now, the baryons follow the dark matter into these newly forming cosmic structures that eventually grow into the galaxies we see today. Without the initial quantum density fluctuations, inflation expanding

these fluctuations to large scales, and the fact that dark matter does not interact with photons, the beautiful cosmic structures we see today would not exist.

Even after baryons begin collapsing, cosmic structures continue growing, driven by dark matter. Once a given perturbation collapses, the dark matter reaches a steady state, which we call a dark matter halo (Navarro et al., 1997). Analytic models can predict the mass distribution of these halos well (Press & Schechter, 1974), and simulations can impressively match the distribution of matter in the Universe (Springel et al., 2006). While dark matter lays the foundation for galaxy formation, we now turn to what happens to the baryons inside a given dark matter halo.

1.1.2 Galaxy Formation on Small Scales

The cosmological framework described in the previous section set the stage for galaxy formation within individual dark matter halos. Here, the physics we will encounter is much more complicated, with gas, radiation, stars, and dark matter all influencing one another through a wide range of processes. To start, the gravitational force of the dark matter pulls gas into the forming halo. This gas dissipates its thermal energy by radiating it away, decreasing the temperature and therefore the pressure support, allowing the gas to collapse to denser configurations in the center of the dark matter halo. While the gas can lose its thermal energy, it cannot lose its angular momentum. Tidal torques in the early Universe gave halos and their gas some spin, and this angular momentum must be conserved as the gas collapses, making a disk the natural configuration of gas within galaxies (Fall & Efstathiou, 1980; Mo et al., 1998).

Within this gas disk, the gas is not totally uniform. The densest parts of this gas can collapse further once they exceed the threshold laid out by Toomre (1964). These regions become dense enough that molecular hydrogen can form, making these objects known as giant molecular clouds (GMCs). In spiral galaxies, GMCs trace the spiral

structure, indicating that star formation is enhanced within the spiral arms. Within GMCs, supersonic turbulence and self-gravity lead to a complex internal structure, with clumps inside the GMC having even higher densities (Urquhart et al., 2018). These clumps are the seeds of star formation. Once the mass of a clump reaches the Jeans mass, runaway collapse begins and continues until stars have formed (Jeans, 1902).

The process of star formation is incredibly complicated, and as an entire field of astronomy there is simply too much to even summarize here (see Krumholz 2014 for a recent review). In the interest of brevity, I will summarize some of the processes most relevant to this dissertation. As the gas of a molecular cloud is converted into stars, we write the star formation rate as

$$\dot{M}_\star = \frac{\epsilon_{\rm ff}}{t_{\rm ff}} M_g \tag{1.1}$$

where $\dot{M}_\star$ is the star formation rate, M_g is the gas mass of the cloud, $t_{\rm ff}$ is the free-fall time of the cloud, and $\epsilon_{\rm ff}$ is the star formation efficiency per free-fall time. $\epsilon_{\rm ff}$ is a very important parameter that describes the efficiency of star formation by quantifying what fraction of gas is turned into stars per free-fall time (Krumholz et al., 2019). As this framework implies, GMCs tend to form large collections of stars in one episode, which typically lasts a few Myr (Da Rio et al., 2010; Beccari et al., 2017). The resulting group of stars is known as a star cluster, and most stars form in clusters (Lada & Lada, 2003).

Once these stars begin to form, feedback from these stars works to disperse the GMC and prevent future star formation. As protostars form, they produce outflows that eject significant amounts of gas and can star to break up the dense regions of molecular clouds (Bally, 2016). Once massive stars form, their UV photons ionize the gas and contribute radiation pressure, continuing the dispersal of the natal cloud

(Williams & McKee, 1997). Finally, massive stars explode as supernovae (SN). Because SN happen in evolved stars, feedback from SN is delayed compared to other methods. They contribute less to the dissolution of individual molecular clouds, especially at low masses (Grudić et al., 2022). However, the large amounts of energy injected by SN lead to dramatic consequences within the galaxy as a whole (Larson, 1974). SN greatly reshape the gas nearby, can launch galactic outflows, and can trigger star formation by contributing to the collapse of nearby gas. On galaxy scales, SN feedback has proved to be one of the essential ingredients that determine the cosmic star formation rate of galaxies (Cole et al., 1994; Hopkins et al., 2014).

Star clusters play an essential role in the story of galaxy formation. As most stars form within clusters, they can be thought of as the building blocks of galaxies. This makes them useful for understanding two major subfields of astronomy, star formation and galaxy evolution. The properties of young star clusters depend strongly on the physics of star formation, then their evolution over time depends on the galactic structure they're embedded in. Star clusters that have survived billions of years can reveal properties of ancient galaxies and help constrain their assembly history. In this dissertation, the overarching goal is to understand star cluster evolution over cosmic time. This requires understanding the initial properties of clusters as well as the processes that govern their evolution. With that in mind, let us now take up some of the properties of star clusters to make the goals of this dissertation more specific.

1.2 Star Clusters

In this section, I will describe some of the different classes of star clusters we observe in the Universe and the processes that shape their evolution over cosmic time.

1.2.1 Young Star Clusters

As the mass of a cluster is its most basic property, one of the key descriptors of the young (<100 Myr) cluster population is the initial cluster mass function. The mass function of young massive clusters (YMCs) in external galaxies is shown in Figure 1.2 from Portegies Zwart et al. (2010). These mass functions are well described by a Schechter (1976) function, which is a power law with an exponential cutoff at high masses. The power law slope α is consistently measured to be around -2 (Bastian, 2008; Krumholz et al., 2019). However, the value of the cutoff mass depends on the star formation rate of the host galaxy, with more intense star formation giving higher cutoff masses. This also has the effect of increasing the maximum cluster mass, letting galaxies with more intense star formation form more massive clusters (Larsen, 2002). This can be seen in Figure 1.2, where the Antennae (a pair of interacting galaxies undergoing vigorous star formation) forms clusters with masses above $10^6\,M_\odot$, unlike all other galaxies shown in this figure.

The radius distribution of young clusters is another important observable. The radius is initially set at the end of the cluster formation process (Murray, 2009; Choksi & Kruijssen, 2021), so measurements of radii can constrain models of cluster formation. In addition, as we will discuss further in Section 1.2.2, the dynamical evolution of star clusters is very sensitive to their density. Measuring the radius (and therefore the density) helps constrain the timescale for cluster disruption.

However, radii measurements are often more difficult to obtain for clusters in external galaxies, as clusters must be at least partially resolved. The Hubble Space Telescope (HST) has been essential for this work, enabling measurements of cluster radii in M31 (Johnson et al., 2012), M51 (Chandar et al., 2016), M82 (Cuevas-Otahola et al., 2020), and M83 (Ryon et al., 2015), among others (e.g. Ryon et al., 2017). While these measurements have led to many interesting results, their heterogeneity can cause difficulties when comparing results. These studies used different instruments,

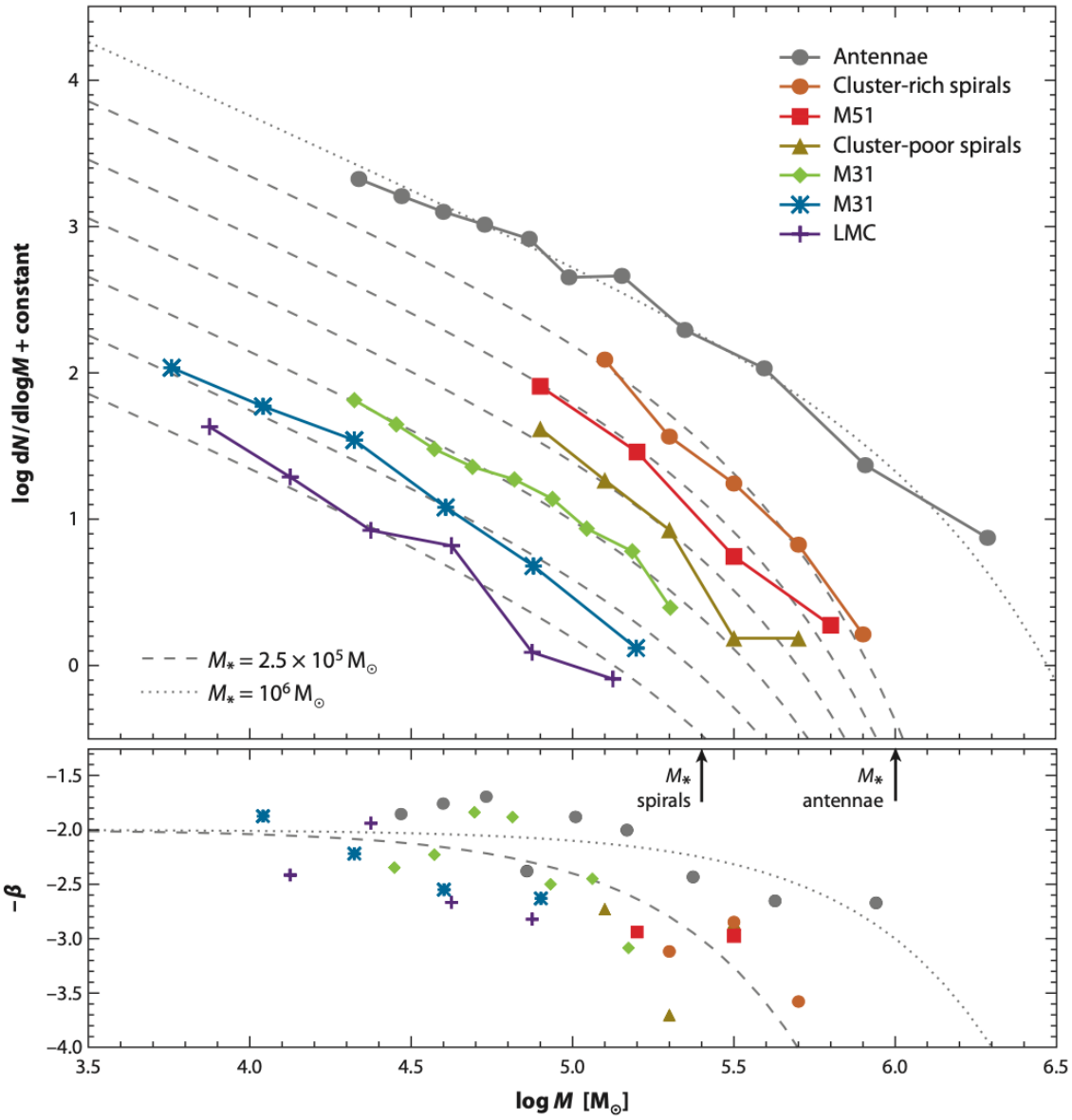

Figure 1.2 Figure from Portegies Zwart et al. (2010) showing the cluster mass functions of clusters younger than 1 Gyr in different galaxies. Dashed and dotted lines indicate Schechter functions, both with $\alpha = -2$ but different cutoff masses. The bottom panel shows the logarithmic slopes of the Schechter functions shown in the top panel.

different PSF construction, different fitting algorithms and different definitions. With the advent of large star cluster surveys such as the Legacy Extragalactic UV Survey (LEGUS, Calzetti et al., 2015; Adamo et al., 2017; Cook et al., 2019), the opportunity exists to create large uniform catalogs of cluster radius and density. In Chapter II, I will discuss my work to create the largest currently available uniform catalog of young star cluster radii.

1.2.2 Globular Clusters

In addition to the young clusters that we observe in the Milky Way and other galaxies, we also find old, massive clusters known as globular clusters (GCs) (Brodie & Strader, 2006). These clusters have masses in the range $\sim 10^4 - 10^6\,M_\odot$ and ages typically older than 10 Gyr (Harris, 1996; Jordán et al., 2007). These old ages make GCs a particularly powerful tool for understanding galaxy formation. When GCs form, they inherit the chemical properties of their host galaxies, giving us a snapshot of the state of the galaxy at a given time. The large masses also requires conditions that don't exist in the Milky Way currently, giving us insight into the properties of galaxies at high redshift (Kruijssen, 2014). GCs also form in satellite galaxies, and when these satellites merge with a central galaxy (such as the Milky Way), their clusters are deposited in this central galaxy. We can then use the distinct properties of these clusters to reconstruct the accretion history of the Milky Way (Kruijssen et al., 2019a).

GCs have sizes of a few parsecs, which means they are consistent with YMCs found in the local Universe (Ryon et al., 2017). This naturally leads to the hypothesis that GCs are the surving subset of a larger population of YMCs that formed at high redshift. However, the mass function of GCs is well characterized by a lognormal distribution with a peak mass of around $2\times10^5\,M_\odot$ (Harris, 1991; Jordán et al., 2007), in contrast to the Schechter (1976) function commonly used to describe YMCs. This

transformation of the mass function over cosmic time in this way would require a preferential destruction of low-mass clusters. The initial population of star clusters forming at high redshift would not fully reach us today at $z = 0$.

This hypothesis has found support from many studies (e.g. Fall & Zhang, 2001; Vesperini et al., 2003; Prieto & Gnedin, 2008; Elmegreen, 2010; Kruijssen, 2015). Clusters face a range of processes that disrupt them, allowing for the transformation in cluster populations over time (Spitzer, 1958, 1987; Gnedin & Ostriker, 1997). The initial gas expulsion decreases the gravitational potential of the cluster, unbinding many low-mass clusters (Hills, 1980). From there, clusters evolve within the gravitational potential of their host galaxy. In particular, the strength of the tidal field determines how much energy individual stars need to become unbound from the cluster and escape. Stronger tidal fields make it easier for stars to leave clusters. This is often parametrized in terms of the tidal radius (or Jacobi radius), which describes the radius within which stars remain bound to the cluster. For an axisymmetric disk, the Jacobi radius is

$$r_{\rm J} = \left(\frac{GM}{\lambda_x}\right)^{1/3} \tag{1.2}$$

where M is the cluster mass and λ_x is the eigenvalue of the tidal tensor along the direction towards the galactic center.

One way in which stars can move outside of this radius and leave the cluster happens because of stellar evolution. As the stars in clusters evolve, supernovae explode and AGB stars shed their outer envelopes, decreasing the stellar mass of the cluster. This makes the cluster expand due to the decreased gravitational potential, but this also shrinks the tidal radius. Some stars that were previously inside this radius are now outside it and are now unbound, decreasing the mass of the cluster further. Relaxation also decreases the mass of the cluster. The distribution of velocities in the cluster tends toward a Maxwellian as stars randomly exchange energy. But stars with too much energy can then escape the cluster, making this high-energy tail continually

underpopulated. This cycle of the velocity distribution refilling the high-energy tail only for these stars to then escape produces a gradual mass loss over time.

In addition to the processes that happen in a constant tidal field, variations in the tidal field also serve to disrupt clusters. Clusters pass by dense GMCs, which provide tidal shocks that can completely unbind the least dense clusters or strip stars from denser clusters (Spitzer, 1958). Many clusters also have orbits that take them above and below the plane of the disk. Passing back through the disk imparts a similar tidal shock (Ostriker et al., 1972). Finally, the orbit of clusters often changes, gradually changing the tidal field and therefore the timescales required for clusters to disrupt. Dynamical fraction descreases the angular momentum of clusters, causing them to spiral towards the galactic center (Chandrasekhar, 1943). Radial migration can move clusters inward or outward (Sellwood & Binney, 2002).

The density of the cluster plays a central role in determining whether clusters survive (e.g. Gieles & Renaud, 2016). Denser clusters are more resistant to tidal disruption. The mass-density relation (or equivalently, the mass-radius relation) of clusters is therefore an essential initial condition that governs how clusters evolve over time. If low-mass clusters are less dense, then they will disrupt more quickly. This is another reason why constraining cluster radii is critical. Radius measurements give us constraints on cluster formation theories, but then also give us the initial conditions for understanding cluster evolution over cosmic time.

To properly model the evolution of star cluster populations over cosmic time, cosmological simulations are needed. These simulations are the only ones that can accurately model the hierarchical growth of galaxies and their star cluster populations, form star clusters with realistic properties in the correct locations within galaxies, and then track cluster evolution over time in a realistic tidal field. All of the processes described in the last few paragraphs must be modeled self-consistently to accurately track cluster evolution over time and recover the properties of observed GCs at $z = 0$.

This is the central goal of my dissertation. In Chapter II I discuss my measurements of star cluster radii, which constrain the initial conditions of cluster evolution. Then in Chapters IV and V, I discuss my work on a suite of cosmological simulations that model star cluster formation and evolution. These simulations have an algorithm for star cluster formation, allowing me to investigate the properties of young star clusters. They also directly model cluster disruption, enabling an investigation of the evolution of star cluster populations over cosmic time.

1.2.3 Nuclear Star Clusters and their Relation to Globular Clusters

Nuclear star clusters (NSCs) are a special class of star cluster located in the innermost regions of most galaxies (Leigh et al., 2012). These clusters are much more massive and luminous than other star clusters, often being the brightest star cluster in their host galaxy (Côté et al., 2006; Jordán et al., 2009; Neumayer et al., 2020), even though their radius can be similar to typical star clusters (Georgiev et al., 2016). In the Milky Way's NSC, studies have found a complicated star formation history (Do et al., 2015; Feldmeier-Krause et al., 2017a). Stars of a wide variety of ages and metallicities are present.

There are two main theories for how NSCs form. First, dynamical friction can cause GCs to drift to the center of the galaxy, where over time multiple clusters merge into a large NSC (Tremaine et al., 1975; Lotz et al., 2001). On the other hand, the presence of stars with ages less than 10 Myr in the Milky Way's NSC suggests a recent episode of in situ star formation (Krabbe et al., 1991; Lu et al., 2013). These two mechanisms likely both contribute, with GC accretion being dominant in low-mass galaxies, while in situ star formation dominates at higher masses (Neumayer et al., 2020).

One particular reason NSCs are interesting is their potential link to the currently-existing GCs of the Milky Way. When a satellite galaxy hosting an NSC merges with

a larger central galaxy, the outskirts of the galaxy will be stripped away, potentially leaving the NSC isolated. It would then appear as a massive GC, but with stars of a wide range of ages and metallicities (Freeman, 1993; Böker, 2008). We see several clusters like this in the Milky Way. As typical GCs have a single value of metallicity, these clusters with measurable iron spreads have become known as "anomalous" or "iron-complex" clusters (Marino et al., 2015). Some examples include ω Cen (Johnson et al., 2009), M54 (Carretta et al., 2010a), and Terzan 5 (Massari et al., 2014).

While the connection between the anomalous GCs and NSCs has long been suspected, it needs to be investigated in a fully cosmological context. Simulations of star cluster formation within this cosmological context would allow a comparison of the properties of simulated NSCs with the observed anomalous GCs in the Milky Way.

1.3 Star Cluster Modeling in Simulations

With this wealth of observational data, we are well-positioned to test theories of cluster formation and evolution. While analytic models can explain some simple aspects of star cluster formation, they have their limitations. On cosmological scales, the analytic framework for structure formation based on linear perturbation theory breaks down once the fluctuations grow large, making future evolution very complex to model analytically. On galactic scales, simple models for gravitational collapse of molecular clouds assume that clouds are spherical and either isolated or embedded in a simple galactic disk (Jeans, 1902; Toomre, 1964; Escala & Larson, 2008). Real molecular clouds are complex, hierarchically structured objects that evolve in similarly complex ways. Molecular clouds are not the only complex structures within galaxies. The galaxy-scale distributions of gas and stars do not always follow simple analytical models. Their evolution is governed by a huge range of physical processes, such as heating, cooling, stellar feedback, dynamical interactions from galactic mergers, and magnetic fields. This complexity makes detailed analytic models impossible.

Numerical simulations are therefore essential for turning an idea into detailed predictions that can be compared against observations.

In simulations of galaxy formation, the simulation code must numerically model the evolution of gas, stars, and dark matter. Stars and dark matter are both typically treated as collisionless particles that move under the influence of gravity only. Gas follows the Eulerian equations for hydrodynamics, which describe the conservation of mass, momentum, and energy. An equation of state is also required. All components contribute to the gravitational potential through the Poisson equation. Numerically, these equations are solved by discretizing the gas into discrete resolution elements. This can be implemented in several ways. A full comparison of methods is beyond the scope of this introduction (see Vogelsberger et al. 2020 for a recent review), but I will summarize two key methods. One popular approach is smoothed particle hydrodynamics (SPH), which is a Lagrangian method that treats gas as particles that follow the local flow (Gingold & Monaghan, 1977). Fields such as the density can be reconstructed through kernel density estimation. Alternatively, grid based approaches divide the simulation volume into cells, which each have defined quantities such as density and temperature. These cells exchange mass and energy through cell interfaces. In the simulations used in this dissertation, I use the grid-based Adaptive Refinement Tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999).

While solving hydrodynamics and gravity is essential to modeling galaxies, simply following the flow of gas under the influence of gravity is not enough. Stars form out of dense regions of GMCs, but no simulation has the dynamic range to follow the gas from the kpc galactic scales down to the AU scales of star formation. Therefore, simulations need some prescriptions for what happens on scales below the resolution limit. These prescriptions are called "subgrid models," and are essential components of astrophysical simulations. For example, most simulations of galaxy formation instantaneously convert gas to stars once it meets some set of criteria, usually ones

requiring the gas to be dense and cold. Similarly, the feedback from massive stars is another process that requires subgrid model. By definition, subgrid models are needed when the simulation cannot resolve a given process, meaning that they are often ad-hoc and have free parameters. As an example, supernovae have been modeled in many different ways over the years (Katz, 1992; Springel & Hernquist, 2003; Stinson et al., 2006; Agertz et al., 2011, 2013; Hopkins et al., 2014; Keller et al., 2014; Hopkins et al., 2018). To reproduce reasonable galaxy properties, the free parameters of subgrid models need to be calibrated against observations. This calibration is an essential part of the process of running modern simulations of galaxy formation.

While the basic framework of simulations is conceptually simple, they are computationally expensive. Millions to billions of computational elements (stars, dark matter particles, gas cells) are required to model galaxies. This expense means that tradeoffs must be made. While high resolution allows for more accurate modeling of galaxies, particularly on small scales, the increase in the number of resolution elements means that these simulations must contain a small volume (Emerick et al., 2019). On the other hand, simulations that aim to simulate a cosmological volume must have lower resolution to be computationally feasible (Vogelsberger et al., 2014; Schaye et al., 2015; Pillepich et al., 2018). No simulation can do everything.

When running a simulation, one must start from an initial condition (IC). The IC specifies the volume contained in the simulation and has an initial configuration of gas properties within this volume. This takes many forms depending on the simulation. Simulations of GMC often use ICs containing a uniform sphere with a random velocity field to emulate turbulence (Lane et al., 2022), while ICs for simulations of isolated galaxies typically include gas, stars, and dark matter following simple analytical profiles (Kim et al., 2016). Of particular interest for this dissertation are ICs for cosmological simulations. These ICs typically are created at very high redshift (e.g. $z = 100$), and contain density fluctuations based on the matter power spec-

trum at that redshift (see Chapter IV for a more complete summary, and Hahn & Abel 2011 for a detailed description). These ICs can easily be made to cover cosmological volumes, with boxes over 100 Mpc on a side being common for the largest hydrodynamic simulations. However, this large volume means that while galaxies are modeled in their full cosmological context, the resolution must be low. On the other hand, simulations of single isolated galaxies can have much higher resolution, but lack the cosmological environment that is essential for properly modeling the hierarchical growth of galaxies.

To bridge these scales, a hybrid approach has been developed, known as "zoom-in" ICs (Katz et al., 1994; Navarro & White, 1994). These simulations model cosmological volumes with very low resolution, while dramatically increasing the resolution in the volume containing a region of interest, such as a galaxy or cluster. For example, one zoom-in IC used in this dissertation has a box size of 35 comoving Mpc, cell sizes in the uninteresting region of 137.5 comoving kpc, then cell sizes as small as 3 physical pc inside the region of interest. This huge dynamic range is challenging for grid based codes like ART to handle. Later in Chapter IV of this dissertation, I describe my work to make zoom-in initial conditions more computationally efficient for the ART code.

With this basic summary of the key aspects of simulations in hand, let us turn to simulations of star cluster formation. There are several approaches currently used to simulate the problem. These range from very high resolution simulations of individual molecular clouds without any galactic context, to simulations of single galaxies that directly model star clusters, to simulations of cosmological volumes with purely subgrid models of cluster formation. In the rest of this section I'll discuss several examples of simulations modeling star cluster formation, discuss the strengths and weaknesses of each, then close with a discussion of where my work fits in.

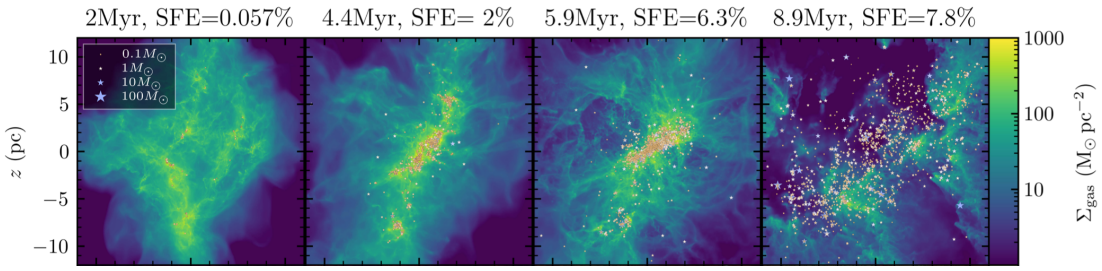

Figure 1.3 Figure adapted from Grudić et al. (2022) showing the evolution of the gas surface density and locations of stars in a simulated star cluster as it forms.

1.3.1 Simulations of Individual Molecular Clouds

To understand how star clusters first form out of molecular clouds, several groups simulate individual molecular clouds with very high resolution, including resolving individual stars (e.g. Klessen & Burkert, 2000; Offner et al., 2009; Fujii et al., 2021; Mathew & Federrath, 2021). Simulating a whole galaxy with this resolution would be too computationally expensive, but a single molecular cloud is feasible. A recent example of this is the STARFORGE project (Grudić et al., 2021). In Grudić et al. (2022), they simulated one molecular cloud of mass $2 \times 10^4\,\mathrm{M}_\odot$ with a mass resolution of $10^{-3}\,\mathrm{M}_\odot$. They were able to reproduce some key star formation observables, such as the IMF and the stellar multiplicity fraction. They also found that the cluster assembles hierarchically, as shown in Figure 1.3. Stars initially form in separate subclusters that collapse along with the gas into a central compact cluster. Eventually, feedback from the newly formed stars disperses the gas, reducing the gravitational potential and allowing the stars to expand again. Their simulated star cluster is unbound, and would continue to disperse if the simulation were continued.

While simulations like these are excellent tools for understanding the process of star cluster formation on small scales, their high cost means that they cannot currently be extended to longer timescales, more massive molecular clouds, or broader portions of the galaxy. These simulations cannot determine the long-term evolution of star clusters, as they do not simulate beyond the formation of the star cluster and do not

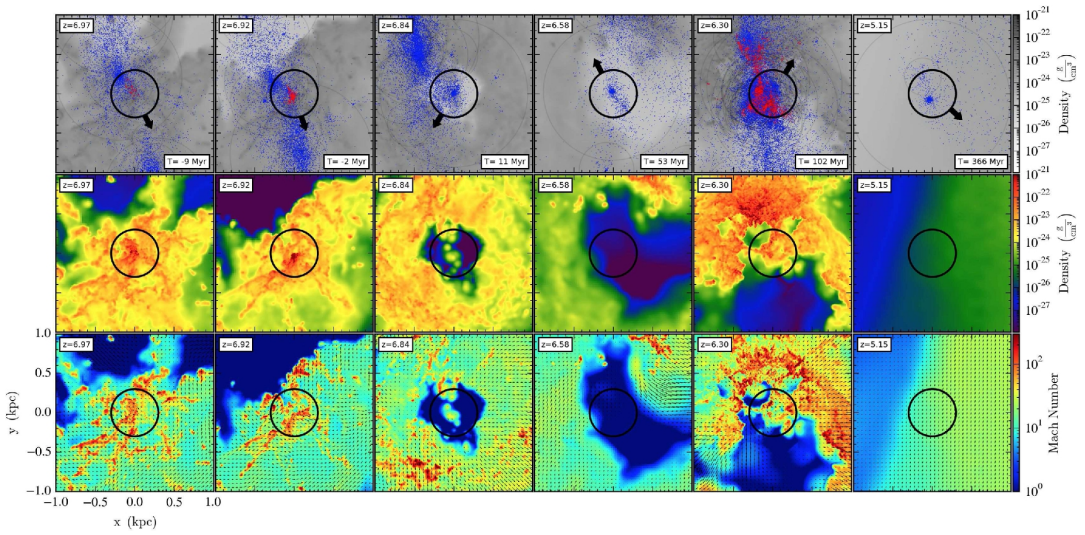

Figure 1.4 Figure from Kim et al. (2018) showing the evolution of a simulated star cluster forming. The grey shading in the top panel shows the projected gas-density, with red points marking star particles younger than 5 Myr and blue points marking all other star particles. The middle panel shows the gas surface density, and the bottom panel shows the gas Mach number. Columns show the evolution of the region with time. In the second column, a very dense cloud forms at the intersection of converging flows, which quickly forms a star cluster that disperses the gas, then remains bound for the entirety of the simulation.

include the surrounding galactic environment.

1.3.2 Directly Resolving Star Clusters in Galaxy Formation Simulations

While simulations of molecular clouds have resolution too high to extend to full galaxies, efforts have been made to run full galaxy simulations with resolution just high enough to resolve star clusters. Of particular interest are simulations from the Feedback In Realistic Environments (FIRE) project as reported on in Kim et al. (2018). These authors use a zoom-in IC around a halo with a virial mass $\sim 10^{10}\,\mathrm{M}_\odot$ at $z = 6$, using a mass resolution of $800h^{-1}\,\mathrm{M}_\odot$ and force resolution of $1.4h^{-1}$ pc. With this high resolution, massive star clusters are resolved by $10^2 - 10^3$ star particles. In particular, they followed the evolution of a massive ($\approx 10^6\,\mathrm{M}_\odot$) gravitationally bound star cluster as it forms at $z \approx 7$. In Figure 1.4 I reproduce a visualization of this cluster.

Simulations like this are excellent tools for understanding the formation of globular clusters at high redshift. They self-consistently model the formation of stars at extremely high resolution, even resolving the half-mass radius of the cluster. Feedback from the new cluster stops future star formation, as it does in simulations of molecular clouds. While simulations such as these are perhaps the future of cluster formation simulations, they are limited by their high resolution. The simulations of Kim et al. (2018) only ran for 420 Myr. Simulations of Milky Way-mass galaxies to $z = 0$ with this resolution are prohibitively expensive. Other types of simulations are needed to follow the evolution of clusters over cosmic time.

1.3.3 Semi-analytic Models of Star Cluster Evolution

In contrast to simulations that directly resolve cluster formation, semi-analytic models are well-suited to following the evolution of populations of star clusters over cosmic time. One example of this class of simulations is the E-MOSAICS model (Kruijssen & Lamers, 2008; Kruijssen, 2009; Kruijssen et al., 2011, 2012; Pfeffer et al., 2018; Keller et al., 2020; Trujillo-Gomez et al., 2021). This model uses the EAGLE hydrodynamic cosmological simulation code as a base (Schaye et al., 2015), then applies a subgrid model of cluster formation and evolution on top. When stars form, the model treats each star particle as a collection of star clusters with an assumed mass function. The tidal evolution of these clusters can be tracked in the simulation over time, directly using the star particle's environment to calculate disruption. These star particles can be tracked to $z = 0$, enabling the modeling of GC kinematics. However, these models are limited by the low spatial resolution of EAGLE. It does not resolve the cold dense phases of the interstellar medium, so it underestimates tidal disruption of clusters. This results in the overproduction of low-mass clusters at $z = 0$.

A simpler model is that developed in a series of papers starting with Muratov & Gnedin (2010) and continuing to Li & Gnedin (2014); Choksi et al. (2018); Choksi

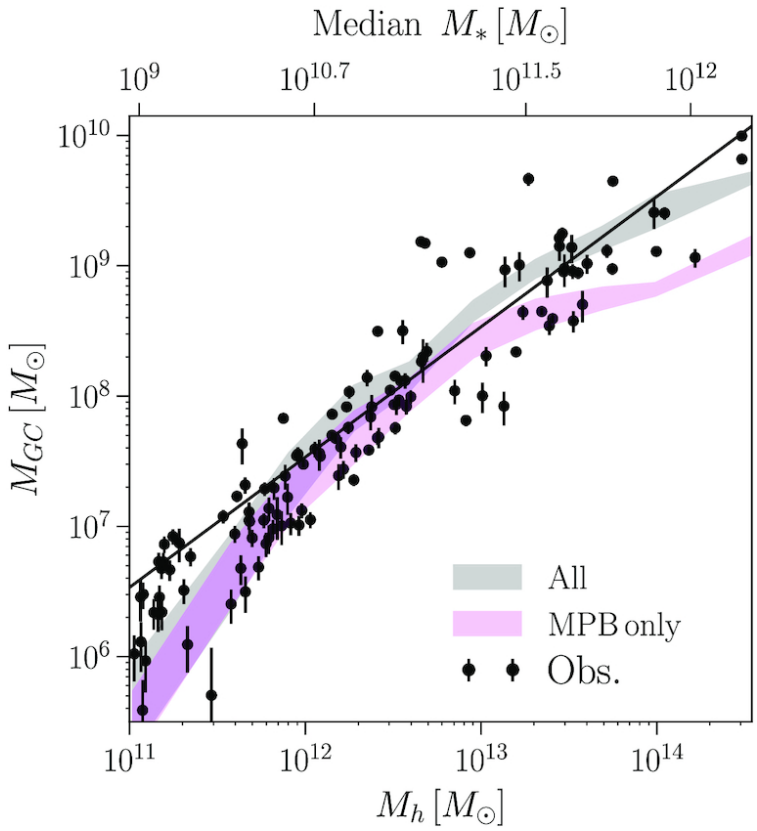

Figure 1.5 Figure from Choksi & Gnedin (2019b) showing relation between galaxy halo mass and mass of the globular cluster system. The gray shaded region shows the fidicual model result, which matches observations well.

& Gnedin (2019a,b). These models are based on dark matter merger trees from simulations such as Illustris (Vogelsberger et al., 2014). Clusters are seeded when the halo growth rate is large, so as to simulate cluster formation in galactic mergers or other rapid star formation events. Scaling relations are used to determine cluster formation rates, with masses of individual clusters drawn from an assumed mass function. Clusters then tidally disrupt at a constant rate. Remarkably, this model only has two free parameters, but can reproduce several attributes of observed star cluster populations, including the relation between mass in GCs and halo mass of the host galaxy. In Figure 1.5 I reproduce a figure from Choksi & Gnedin (2019b) showing this relation. However, this model currently contains no information about the locations of clusters within galaxies.

While semi-analytic models such as these are powerful tools for understanding the evolution of star cluster populations over time, they are less able to inform our understanding of star cluster formation. Both models presented here assume a cluster initial mass function. While E-MOSAICS does model the location of clusters within galaxies, it does not self-consistently account for the feedback from star clusters when setting cluster properties. Clusters are seeded based on analytic prescriptions using local gas conditions. A different class of simulations are needed to understand cluster formation and evolution together.

1.3.4 Cosmological Simulations with Continuous Cluster Formation

The final category of simulations I'll discuss are the unique simulations presented in Li et al. (2017, 2018); Li & Gnedin (2019) by former University of Michigan Astronomy Ph.D. student Hui Li. All other cosmological simulations treat star formation as an instantaneous process, with the prescription not changing significantly since the 1990s (Katz, 1992; Cen & Ostriker, 1992). Even though star formation happens over several Myr, and these simulations have timesteps small enough to resolve this process temporally, they do not. In contrast, this new formalism does resolve star formation, letting a star particle form within a GMC over time until its feedback disperses the gas. This treats star clusters as the unit of star formation.

Star particles are seeded in dense molecular gas, and accrete material over time from a region repreenting the GMC. Figure 1.6 shows a cartoon from Li et al. (2018) illustrating how the star cluster particle can accrete gas from this region of fixed physical size (with a radius of 5 pc) representing the GMC, which can include multiple cells within the simulation volume. The growth rate of a cluster is given by

$$\dot{M} = \frac{\epsilon_{\rm ff}}{t_{\rm ff}} \sum_{\rm cell} f_{\rm GMC} V_{\rm cell} f_{\rm H_2} \rho_{\rm gas} \tag{1.3}$$

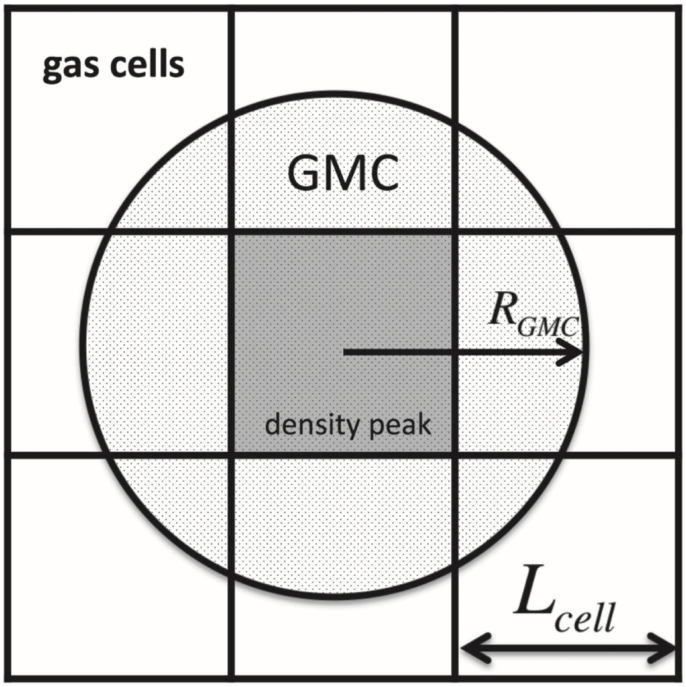

Figure 1.6 Figure from Li et al. (2018) showing a sketch of the star forming GMC sphere within the cell structure of the simulation.

where $\epsilon_{\rm ff}$ is the local star formation efficiency per freefall time $t_{\rm ff}$, $f_{\rm GMC}$ is the fraction of cell volume $V_{\rm cell}$ included in the GMC, $f_{\rm H_2}$ is the local mass fraction of molecular gas, and $\rho_{\rm gas}$ is the local total gas density. $\epsilon_{\rm ff}$ is a key parameter that determines that rate of cluster growth, influencing many other cluster properties in turn. Gas accretion continues until feedback from the newly formed stars terminates future accretion, ending the cluster's growth. This self-consistently sets the mass of the star particles. However, not all stars are bound to the newly formed cluster. To account for this, a prescription determines the initial bound fraction based on the integrated star formation efficiency of the formed cluster. Clusters that turn more of their gas into stars have a higher fraction of their stars gravitationally bound in a cluster (Li et al., 2019). These prescriptions self-consistently model star cluster formation, with particles that can be interpreted as individual star clusters.

Importantly, this direct model for cluster formation does not require the assumptions that semi-analytic models like E-MOSAICS do. The cluster mass function arises

naturally, rather than being imposed. The later evolution of clusters is also tracked well, as clusters evolve in the tidal field of the galaxy, which is simulated at high resolution. Unlike very high resolution simulations of star cluster formation such as Kim et al. (2018), these simulations are able to span many Gyr of evolution, potentially reaching $z = 0$. For these reasons, this model occupies a niche not otherwise occupied by simulations of cluster formation.

Simulations using this prescription for star cluster formation used an initial condition containing a single MW-mass galaxy and were run to $z = 1.5$ (Li et al., 2017, 2018; Li & Gnedin, 2019). These simulations were able to reproduce many key observed properties of star clusters, including the shape of the initial cluster mass function, the total mass of stars contained in bound clusters, the relationship between the maximum cluster mass and the star formation rate density, and the formation timescales of star clusters.

This simulation suite presented a novel approach for modeling star cluster formation in cosmological simulations and showed that it produces realistic star cluster properties, making these simulations a success. However, they still had some limitations. The biggest limitation is that the simulations only reached $z = 1.5$. Direct comparisons with the observed cluster population at $z = 0$ is not possible. Importantly, cluster disruption cannot be fully followed to $z = 0$ to see which clusters survive as GCs. While analytical calculations are possible, they are naturally limited (Li & Gnedin, 2019). To fully understand how GCs form, we need to self-consistently track their evolution over cosmic time. Also, these simulations only contained one Milky Way-sized galaxy, limiting the statistical power of the comparisons that can be made.

In this dissertation, I will describe my work to run the next-generation simulations of star cluster formation in a cosmological context using this model. I have made several improvements to the cluster formation and feedback prescription, then run a

new suite of simulations. These runs use two Local Group analogs, giving four Milky Way mass galaxies rather than a single one. More importantly, these runs are set up with the goal of reaching $z = 0$ to enable the comparisons with the Milky Way laid out in the previous paragraph. While they have not reached $z = 0$ yet, my work sets the stage for this important comparison.

1.4 Dissertation Overview

In the rest of this dissertation, I will describe my work investigating star cluster formation in simulations and observations. On the broadest scales, my goal is to understand how clusters evolve over cosmic time. While I obviously cannot achieve that fully in one dissertation, I have made advances in several key areas that both increase our current knowledge and set the stage for future work.

Chapter II describes my work to measure the radii of young star clusters. These measurements can be used to place quantitative constraints on theories of cluster formation, while also setting the initial conditions for cluster evolution. In Chapters III, IV, and V, I describe my work modeling star cluster formation in cosmological simulations. Chapter III presents my work to test the hypothesis that NSCs of satellite galaxies could become the progenitors of the anomalous GCs seen in the Milky Way. In Chapter IV, I present my efforts to modify ICs for cosmological simulations to make simulations using them more computationally efficient. This sets the stage for the capstone of this dissertation, Chapter V. Here I describe my new suite of cosmological simulations that model cluster formation. I perform a detailed analysis of different prescriptions for cluster formation and feedback, investigating how subgrid models should be tuned to optimally match observations. I also show initial results for the evolution of star clusters, showing that cluster disruption over time can transform the mass function of young clusters to the mass function of GCs in the MW.

CHAPTER II

Radii of Young Star Clusters in Nearby Galaxies

This chapter was published as: ***Brown, G.****, & Gnedin, O. Y. 2021, MNRAS, 508, 5935*

Abstract

We measure the projected half-light radii of young star clusters in 31 galaxies from the Legacy Extragalactic UV Survey (LEGUS). We implement a custom pipeline specifically designed to be robust against contamination, which allows us to measure radii for 6097 clusters. This is the largest sample of young star cluster radii currently available. We find that most (but not all) galaxies share a common cluster radius distribution, with the peak at around 3 pc. We find a clear mass-radius relation of the form $R_{\rm eff} \propto M^{0.24}$. This relation is present at all cluster ages younger than 1 Gyr, but with a shallower slope for clusters younger than 10 Myr. We present simple toy models to interpret these age trends, finding that high-mass clusters are more likely to be not tidally limited and expand. We also find that most clusters in LEGUS are gravitationally bound, especially at older ages or higher masses. Lastly, we present the cluster density and surface density distributions, finding a large scatter that appears to decrease with cluster age. The youngest clusters have a typical surface density of 100 $\rm M_\odot\,pc^{-2}$.

2.1 Introduction

Young star clusters bridge the small scales of star formation and the large scales of galaxy formation. They are easily detected in nearby star-forming galaxies and contain the majority of massive stars (Krumholz et al., 2019). Their almost universal luminosity function, and corresponding mass function (Adamo et al., 2020), is used both as a test of molecular cloud collapse models (Ballesteros-Paredes et al., 2020) and as a constraint on the star formation modeling in cosmological simulations (Li et al., 2018). However, our understanding of the origin and long-term evolution of star clusters is still hindered by lack of reliable measurements of their density distribution.

Depending on the initial density, young massive clusters may evolve into globular clusters or may dissolve into the smooth stellar field (Portegies Zwart et al., 2010). As clusters form in giant molecular clouds (GMCs), feedback from massive stars ejects the residual gas, making the cluster expand to re-establish virial equilibrium (Goodwin & Bastian, 2006). Additionally, tidal shocks from encounters with other GMCs or spiral arms kinematically heat the cluster and lead to its disruption (Spitzer, 1958). The density of young clusters determines whether they can survive this intense phase as bound stellar systems.

To calculate the density of young star clusters, one needs to measure their radii. Currently published measurements are limited to only a few galaxy samples. Kharchenko et al. (2013) measured the radii for $\sim$ 1100 open clusters in the Milky Way, while samples from external galaxies include Bastian et al. (2012) and Ryon et al. (2015) who measured radii for several hundred clusters in M83, about 3800 clusters in M51 measured by Chandar et al. (2016), and 514 clusters in M31 measured by the PHAT survey (Johnson et al., 2012; Fouesneau et al., 2014). Of particular note for this paper, Ryon et al. (2017) (hereafter R17) measured the radii of several hundred clusters spread between two galaxies in the Legacy Extragalactic UV Survey (LEGUS).

In this paper we measure the projected half light radius (effective radius) of clus-

ters in the 31 galaxies with publicly available cluster catalogs from LEGUS. Our method for fitting the radii is described in Section 2.2. In Section 2.3 we describe our findings of a cluster radius distribution common to most galaxies and a clear cluster mass-radius relation. In Section 2.4 we discuss how selection effects affect our results, calculate cluster densities, and present a toy model of cluster evolution. We close with a summary in Section 2.5.

2.2 Methods

2.2.1 The LEGUS Sample

We use the publicly available LEGUS dataset[1] to extend the sample of clusters with uniformly measured radii and densities to the 31 galaxies with currently available cluster catalogs. We summarize some of the key details of LEGUS in this section (see Calzetti et al. 2015 for more on the LEGUS survey description, Adamo et al. 2017 and Cook et al. 2019 for more on the cluster catalogs).

LEGUS is a Cycle 21 Treasury program on HST that collected imaging with WFC3/UVIS to supplement archival ACS/WFC imaging, producing five-band coverage from the near-UV to the I band for 50 galaxies. Within each field, a uniform process was used to identify cluster candidates.

First, SourceExtractor (Bertin & Arnouts, 1996) is used to find sources in the white-light images (a combination of the images in all five filters, weighted by S/N). Next, a user selects a training set of objects that are clearly clusters or stars. The pipeline calculates the concentration index (CI), which is the magnitude difference between apertures of radius 1 pixel and 3 pixels (Holtzman et al., 1992; Whitmore et al., 2010). The user then selects a value of the CI that separates stars from clusters. The aperture for science photometry is chosen as the integer number of

[1] `https://archive.stsci.edu/prepds/legus/dataproducts-public.html`

pixels containing at least 50% of the cluster flux, based on the curve of growth of the clusters. This science photometry is done with the same aperture in all filters, using a sky annulus located at 7 pixels with a width of 1 pixel.

Next, aperture corrections for each filter are determined using an averaged method and CI-based method. In this study we use catalogs using the averaged aperture correction method, so we describe that here. The correction is the difference between the magnitude of the source obtained using an aperture of 20 pixels and the magnitude obtained from the science aperture. The average aperture correction for a user-defined set of well-behaved clusters is used for all clusters. This photometry is corrected for galactic foreground extinction (Schlafly & Finkbeiner, 2011). Sources that are detected in at least four filters with a photometric error of less than 0.3 mag, have an absolute V-band magnitude brighter than -6, and have a larger CI than the limit determined earlier, are then visually inspected by the LEGUS team. Sources that do not pass some of these cuts can be manually added by LEGUS team members, but the number is small.

Three or more team members visually inspect each cluster candidate, classifying it into one of the following four classes. Class 1 objects are compact and centrally concentrated with a homogeneous color. Class 2 clusters have slightly elongated density profiles and a less symmetric light distribution. Class 3 clusters are likely compact associations, having asymmetric profiles or multiple peaks on top of diffuse underlying wings. Class 4 objects are stars or artifacts.

The age, mass, and extinction of each cluster are determined by the SED fitting code *Yggdrasil* (Zackrisson et al., 2011). Versions of the catalogs are created with different stellar tracks and extinction laws, but in this paper we select the catalogs that use the MW extinction law (Cardelli et al., 1989) and Padova-AGB tracks available in Starburst99 (Leitherer et al., 1999; Vázquez & Leitherer, 2005).

The survey targeted 50 galaxies. Currently, public cluster catalogs are available

for 31 of these galaxies. Some galaxies are observed with multiple fields, resulting in 34 total fields with cluster catalogs. Table 2.1 lists these fields along with some key properties of the galaxies.

2.2.2 Outline of the Measurement Procedure

To fit the cluster radii, we implement a custom pipeline written in Python. We choose to implement our own method to have full control over the fitting process and better quantify the distribution of errors of the fit parameters. It is analogous to that in the popular package GALFIT (Peng et al., 2002, 2010), but with automated masking and several other features described in Section 2.2.4 that make it more robust against contamination from other nearby sources and ensure a good estimate of the local background. We assume an EFF density profile for young clusters (Elson et al., 1987; Larsen, 1999; McLaughlin & van der Marel, 2005), then convolve it with the empirically-measured point spread function (PSF) before comparing to the data.

We use the F555W images in 25 of the 34 fields, but for the other 9 fields that were not observed in F555W, we instead use F606W. All images have a pixel scale of 39.62 mas pixel^{-1}. The LEGUS mosaics have units of $e^{-}s^{-1}$, and we multiply by the exposure time to convert to the electron count. We use the recommended LEGUS cluster catalogs that adopt the MW extinction, Padova stellar evolutionary tracks, and the averaged aperture correction method (Adamo et al., 2017). The LEGUS team created visual classification tags by visually inspecting each cluster with multiple team members. We select clusters that were identified as being concentrated and either symmetric (class 1) or with some degree of asymmetry (class 2). We use the mode of the classifications from multiple team members. Additionally, machine learning classifications are available for several galaxies (Grasha, 2018; Grasha et al., 2019). For NGC 5194 and NGC 5195, we use the human classifications for clusters where those are available, and supplement with machine learning classifications for clusters

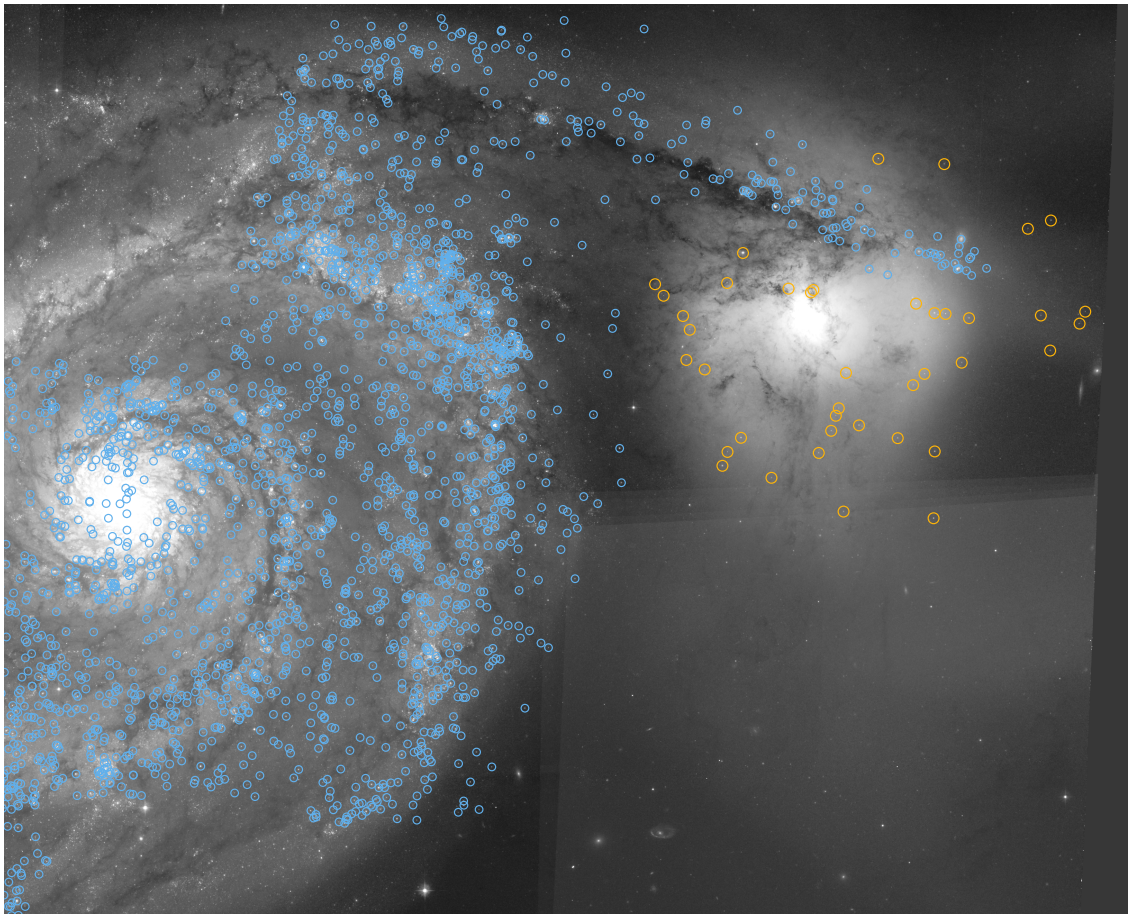

Figure 2.1 A portion of the F555W image of the LEGUS field with NGC 5194 and NGC 5195, illustrating our selection of cluster belonging to NGC 5194 (blue circles) and NGC 5195 (orange circles).

not inspected by humans. In NGC 1566, we use the hybrid classification system created by the LEGUS team, where some clusters are inspected by humans only, some by machine learning only, and some with a machine learning classification verified by humans. We do not use the machine learning classifications for NGC 4449, as we find that the machine-classified clusters have significantly different concentration index and radius distributions from the rest of the sample. This selection produces a final sample of 7242 clusters. As NGC 5194 and NGC 5195 are in the same LEGUS field, we manually identify their host galaxy. Figure 2.1 shows this selection.

Table 2.1: List of the LEGUS fields included in the sample, with some key properties of their cluster populations. We use the TRGB based distances from Sabbi et al. (2018) except for NGC 1566 – see the discussion at the end of Section 2.2.4. Stellar masses and specific star formation rates are from Calzetti et al. (2015), who obtained SFR from *GALEX* far-UV corrected for dust attenuation, as described in Lee et al. (2009), and stellar mass from extinction-corrected B-band luminosity and color information, as described in Bothwell et al. (2009) and using the mass-to-light ratio models of Bell & de Jong (2001).

LEGUS Field	N	log $M_\star$ ($M_\odot$)	log sSFR (yr^{-1})	Distance (Mpc)	PSF size (pc)	Cluster $R_{\rm eff}$: 25–50–75th percentiles (pc)
IC 4247	5	8.08	-10.18	5.11 ± 0.4	1.48	2.30 – 3.52 – 4.36
IC 559	21	8.15	-10.45	10.0 ± 0.9	2.74	4.05 – 4.85 – 6.34
NGC 1313-e	137	9.41	-9.35	4.30 ± 0.24	1.26	1.41 – 2.35 – 3.00
NGC 1313-w	276	9.41	-9.35	4.30 ± 0.24	1.26	1.13 – 2.54 – 3.58
NGC 1433	112	10.23	-10.80	9.1 ± 1.0	2.53	1.11 – 1.79 – 3.20
NGC 1566	881	10.43	-9.68	15.6 ± 0.6	4.33	3.00 – 4.30 – 6.28
NGC 1705	29	8.11	-9.07	5.22 ± 0.38	1.35	2.73 – 3.26 – 4.20
NGC 3344	237	9.70	-9.76	8.3 ± 0.7	2.37	1.72 – 2.40 – 3.53
NGC 3351	19	10.32	-10.13	9.3 ± 0.9	2.75	0.88 – 2.55 – 5.17
NGC 3738	142	8.38	-9.54	5.09 ± 0.40	1.52	2.17 – 3.25 – 4.30
NGC 4242	12	9.04	-10.04	5.3 ± 0.3	1.54	1.28 – 2.47 – 3.11
NGC 4395-n	20	8.78	-9.25	4.54 ± 0.18	1.24	1.34 – 1.84 – 2.42
NGC 4395-s	95	8.78	-9.25	4.54 ± 0.18	1.33	0.49 – 0.82 – 1.76
NGC 4449	425	9.04	-9.07	4.01 ± 0.30	1.17	1.68 – 2.40 – 3.32
NGC 45	14	9.52	-9.97	6.8 ± 0.5	1.84	2.41 – 3.25 – 4.61
NGC 4656	184	8.60	-8.90	7.9 ± 0.7	2.11	2.35 – 3.32 – 4.15
NGC 5194	2921	10.38	-9.54	7.40 ± 0.42	2.16	1.29 – 2.17 – 3.29
NGC 5195	40	10.36	-10.82	7.40 ± 0.42	2.16	2.69 – 3.39 – 4.88
NGC 5238	8	8.15	-10.15	4.43 ± 0.34	1.29	2.29 – 2.63 – 2.94

LEGUS Field	N	log $M_\star$ ($M_\odot$)	log sSFR (yr^{-1})	Distance (Mpc)	PSF size (pc)	Cluster $R_{\rm eff}$: 25–50–75th percentiles (pc)
NGC 5253	57	8.34	-9.34	3.32 ± 0.25	0.98	1.09 – 1.86 – 2.66
NGC 5474	143	8.91	-9.48	6.6 ± 0.5	1.95	2.28 – 3.25 – 4.13
NGC 5477	14	7.60	-9.12	6.7 ± 0.5	1.98	2.40 – 3.64 – 4.78
NGC 628-c	691	10.04	-9.48	8.8 ± 0.7	2.59	1.70 – 2.43 – 3.44
NGC 628-e	172	10.04	-9.48	8.8 ± 0.7	2.49	1.95 – 2.64 – 4.05
NGC 6503	167	9.28	-9.77	6.3 ± 0.5	1.69	1.31 – 2.12 – 3.34
NGC 7793-e	108	9.51	-9.79	3.79 ± 0.20	0.92	0.60 – 1.02 – 2.20
NGC 7793-w	135	9.51	-9.79	3.79 ± 0.20	1.11	0.66 – 1.26 – 2.20
UGC 1249	48	8.74	-9.56	6.4 ± 0.5	1.88	1.83 – 2.58 – 3.26
UGC 4305	45	8.36	-9.28	3.32 ± 0.25	0.97	0.69 – 1.15 – 1.91
UGC 4459	7	6.83	-8.99	3.96 ± 0.30	1.23	1.83 – 2.51 – 3.13
UGC 5139	9	7.40	-9.10	3.83 ± 0.29	1.12	1.07 – 1.54 – 4.00
UGC 685	11	7.98	-10.13	4.37 ± 0.34	1.31	1.77 – 2.23 – 2.96
UGC 695	11	8.26	-9.95	7.8 ± 0.6	2.08	1.39 – 6.21 – 7.49
UGC 7408	35	7.67	-9.67	7.0 ± 0.5	2.08	3.75 – 4.41 – 5.68
UGCA 281	11	7.28	-8.98	5.19 ± 0.39	1.51	3.12 – 3.50 – 4.25
Total	7242	–	–	–	–	1.53 – 2.48 – 3.69

2.2.3 PSF Creation

Like GALFIT, our method convolves the PSF with the model image, and compares the result with the observed data. To produce this PSF, we use `Photutils` (Bradley et al., 2019), an Astropy package (Astropy Collaboration et al., 2013, 2018). We manually select bright isolated stars in each field, then use the `EPSFBuilder` class of `Photutils` to create a separate PSF for each field. `EPSFBuilder` follows the prescription of Anderson & King (2000). The final PSF images are 15 pixels (0.59") wide, and we spatially subsample the PSF by a factor of two, producing a PSF with twice the spatial resolution of the input image. We do not choose higher values, as this significantly increases the computational cost of the fitting procedure (particularly the convolution). As shown in Figure 2.9, our results are consistent with those of R17 for NGC 1313 and NGC 628, even for the smallest clusters, indicating that this oversampling factor is adequate. We use `Photutils`'s "quadratic" smoothing kernel, which is the polynomial fit with degree=2 to 5x5 array of zeros with 1 at the center. We found that the other options gave unphysically non-smooth PSFs. Figure 2.2 illustrates our created PSFs.

2.2.4 Fitting Cluster Parameters

We fit the clusters with the EFF surface brightness profile, as it accurately describes the light profiles of young star clusters (Elson et al., 1987; Larsen, 1999; McLaughlin & van der Marel, 2005; Bastian et al., 2012; Ryon et al., 2015, 2017; Cuevas-Otahola et al., 2020). Assuming circular symmetry, it takes the form

$$\mu(r) = \mu_0 \left(1 + \frac{r^2}{a^2}\right)^{-\eta} \tag{2.1}$$

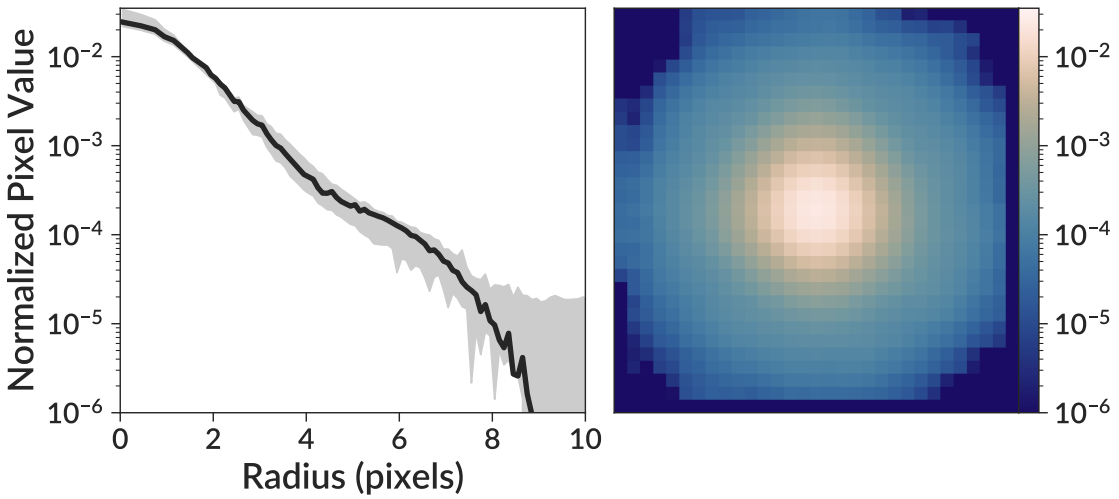

Figure 2.2 Visualization of one of our PSFs, from the NGC 1313-e field. The solid line in the left panel shows the azimuthally-averaged radial profile of this PSF, while the shaded region shows the range of the PSF profiles from all other fields, normalized to integrate to unity. The right panel shows an image of the NGC 1313-e PSF.

where μ is the surface brightness, a is the scale radius, and η is the power law slope. As real clusters are typically not circularly symmetric, we include ellipticity as follows:

$$\mu(x,y) = \mu_0 \left(1 + \left[\frac{x'(x,y)}{a}\right]^2 + \left[\frac{y'(x,y)}{aq}\right]^2\right)^{-\eta} \tag{2.2}$$

where q is the ratio of the minor to major axes of the ellipse ($0 < q \leq 1$). We have rotated the image coordinate system by angle θ about the cluster center (x_c, y_c) to new coordinates (x', y') as follows:

$$x'(x,y) = (x - x_c)\cos\theta + (y - y_c)\sin\theta \tag{2.3}$$

$$y'(x,y) = -(x - x_c)\sin\theta + (y - y_c)\cos\theta. \tag{2.4}$$

Here x' is aligned with the cluster major axis, while y' is aligned with the minor axis. This gives 7 cluster parameters: μ_0, x_c, y_c, a, q, θ, η. We also leave the local background $f_{\rm BG}$ as a free parameter, giving 8 total parameters to fit.

We perform this fit on a 30×30 pixel snapshot centered on the cluster, follow-

ing R17. We tested larger snapshot sizes (40 and 50 pixels) but found no significant differences in fitted cluster radii, even for the biggest clusters where a larger snapshot could potentially allow for a better determination of the local background. These larger snapshot sizes also included more contaminating sources, leading to more catastrophic fit failures. The 30-pixel snapshot minimizes these failures while still performing well on the largest clusters.

To account for contaminating sources inside this 30×30 pixel snapshot, we mask star-like sources identified by the `IRAFStarFinder` class of `Photutils`. Any pixels within 2×FWHM of the stars are masked. However we discard any stars whose masked region would extend within 3 pixels of the cluster center, as well as any sources with a peak pixel value less than 2 times the local sky background identified by `IRAFStarFinder`. This second criterion was added to stop the masking of substructure in the most extended clusters. We also mask any pixels that are within 6 pixels of another star cluster, in cases where two star clusters are close to each other.

However, an automated masking system cannot solve all issues with contamination. To make our fitting method robust to potential contamination, we have made substantial modifications compared to a GALFIT-like method.

Our best-fit parameters maximize the posterior distribution, defined as

$$\log P_{\text{posterior}} = -\frac{1}{2}\sum_{x,y} w(x,y)\left|\frac{f_d(x,y) - f_m(x,y)}{\sigma(x,y)}\right| + \log P_{\text{prior}} \tag{2.5}$$

where x and y are pixel coordinates, w are pixel weights, f_d is the data value at pixel (x, y), f_m is the model, $\sigma(x, y)$ is the pixel uncertainty, and P_{prior} is the prior distribution. We will expand on each of these components in turn. Note that we use the absolute value of the differences between the model and data rather than the more typical square. As the square weights large differences more heavily, it has the effect of increasing the attention the fit pays to unmasked contamination, as these

pixels have large deviations. Using the absolute value instead produces fits that are less affected by contamination.

In addition, pixel weights are used to reduce the effect of contamination, particularly at large distances from the cluster. We weight each pixel proportional to $1/r$ so that each annulus has the same weight. To avoid giving dramatically more weight to the most central pixels, all pixels within 3 pixels from the center receive the same weight. We use the distance from the center of the cluster to determine the radius, giving $r^2 = (x - x_c)^2 + (y - y_c)^2$ and

$$w(r) = \begin{cases} 1 & \text{if } r \leq 3 \\ 3/r & \text{if } r > 3 \end{cases} \tag{2.6}$$

Giving equal weight to each annulus stops the large number of pixel values at large radius from dominating the fit, effectively increasing the focus on the cluster at the center.

We break the pixel uncertainty $\sigma(x, y)$ into two components: image-wide sky noise plus Poisson noise from individual sources:

$$\sigma^2(x, y) = \sigma^2_{\text{sky}} + f_d(x, y) \tag{2.7}$$

where $f_d(x, y)$ is the pixel value in electrons, and equals the Poisson variance. To calculate the global sky noise, we use the 3-sigma clipped standard deviation of the pixel values of the entire image.

The model component $f_m(x, y)$ is the convolution of the underlying cluster model with the PSF plus the local background f_{BG}, which we assume to be constant over the fitting region. We subsample the pixels for both the model and the empirical

PSF, then rebin the resulting model image to the same scale as the data:

$$f_m(x, y) = \sum_{x_s \in x} \sum_{y_s \in y} (\mathrm{PSF} * \mu)(x_s, y_s) + f_{\mathrm{BG}} \tag{2.8}$$

where x_s and y_s represent subpixel positions that are not integer pixel values like x and y, $*$ represents convolution, and μ is the functional form of the fitted profile given by Equation 2.2.

The last component of Equation 2.5 is the prior distribution. We employ a prior on the local background. This is needed because at very low values of η (shallow power law slopes), the background can be incorrectly fit by this cluster component rather than a truly flat background. This attributes light to the cluster that should be attributed to the background, incorrectly inflating the enclosed light and therefore R_{eff}. Additionally, η is strongly degenerate with a, so as to give the same value of μ at some typical radius. Low values of η caused by incorrect background fits also lead to unphysically small values for a. We find that constraining the background addresses these issues. We first estimate the background and background uncertainty by using sigma clipping to calculate the mean (μ_{BG}) and standard deviation (σ_{BG}) of all pixels farther than 6 pixels from the cluster center. The mean is used as the mean value of a Gaussian prior on the background. As the background becomes an issue for low values of η, we condition the width of our prior on it. We use a logistic function that produces a tight prior $\sigma_{\mathrm{prior}} = 0.1\sigma_{\mathrm{BG}}$ for low values of η, while giving a looser constraint $\sigma_{\mathrm{prior}} = \sigma_{\mathrm{BG}}$ for higher values of η. This takes the form

$$\log P_{\mathrm{prior}} = -\frac{1}{2}\left(\frac{f_{\mathrm{BG}} - \mu_{\mathrm{BG}}}{\sigma_{\mathrm{prior}}}\right)^2 \tag{2.9}$$

where

$$\sigma_{\mathrm{prior}}(\eta) = \sigma_{\mathrm{BG}}\left(0.1 + \frac{0.9}{1 + e^{10(1-\eta)}}\right) \tag{2.10}$$

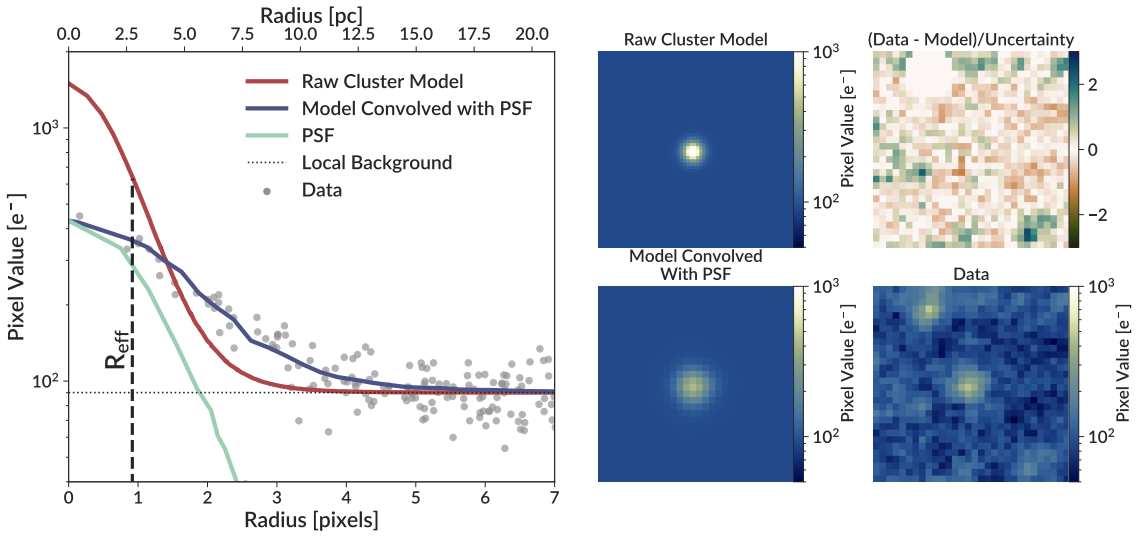

Figure 2.3 An example of our fitting process. The right set of panels show (in counterclockwise order, starting from the top left) the 2D pixel values for the raw cluster model, the cluster model after being convolved with the PSF, the data, then the residuals after subtracting the PSF-convolved model from the data. Note the masking of the contaminating object in the top right panel. The left panel shows the radial profiles for these components (with the PSF normalization adjusted accordingly). Note that the radial profile is included for illustrative purposes only, the fitting is done using the 2D images.

The combined effect of using the absolute value of differences, the pixel weights, and the prior on the background is to produce cluster fits that more closely match the cluster itself. They prevent the fit from being drawn towards any surrounding structures, while also allowing the local background to be fit appropriately. This produces trustworthy cluster parameter values. Lastly, to ensure full numerical convergence of the fit, we use multiple sets of initial values for the fit parameters, selecting the final parameter set corresponding to the highest likelihood. An example of a cluster fitted with our method is shown in Figure 2.3.

To determine the distribution of errors of the fit parameters, we perform bootstrapping on the pixels in the snapshot. We split the snapshot into two regions: pixels closer than 9 pixels from the cluster center and pixels outside this region. For the inner region, we resample individual pixels with replacement. For the outer region,

we group the pixels into 5x5 blocks, then resample those blocks with replacement. Using blocks in the outskirts does a better job accounting for any missed or faint contaminating sources. If we were to use individual pixels, at least some of the pixels from these sources would be included in a given resampling, while using blocks allow us to exclude these sources completely in certain iterations, giving a better estimate of how these sources affect the cluster fit. Using individual pixels in the center is necessary as the cluster itself may be roughly the size of the 5x5 chunk. Figure 2.4 shows an example of the pixels included in one randomly selected bootstrap realization.

We run the bootstrap realizations until convergence of the fit parameters. Every 20 iterations, we calculate the standard deviation of the distributions of all 8 fit parameters in the accumulated iterations, then compare them to the standard deviations from the last time it was calculated. We stop bootstrapping when the standard deviation of each parameter changes by less than 10 percent. Most clusters required 100–140 iterations to converge. Our reported uncertainties on $R_{\rm eff}$ are marginalized over all other parameters. We use $R_{\rm eff}$ calculated using the original snapshot as the best fit value, then take the 16–84th percentile range of $R_{\rm eff}$ from the bootstrap iterations as the uncertainty.

As we measure everything in pixel values, we need to convert to physical length units. To do this we use the TRGB distances to all LEGUS galaxies provided by Sabbi et al. (2018). That work provides independent estimates of distance for each field. For galaxies split between two fields, we use the mean of the two distance estimates for both fields. Lastly, NGC 1566 has an unreliable distance estimate. The TRGB was too faint to be detected in Sabbi et al. (2018), and available values in the literature span a wide range (from 6 to 20 Mpc). However, NGC 1566 was identified as being part of the group centered on NGC 1553, which has a measured distance and group radius (Kourkchi & Tully, 2017; Tully et al., 2016). For NGC 1566, we adopt the distance to NGC 1553 with uncertainty of the group radius.

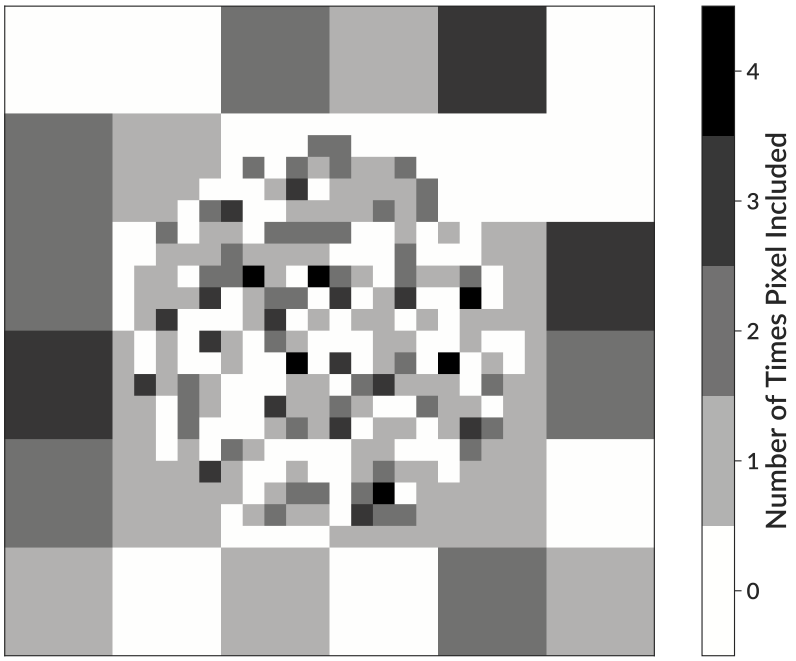

Figure 2.4 An example of the pixels included in one randomly selected bootstrap iteration. The central region is resampled on a pixel-by-pixel basis, while the outskirts are resampled on 5×5 pixel blocks.

2.2.5 Converting to Effective Radius

The effective radius is defined to be the circular radius that contains half of the projected light of the cluster profile. For a circularly symmetric EFF profile, this is

$$R_{\rm eff} = a\sqrt{2^{\frac{1}{\eta-1}} - 1} \tag{2.11}$$

However, this equation asymptotically approaches infinity as η approaches unity, and the total light of the profile is infinite for $\eta \leq 1$. As some cluster fits prefer values of η near or below unity, we implement a maximum radius for the cluster profiles, removing this infinity and allowing the effective radius to be well defined for any value of η. We choose the size of our box (15 pixel radius) as $R_{\rm max}$. When using a maximum radius, the effective radius for the circularly symmetric EFF profile is

$$2\left[1 + \left(\frac{R_{\rm eff}}{a}\right)^2\right]^{1-\eta} = 1 + \left[1 + \left(\frac{R_{\rm max}}{a}\right)^2\right]^{1-\eta}. \tag{2.12}$$

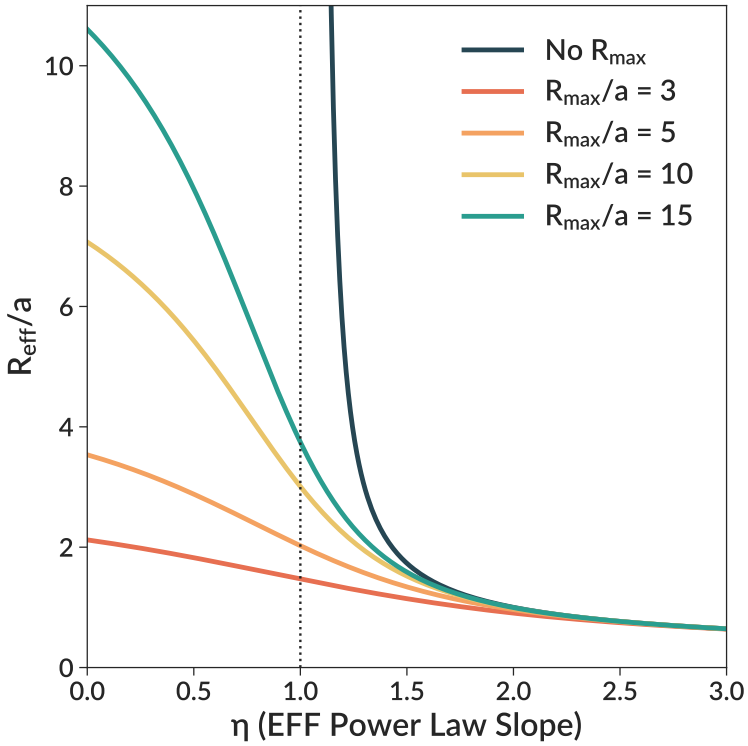

Figure 2.5 Comparison of the effective radius when calculated with (Eq. 2.12) or without (Eq. 2.11) a maximum radius. We show this for several representative values of $R_{\rm max}/a$, as clusters have values across this full range.

For $\eta \gtrsim 1.5$, this agrees well with Equation 2.11, as shown in Figure 2.5.

A correction is required for an elliptical profile. We empirically determine this correction for the EFF profile as a function of η and q, by performing numerical integration of elliptical EFF profiles. We determine the circular aperture that contains half of the total light. We use a circular maximum radius $R_{\rm max}/a = 10$ for this integration, although we find that the results do not depend strongly on the chosen $R_{\rm max}$. At a given η, we find that the relation between true effective radius and the effective radius calculated assuming a circularly symmetric profile (Equation 2.12) is linear with q, so we parametrize the correction as

$$\frac{R_{\rm eff,true}}{R_{\rm eff,circ}} = 1 + m(q-1) \tag{2.13}$$

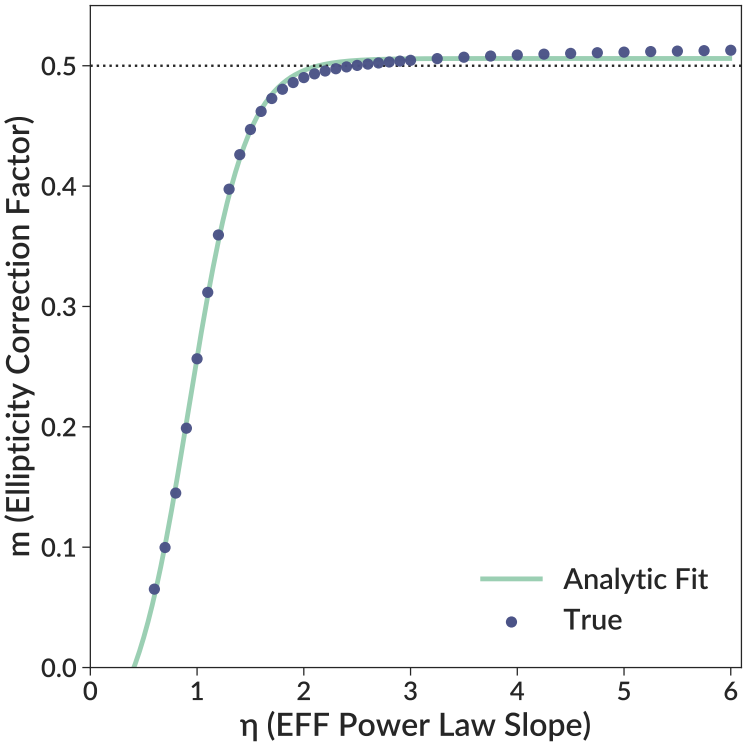

Figure 2.6 Numerical calculations of the ellipticity correction factor m (see Equation 2.13), along with the analytic fit (Equation 2.15). The horizontal dotted line shows $m = 0.5$, as used in Equation 2.14.

where m is a function of η. With $m = 0.5$, we obtain the commonly used correction

$$\frac{R_{\text{eff,true}}}{R_{\text{eff,circ}}} = 0.5(1 + q) \tag{2.14}$$

as found in the ISHAPE manual (Larsen, 1999). We measure m as a function of η, and find it is well fit by a logistic-type function, with an RMS deviation of only 0.0043. This gives a final correction of the form:

$$\frac{R_{\text{eff,true}}}{R_{\text{eff,circ}}} = 1 + \left(\frac{0.579}{1 + \exp\left(\frac{0.924 - \eta}{0.266}\right)} - 0.073 \right) (q - 1) \tag{2.15}$$

The results of the numerical integration and the fit are shown in Figure 2.6.

2.2.6 Cluster Fit Quality

While our fitting procedure is designed to be robust, it does not perform perfectly on all clusters. We exclude clusters with unrealistic parameter values, which we define as a scale radius $a < 0.1$ pixels, $a > 15$ pixels, or an axis ratio $q < 0.3$. We also exclude any clusters where the fitted center is more than 2 pixels away from the central pixel identified by LEGUS. This eliminates 6.7% percent of the sample.

Additionally, to quantitatively evaluate which clusters have poor fits, we implement a quality metric based on a comparison of the cumulative light profiles of the cluster data and the model (after subtracting the best-fit background from both the data and model). As our primary goal is to evaluate the reliability of $R_{\rm eff}$, the cumulative profile is a strong indicator as it probes all light enclosed within a given radius.

Specifically, our metric uses the cumulative light profile to estimate the half-light radius of the cluster non-parametrically, then compares the enclosed light of the model and data within this radius. The relative difference is

$$d = \left| \frac{F_{\rm model}(< R_{1/2}) - F_{\rm data}(< R_{1/2})}{F_{\rm data}(< R_{1/2})} \right| \tag{2.16}$$

where the non-parametric radius $R_{1/2}$ is defined by

$$F_{\rm data}(< R_{1/2}) = 0.5 F_{\rm data}(< 15 \text{ pixels}). \tag{2.17}$$

Here $F(< R)$ is the cumulative flux enclosed within a circular radius R. We use 15 pixels as the maximum radius as it is the radius of the individual cluster snapshots.

We then calculate the distribution of this metric d for clusters that pass the cuts mentioned at the first paragraph of this section, shown in Figure 2.7. The knee of the cumulative distribution is at approximately the 90th percentile, so we use that

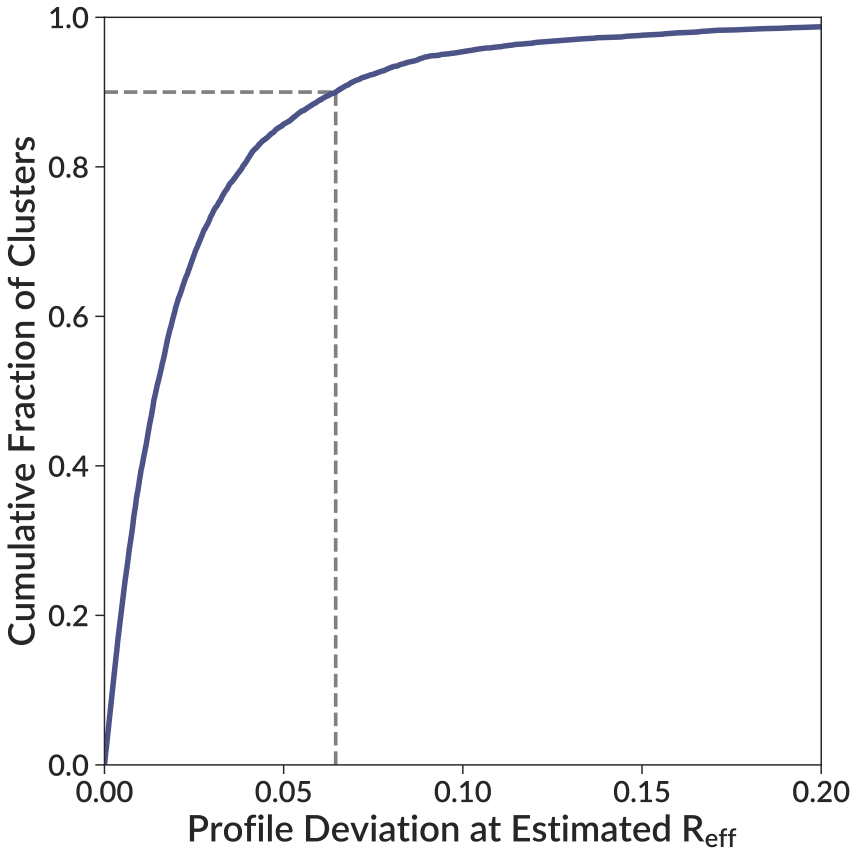

Figure 2.7 Cumulative distribution for the deviation in the cumulative light profile as described in Section 2.2.6. Our analysis excludes clusters that fall above the 90th percentile of this metric, corresponding to approximately 6.5% deviation of the integrated light within non-parametric $R_{1/2}$.

percentile as our cut. Any cluster above the 90th percentile will be excluded from the analysis in the rest of this paper. This results in a final sample of 6097 clusters with reliable radius measurements. This 90th percentile cut corresponds to about 6.5% error on the light enclosed within the estimated effective radius, indicating the high quality of the fits we keep.

2.2.7 Artificial Cluster Tests

To test the ability of the pipeline to recover the effective radius of small clusters, we perform artificial cluster tests. We generate 150 synthetic clusters following the EFF profile. These clusters have magnitudes from 20 to 24, $1.25 \leq \eta \leq 2.5$, and

$0.03 \leq R_{\rm eff} \leq 3$ pixels. All parameters are uniformly distributed within these ranges. To calculate the cluster magnitude, we follow the LEGUS pipeline as described in Adamo et al. (2017). We use a circular aperture with a radius of 4 pixels (chosen as the integer pixel value that contains 50% of the flux of typical clusters) and a local sky annulus located at 7 pixels with a width of 1 pixel, then apply the average aperture corrections for the NGC 628-c field. The artificial clusters span the range of the magnitudes of real clusters in this field.

We convolve these models with the PSF for the NGC 628-c field, add Poisson noise, and insert these artificial clusters into the NGC 628-c field. We then run our pipeline on this new image to measure their effective radii. The results of this test are shown in Figure 2.8.

The pipeline is able to accurately measure cluster radii down to about 0.3 pixels. Below this point, the pipeline systematically overestimates the true radius. The performance of the pipeline does depend on magnitude, as faint clusters with magnitude 24 (the limit of the clusters in the LEGUS catalog for NGC 628-c) have a much wider dispersion than brighter clusters, even at larger radii. A visual examination of these fits shows that contamination and noise are the primary causes of this dispersion. Faint sources rise above the background less than bright sources, so variations in the background can influence the fit more. Additionally, the Poisson pixel noise of the source itself can influence the fit, even for artificial clusters placed in a region where the background is smooth. Due to the nature of Poisson noise, this affects faint clusters the most.

Figure 2.8 also shows the ability of the pipeline to detect when clusters have poor fits. Many of the catastrophic failures are correctly identified as failures. However, for the very compact clusters the pipeline identifies many fits that actually overestimate $R_{\rm eff}$ as reliable. This is likely because when $R_{\rm eff}$ is much smaller than the PSF, the observed cluster is not very different from the PSF. A slightly larger model is still very

similar to the data, making the pipeline identify the fit as a success. This also means that generally the error bars may be underestimated for clusters with $R_{\rm eff} < 0.3$ pixels. Nevertheless, for most clusters our pipeline gives reliable measurements. Importantly, if $R_{\rm eff}$ is measured to be small, then it really is small. These trends are true for all values $\eta \gtrsim 1.25$. For values of $\eta \sim 1$, we found larger scatter.

2.3 Results

2.3.1 Comparison to Ryon et al. (2017)

R17 used GALFIT to measure the effective radii of clusters in NGC 1313 and NGC 628, two of the galaxies in the LEGUS sample. To validate our method, we compare our measurements to those of R17. When making this comparison, we perform a separate round of fitting with several modifications to our method to match what was done in R17. We do not mask any contaminating sources, do not use radial weighting, and use the square of differences rather than the absolute value. When postprocessing these results, we do not use a maximum radius to calculate the effective radii (instead using Equation 2.11), we use the simple ellipticity correction (Equation 2.14), and we use the same distances to NGC 1313 and NGC 628 as R17 did (Jacobs et al., 2009; Olivares E. et al., 2010). These changes ensure consistency with the R17 method.

The left panel of Figure 2.9 shows the results of that comparison. Following R17, only clusters with $\eta > 1.3$ are shown in this plot, as $R_{\rm eff}$ given by Equation (2.11) is unreliable for lower values. To quantify the deviation, we use the RMS error, defined as

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum \frac{\left(R_{\rm eff,R17} - R_{\rm eff}\right)^2}{\sigma_{\rm R17}^2 + \sigma^2}} \tag{2.18}$$

where σ and $\sigma_{\rm R17}$ are the error on $R_{\rm eff}$ in this work and R17, respectively. This RMS deviation is 0.55, indicating excellent agreement.

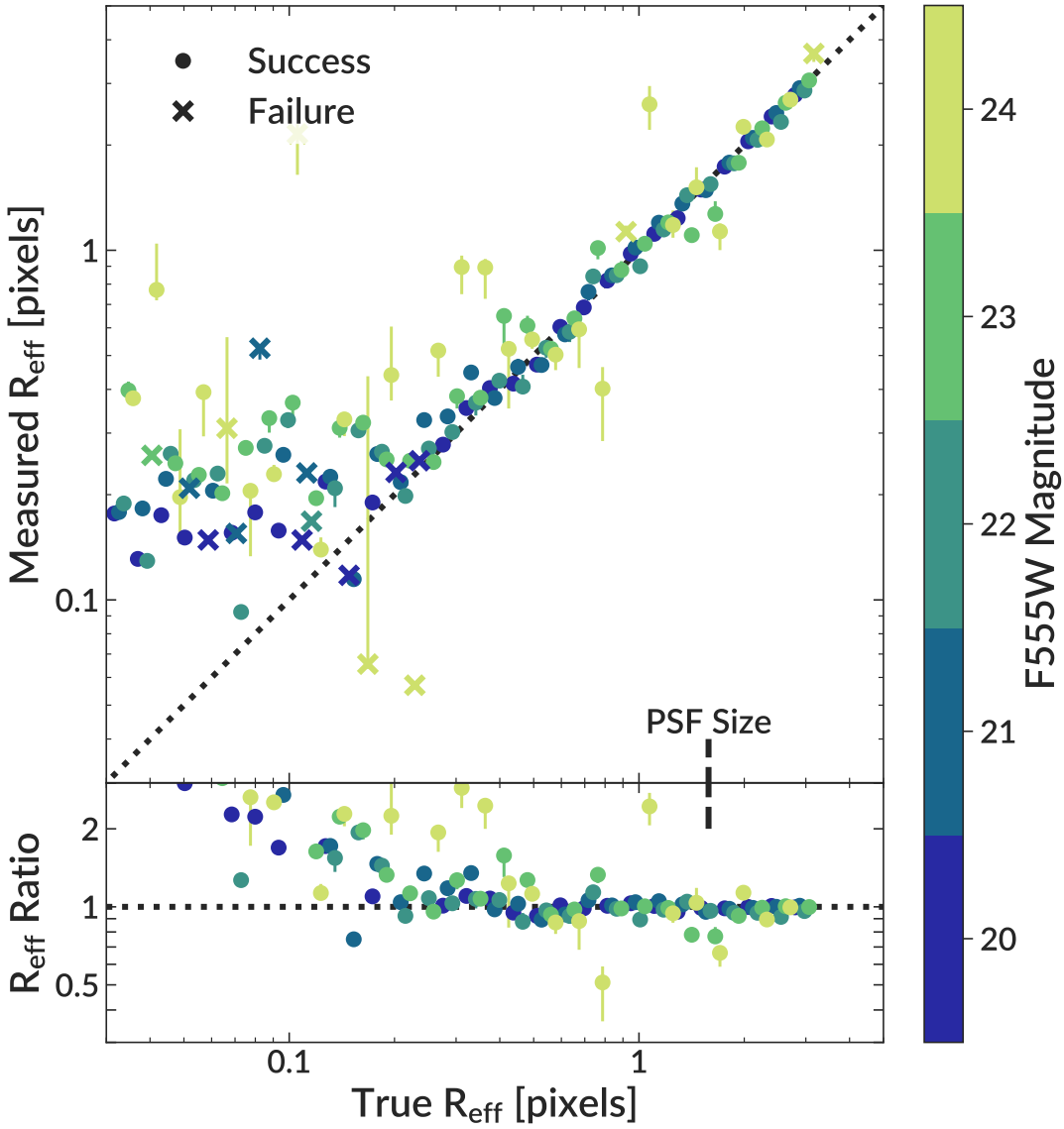

Figure 2.8 Results of artificial cluster tests. The top panel shows a comparison of the true effective radius to that measured by our pipeline. Solid circles show fits that the pipeline identified as successful, while the crosses show failures. The bottom panel shows the ratio of the measured effective radius to the true effective radius. Only successful fits are shown in this bottom panel. The dashed line spanning both panels indicates the PSF size in pixels for the NGC 628-c field, which is the field into which clusters were inserted.

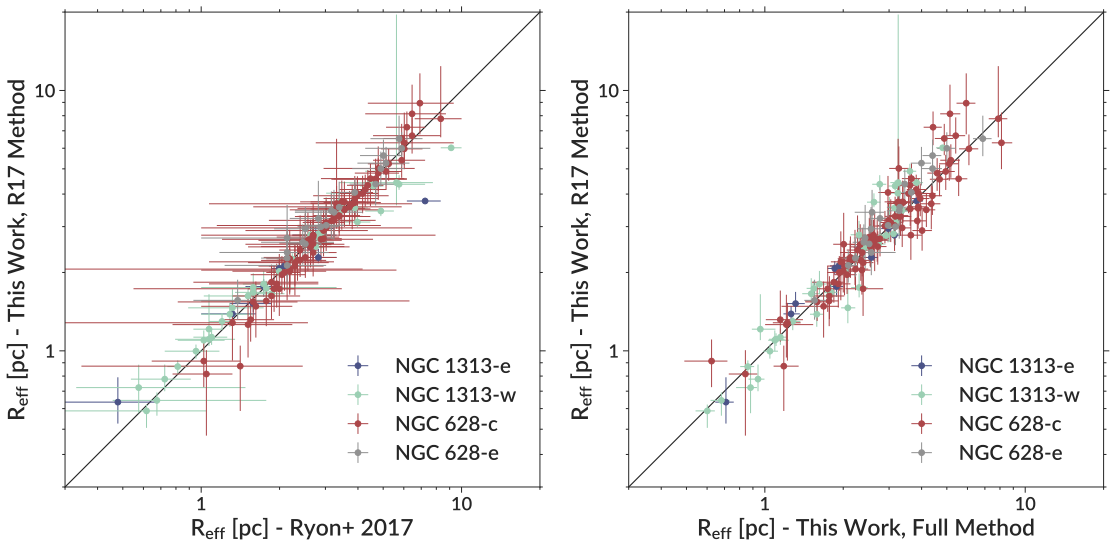

Figure 2.9 Comparison of our cluster effective radii using different fitting methods for NGC 628 and NGC 1313. The left panel shows a comparison to those of Ryon et al. (2017) using the same fitting method. The right panel shows a comparison of this method to our full method used in the rest of the paper.

In the right panel of Figure 2.9 we show a comparison of our full method to our R17-like method. These two methods show good agreement across the full radius range, with no significant deviations. The RMS deviation here is 1.62. This higher value is primarily driven by the error bars, which are smaller than those obtained in R17. In addition, a small number of clusters have significantly different radii between the two methods. The majority of these discrepant clusters have a small value for η (typically just above the cutoff of $\eta > 1.3$ for inclusion in this plot), where the use of a maximum radius has the greatest effect on $R_{\rm eff}$ (see Figure 2.5).

In the rest of this paper we analyze the results obtained with our full fitting method.

2.3.2 Cluster Radius Distribution

In Figure 2.10 we show the distribution of effective radii measured in the entire LEGUS sample with our full fitting method. For most galaxies the distributions are

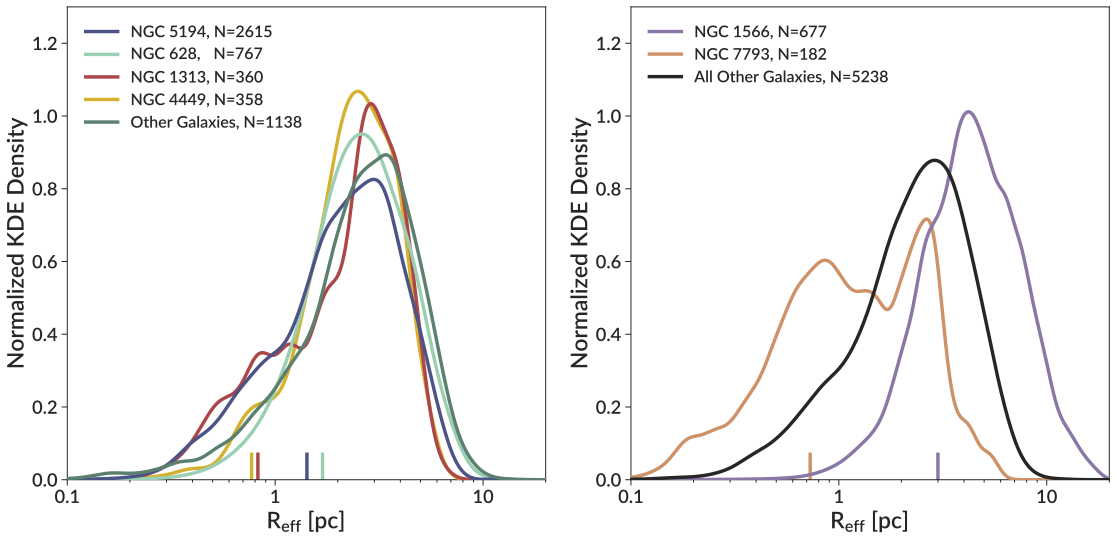

Figure 2.10 Kernel density estimation of the cluster radius distributions of the galaxies with the most clusters. The line for each galaxy shows the summed Gaussian kernels representing its clusters, where we use a width equal to twice each cluster's radius error. Each curve is normalized to the same area for comparison purposes. The left panel shows galaxies with similar cluster distributions, while the right panel shows two galaxies with different distributions. Note that the "Other Galaxies" in the left panel do not include the two discrepant galaxies NGC 1566 and NGC 7793 shown in the right panel, while the "All Other Galaxies" in the right panel include all galaxies shown in the left panel. The tick marks at the bottom show the pixel size in parsecs for the images of each galaxy.

remarkably similar. They have a peak at $R_{\rm eff} \approx 3$ pc, an extended tail to below 1 pc, and a sharper cutoff at large radii. This peak at $R_{\rm eff} \approx 3$ pc has been seen consistently in other studies of young clusters (e.g. Meurer et al., 1995; Larsen, 1999; Scheepmaker et al., 2007; Bastian et al., 2012; Ryon et al., 2015, 2017; Cantat-Gaudin et al., 2018). Galaxies with similar distributions are shown in the left panel of Figure 2.10, while two galaxies NGC 1566 and NGC 7793 with discrepant shapes are shown in the right panel and will be discussed below. While we only show several galaxies individually in the left panel of Figure 2.10 for clarity, an examination of all galaxies shows that they have very similar distributions. In Table 2.1 we include the quartiles of the cluster radius distribution of each galaxy as another method of quantifying their distributions.

To characterize this common distribution, we create a stacked distribution of clusters from all galaxies other than the two discrepant. The distribution is shown as the comparison line in the right panel of Figure 2.10 and has a sharp peak at 2.9 pc. We compared it to several common analytical distributions and found that neither normal nor lognormal functions are good fits, due to the asymmetric shape of the observed distribution. Instead we find that the Weibull distribution produces an excellent match:

$$\frac{dN}{dR_{\rm eff}} = \frac{k}{\lambda}\left(\frac{R_{\rm eff}-R_0}{\lambda}\right)^{k-1} \exp\left[-\left(\frac{R_{\rm eff}-R_0}{\lambda}\right)^{k}\right] \tag{2.19}$$

with $k = 2.22$, $\lambda = 3.64$ pc, and $R_0 = 0.185$ pc.

To quantitatively test whether the individual galaxy samples are statistically consistent with being drawn from the same distribution, we employ the one-sided Kolmogorov–Smirnov test. We find that of the 31 galaxies, 16 have $p-$value > 0.01 (13 with $p-$value > 0.05), indicating that they are not inconsistent with the stacked distribution. However, of the 13 galaxies with more than 50 clusters, only 4 have

$p > 0.01$ (2 with $p > 0.05$). The large number of clusters in these galaxies provides high statistical significance to formally distinguish the distributions. Still, the individual distributions exhibit strong visual similarity.

Two galaxies show cluster distributions significantly different from the rest. NGC 1566 appear shifted to larger radii than other galaxies. It has less low radius clusters, a peak at larger radii (4.2 pc compared to 2.9 pc for the stacked distribution), and more high radius clusters than any other galaxy. Selection effects may be partly responsible. At the adopted distance of NGC 1566 of 15.6 Mpc, 1 pixel covers 3 pc and our PSF model has an effective radius of 4.3 pc. Small clusters may not be resolved and therefore not included in the LEGUS catalog. While a full characterization of the LEGUS selection effects is beyond the scope of this paper, Adamo et al. (2017) examined the completeness as a function of cluster radius in NGC 628 at a distance of 8.8 Mpc. The concentration index cut excluded roughly 50% of clusters with $R_{\rm eff} = 1$ pc. This selection effect is not significant for most galaxies, as the peak of the cluster radius distribution is at higher radii, and most galaxies are closer than NGC 628. But since NGC 1566 is at approximately twice the distance of NGC 628, we can expect that its observations will be incomplete below 2 pc. This could explain the dearth of small clusters in NGC 1566, but would not explain the overabundance of large clusters. Our adopted distance could be responsible for this. As mentioned at the end of Section 2.2.4, the distance to NGC 1566 is uncertain, with distance estimates ranging from 6 to 20 Mpc. If our adopted distance of 15.6 Mpc is an overestimate, our cluster radius measurements will also be overestimated. Adopting a distance of 11 Mpc rather than 15.6 Mpc would bring it in line with the distributions of other galaxies, thus effectively treating this distribution like a standard ruler (Jordán et al., 2005). Future distance measurements may be able to resolve this and determine whether the cluster radius distribution in NGC 1566 is significantly different than that of other galaxies.

The other discrepant galaxy, NGC 7793, has a double peaked distribution that is much broader than that in other galaxies. One peak is near the 3 pc peak seen in other galaxies, while another is at $\sim$ 0.8 pc. The reason for this is unclear. While NGC 7793 is split into two fields, both fields show the same double peaked distribution. Its specific star formation rate is within the range of other galaxies. It is closer than most other galaxies in the sample, meaning the smallest clusters are more likely to be included, but other galaxies with similar distances do not show this bimodal distribution. A visual examination of the spatial distribution within NGC 7793 of the clusters belonging to each peak does not show any striking trends. An examination of the age and mass of the clusters shows that, compared to other galaxies, NGC 7793 has more small young clusters and no large young clusters. Additionally, the age distribution is bimodal, with a deficit of intermediate age clusters. Roughly speaking, this results in the low-radius peak being mostly young, low-mass clusters, while the high-radius peak is mostly old, high-mass clusters. Future detailed studies of NGC 7793 may be needed to understand its cluster population in more detail.

2.3.3 Cluster Mass-Radius Relation for All LEGUS Galaxies

Figure 2.11 shows the mass-radius relation for the clusters in our sample. A clear mass-radius relation is visible, albeit with a shallow slope. To guide the eye, lines indicate the 5, 25, 50, 75, and 95th percentiles of the radius at a given mass.

This plot shows the full mass range of the LEGUS clusters. However, masses below 5000 $M_\odot$ measured by the deterministic method used in LEGUS may be unreliable, as the assumption of a fully-sampled IMF is no longer valid (Maíz Apellániz, 2009; Silva-Villa & Larsen, 2011; Krumholz et al., 2015). To account for this, previous work with the LEGUS data, such as Adamo et al. (2017) and R17, restricted to clusters with masses above 5000 $M_\odot$. As LEGUS is not complete for 5000 $M_\odot$ clusters at old ages, these papers selected clusters younger than 200 Myr. This produces a complete

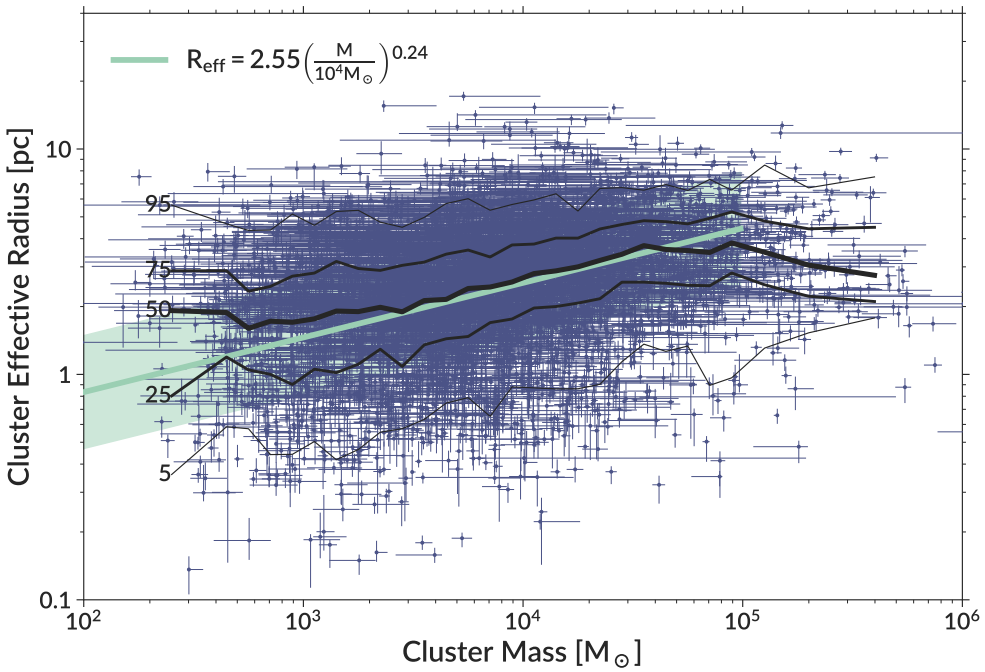

Figure 2.11 The mass-radius relation for the clusters in LEGUS. Clusters of all ages are included in this plot. The black lines show then 5, 25, 50, 75, and 95th percentiles of radii at a given mass. The solid line shows the best fit linear relation, with the shaded region showing the intrinsic scatter.

sample using clusters with the most reliable masses. We take a different approach in this paper, and use the full mass and age ranges of clusters with good SED fits, adjusting their errors to account for the systematic errors in the SED fitting. If we were to consider only clusters more massive than 5000 $M_\odot$, it would exclude about a third of the complete sample and greatly decrease the dynamic range of the mass-radius relation. Section 2.4.1 below discusses possible effects of incompleteness.

As discussed in Krumholz et al. (2015), clusters have poor fits – as quantified by the $Q(\chi^2)$ statistic – when the assumption of a fully-sampled IMF is violated. To avoid these unreliable low-mass clusters, we restrict ourselves to clusters with $Q(\chi^2) > 10^{-3}$. Of the 6097 clusters with successfully measured radii, 5105 pass this further cut. In addition to low-mass clusters, we also find qualitatively that this cut removes many clusters with high masses ($> 10^5\,M_\odot$) and very small radii ($R_{\rm eff} < 1$ pc) that were outliers from the mass-radius relation due to their unreliable mass.

In addition, Krumholz et al. (2015) show that the deterministic LEGUS mass uncertainties are likely underestimated for low-mass clusters. To correct for this, we compare the mass uncertainty in Krumholz et al. (2015) to the uncertainty in the LEGUS catalogs. For clusters below 5000 $M_\odot$, the median difference in uncertainty is 0.16 dex. We add this 0.16 dex correction to the mass uncertainty of all clusters below 5000 $M_\odot$ when performing our fits.

This produces a sample across the full mass range that includes only the most reliable low-mass clusters and adjust their errors to account for the systematic error in the deterministic LEGUS SED fitting. Our resulting sample may not be complete (as we will be missing old, low-mass clusters), but we will discuss these selection effects throughout the rest of the paper.

We fit this mass-radius distribution assuming a power law relation:

$$\hat{R}_{\rm eff}(M) = R_4 \left(\frac{M}{10^4\,{\rm M}_\odot} \right)^\beta \tag{2.20}$$

such that the normalizing factor R_4 is the effective radius at $10^4\,M_\odot$. We incorporate errors in both mass and radius and account for the intrinsic scatter by minimizing projected displacements of data points from the fit line as outlined in Hogg et al. (2010). In Appendix A we describe this in more detail and compare it to a hierarchical Bayesian model that includes a treatment of selection effects. We found that neither the fitting method nor the inclusion of selection effects made a difference in the fit parameters, so we use this simpler method. We restrict the fitting to clusters below $10^5\,M_\odot$, as the relation appears to flatten above that mass, possibly because of small-number statistics. We determine errors on the fit via bootstrapping. For this full sample, our best fit power law slope is $\beta = 0.24$, with an intrinsic scatter of 0.25 dex. Restricting to clusters with ages younger than 1 Gyr produces the same result, primarily because most clusters in LEGUS are younger than 1 Gyr.

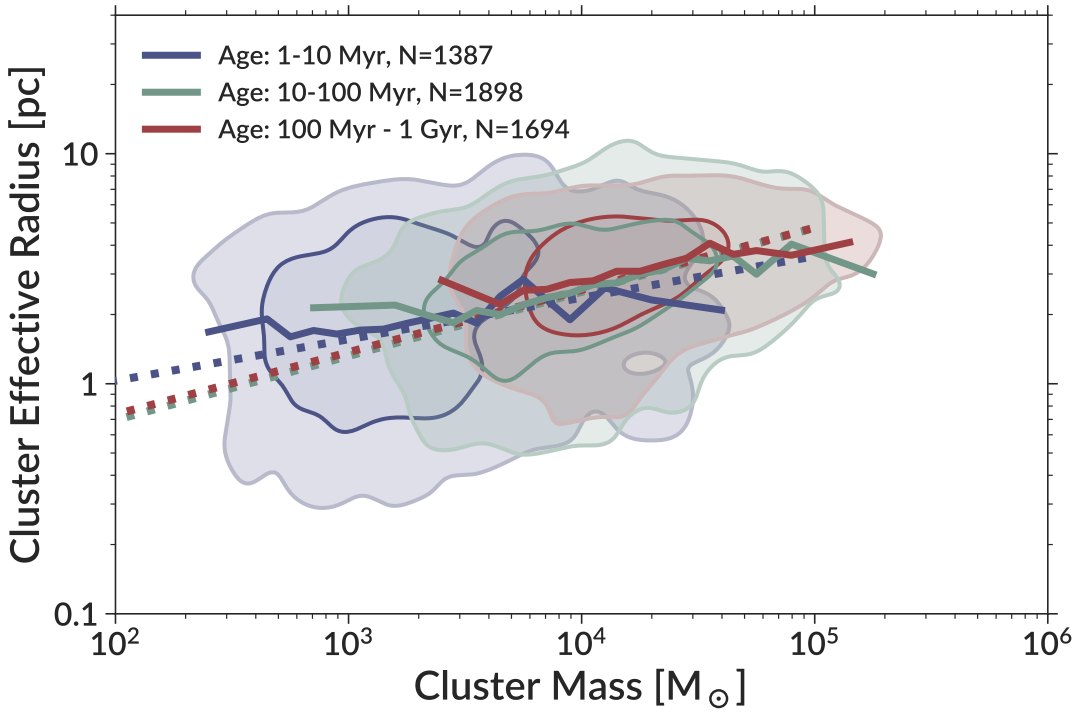

Figure 2.12 The mass-radius relation for the clusters in LEGUS, split by age. Contours enclose 50 and 90% of the clusters in each bin, and are smoothed by a kernel of 0.08 dex. The dashed lines show the fits, while the solid lines show the running median in each age bin.

With this large sample, we can investigate the cluster mass-radius relation for various subsamples of the data. For all cases below, we restrict our sample to clusters younger than 1 Gyr and less massive than $10^5\,M_\odot$. Table 2.2 shows the fitting parameters for all the subsets discussed below.

Table 2.2 Fits to the cluster mass-radius relation for various subsets of the data. All fits were done using clusters with masses below $10^5\,M_\odot$, while all fits other than the first row were done using clusters younger than 1 Gyr. Note that the fits including other datasets are sensitive to how these datasets are weighted, adding systematic uncertainties to the fit parameters.

Selection	N	β: Slope	R_4: $R_{\rm eff}$(pc) at $10^4\,M_\odot$	Intrinsic Scatter	$\log M$ percentiles: 1–99
Full LEGUS Sample	5105	0.242 ± 0.010	2.548 ± 0.022	0.250 ± 0.003	2.57 – 5.40
1 Myr – 1 Gyr	4979	0.246 ± 0.010	2.554 ± 0.022	0.250 ± 0.003	2.57 – 5.26
Age: 1–10 Myr	1387	0.180 ± 0.028	2.365 ± 0.106	0.319 ± 0.006	2.40 – 4.89
Age: 10–100 Myr	1898	0.279 ± 0.021	2.506 ± 0.035	0.238 ± 0.005	2.91 – 5.24
Age: 100 Myr – 1 Gyr	1694	0.271 ± 0.027	2.558 ± 0.048	0.198 ± 0.005	3.46 – 5.40
LEGUS + MW	6158	0.296 ± 0.002	2.555 ± 0.022	0.225 ± 0.003	0.93 – 5.22
LEGUS + External Galaxies	5874	0.229 ± 0.008	2.561 ± 0.020	0.244 ± 0.003	2.36 – 5.31
LEGUS + MW + External Galaxies	7053	0.292 ± 0.002	2.567 ± 0.020	0.222 ± 0.003	0.97 – 5.26

Figure 2.12 shows the mass-radius relation split by cluster age. The mass range of the three bins is clearly different, due to LEGUS's absolute V band magnitude cut. Evolutionary fading results in only massive clusters being detected at older ages. To demonstrate this, we include the mass range spanned by the 1–99th percentile of each sample in Table 2.2. In all 3 bins, we detect a significantly nonzero slope of the mass-radius relation. The running medians of each panel are quite similar, especially for the two oldest age bins, which deviate from the 1–10 Myr bin at $M > 10^4\,\mathrm{M}_\odot$. This matches what we find in the formal fit, where the two older bins have a slope and normalization indistinguishable from each other, while the 1–10 Myr bin has a significantly shallower slope. In addition, the intrinsic scatter decreases with age.

Additionally, we supplement the LEGUS sample with several other large samples for young star clusters, mostly from the compilation of Krumholz et al. (2019). In all the samples below, we restrict to clusters with an age less than 1 Gyr and masses below $10^5\,\mathrm{M}_\odot$, as done for our main fits. This age cut means that we do not include any globular clusters. Additionally, some of the samples in Krumholz et al. (2019) are for galaxies already included in this paper (namely NGC 628, NGC 1313, and NGC 5194), so we do not include them again here.

We include Milky Way open clusters (OCs) within 2 kpc of the Sun from Kharchenko et al. (2013), who measured King parameters for these clusters. Following Krumholz et al. (2019), we calculate mass using Equation 3 of Piskunov et al. (2007), Oort constants from Bovy (2017), and the distance from the Sun to the Galactic center from Bland-Hawthorn & Gerhard (2016). We also include the sample of 12 Milky Way YMCs compiled in Krumholz et al. (2019).

We additionally include samples from several external galaxies. Mackey & Gilmore (2003b,a) measured radii for 53 clusters in the LMC and 10 clusters in the SMC. EFF profiles were fit to the surface brightness profiles of clusters. These surface brightness profiles were also used to obtain the total luminosity of each clusters, which was

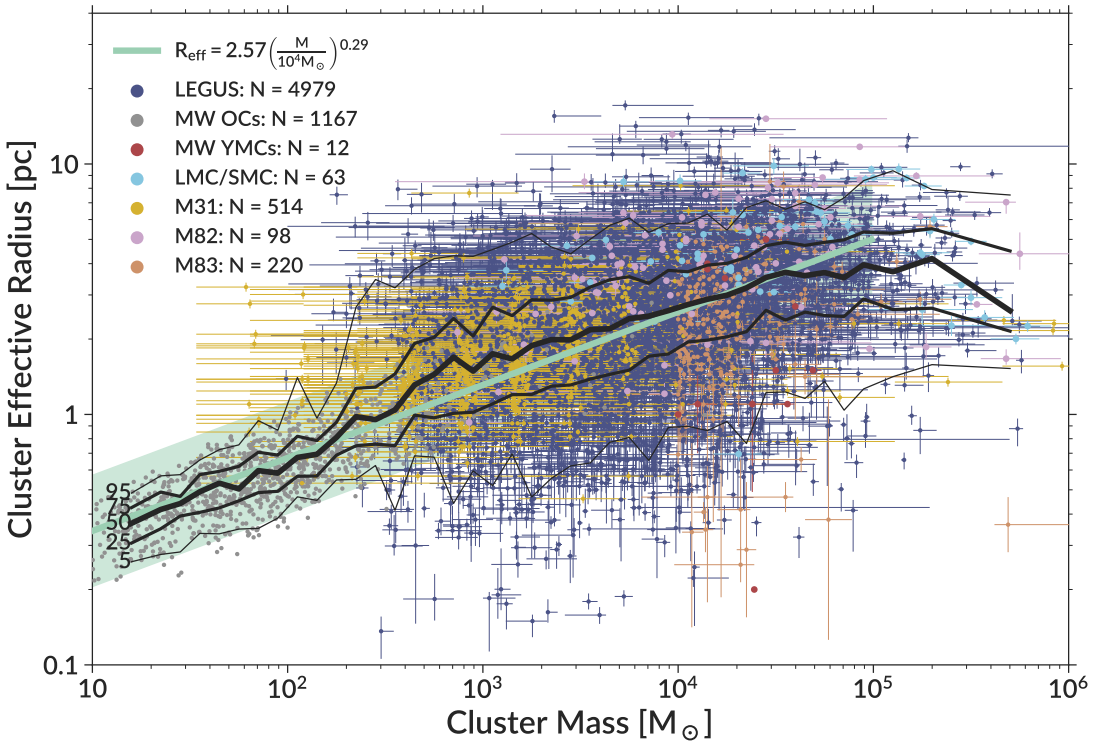

Figure 2.13 The mass-radius relation for the clusters in LEGUS as well as external datasets, as described in the text. Only clusters with ages less than 1 Gyr are shown here. The black lines show then 5, 25, 50, 75, and 95th percentiles of radii at a given mass. Note that a small random shift has been applied to the original discrete M31 masses for visual purposes.

converted into the cluster mass by using mass-to-light ratios.

The Panchromatic Hubble Andromeda Treasury (PHAT) survey identified stellar clusters in M31 (Johnson et al., 2012; Fouesneau et al., 2014). Half-light radii were determined by interpolating the flux profile to find the radius in arcseconds that includes half of the light (Johnson et al., 2012). We then use a distance of 731 kpc to M31 (Wagner-Kaiser et al., 2015) to convert the radii to parsecs. Masses were determined using a Bayesian SED fitting method that explicitly accounts for the stochastic sampling of the IMF (Fouesneau et al., 2014).

In a series of papers, Cuevas-Otahola et al. (2020) and Cuevas-Otahola et al. (2021) calculated structural parameters for 99 star clusters in M82. They fit EFF, King, and Wilson profiles to the surface brightness profiles, finding that the EFF profile best represents the clusters in their sample. Similarly to Mackey & Gilmore (2003b,a), masses are determined by applying a mass-to-light ratio to fitted luminosities.

We also include the clusters in M83 measured by Ryon et al. (2015). Radii are measured by fitting an EFF profile to the 2D light profile using GALFIT, as in R17. Masses are derived in Silva-Villa et al. (2014) and are done using an SED-fitting method similar to that used in LEGUS (Adamo et al., 2010).

Figure 2.13 shows the mass-radius relation including these data sources. In the fit shown in this figure, we give each cluster equal weight, no matter which dataset it comes from. The addition of the MW OCs in particular extends the mass-radius relation down to very small masses, while the other samples overlap nicely with LEGUS. Including the MW clusters produces a steeper slope than the fit using only LEGUS clusters ($\beta = 0.296$). In addition, the intrinsic scatter decreases, likely due to the smaller scatter in the MW OC data. Including the data from external galaxies produces a shallower slope ($\beta = 0.229$), likely due to fewer low-mass clusters with small radii in the M31 PHAT sample. Including all data produces a fit similar to the

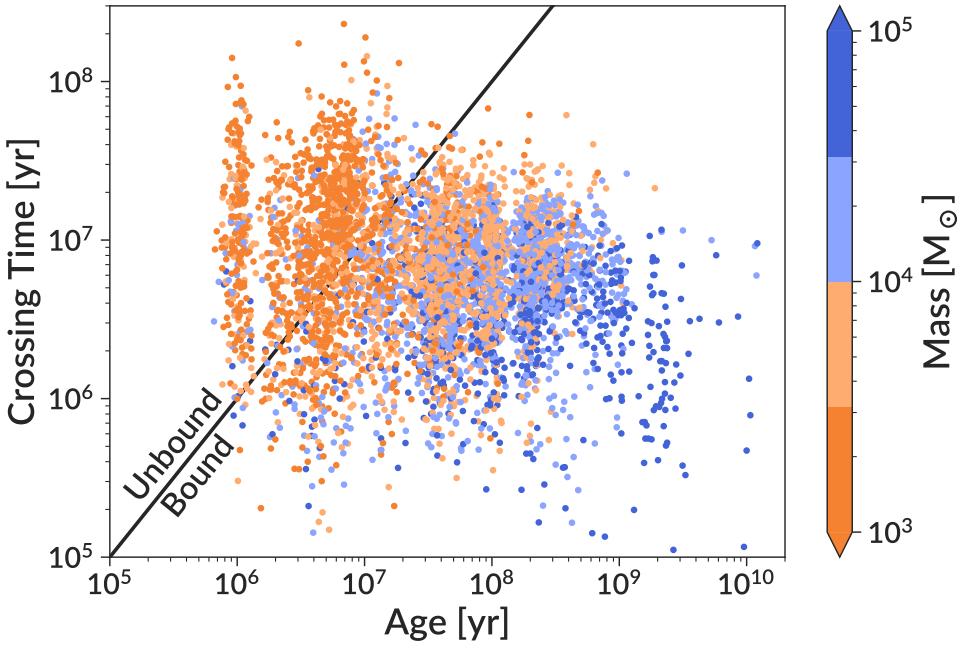

Figure 2.14 A comparison of the crossing time to the age of clusters in LEGUS. Clusters are color coded according to their mass, and a small random offset was added to the discrete ages for visual purposes. The black line shows where these times are equal. Clusters where the age is longer than the crossing time are likely bound, while those where the age is smaller are likely unbound.

fit for LEGUS + MW, likely due to the leverage and large numbers provided by the low-mass clusters in the MW sample. In all cases, R_4 is consistent with that measured in the LEGUS sample alone. Note that due to the asymmetric shape of the cluster radius distributions, R_4 may be different from the peak value of the distribution quoted in Section 2.3.2.

2.3.4 Are Clusters Gravitationally Bound?

Gieles & Portegies Zwart (2011) suggest a distinction between bound clusters and unbound associations, where bound clusters are older than their instantaneous crossing time:

$$t_{\rm cross} = 10 \left(\frac{R_{\rm eff}^3}{GM} \right)^{1/2} \tag{2.21}$$

Clusters with $t_{\rm age} > t_{\rm cross}$ have remained together for their lifetimes, indicating that

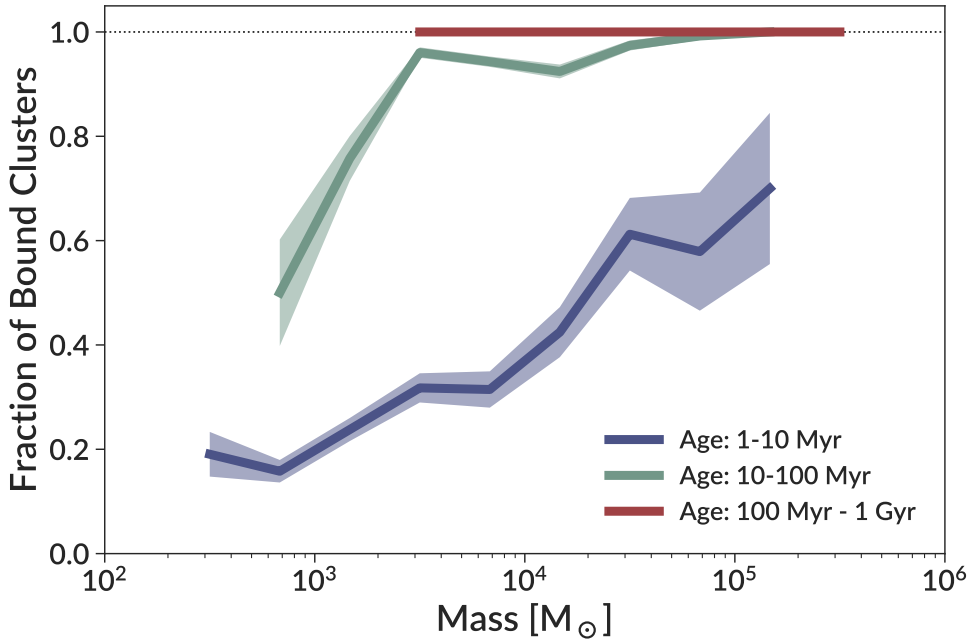

Figure 2.15 The fraction of clusters that are older than their crossing times (indicating that they are bound objects) as a function of mass. Shaded regions show the 68% confidence region.

they are gravitationally bound. Unbound objects expand with time, causing the crossing time to increase proportionally.

Figure 2.14 shows the comparison of these timescales for the clusters in the LEGUS sample. The majority of clusters are bound. The only unbound clusters are young (less than 10 Myr) and tend to be less massive. At a given age, the less massive clusters are more likely to be unbound. We find that 78% of all clusters, 92% of clusters with $M > 5000\,\mathrm{M}_\odot$, and 97% of clusters older than 10 Myr are bound.

Figure 2.15 shows how the fraction of clusters that are bound changes with age and mass. 100% of objects older than 100 Myr are bound, while in the other age bins a clear trend with mass is seen. For the 10–100 Myr bin, nearly all clusters above $3000\,\mathrm{M}_\odot$ are bound, while the youngest clusters show a steadily increasing fraction of bound clusters with mass. This confirms that the LEGUS pipeline selects gravitationally bound objects, especially for clusters with higher masses or older ages.

2.4 Discussion

2.4.1 Selection Effects

When selecting clusters, the LEGUS survey used an absolute magnitude cut, selecting clusters with a V band magnitude brighter than -6 mag. As massive stars in clusters die, the cluster fades. This means that older clusters must be more massive to be detected, producing a significant selection effect in the LEGUS sample.

This is particularly visible in Figure 2.12. In the oldest age bin, there are nearly no clusters below a few $10^3\,M_\odot$, where the bulk of the youngest clusters are. The mass ranges seen in Figure 2.12 should not be simply interpreted as evidence for disruption of low-mass clusters, as old low-mass clusters would not be detected even if they existed. This also complicates an examination of cluster evolution, as without old low-mass clusters to compare, it is difficult to test predictions of low-mass cluster evolution.

In the results above and discussion that follows, we present results using all LEGUS clusters. We want to be clear that this is not a complete sample. Where relevant, we discuss how these selection effects may bias the results presented.

LEGUS also uses a cut in concentration index (with a value that varies for each galaxy). This cut may result in the removal of the smallest clusters. This will depend on galaxy distance, as the smallest clusters will be possible to resolve in nearby galaxies but not distant ones. As mentioned above in Section 2.3.2, this is not likely to affect many of our galaxies. Adamo et al. (2017) examined the completeness as a function of radius for NGC 628 at a distance of 8.8 Mpc, finding that LEGUS includes roughly 50% of clusters with $R_{\rm eff} = 1$ pc. As most of our galaxies are closer than NGC 628, they will be less affected. The radius distribution shows a clear peak significantly above the radius where we may be incomplete (Figure 2.10), showing that the potential removal of small clusters likely will not dramatically change our

results.

We also note that because of the inability of our pipeline to pick up extremely small objects (smaller than 0.3 pixels; see Figure 2.8), the smallest objects may be even more compact than reported.

The inclusion of small, low-mass clusters would have the effect of steepening the mass-radius relation. Interestingly, a comparison of the slopes in Figures 2.11 and 2.13 shows that a steeper slope better matches the MW OCs. If the true mass-radius relation is steeper than we measure here, it may make our results more consistent with the measurements in the MW.

Lastly, we note that the radial coverage of each galaxy varies. About half of the LEGUS galaxies are compact enough that they can be covered with one HST WFC3/UVIS pointing, but larger galaxies are not completely covered. Some galaxies include only central regions (e.g. NGC 1566), while others include the central regions and part of the disk (e.g. NGC 628). See Figure 3 of Calzetti et al. (2015) for the full footprints for all LEGUS galaxies. This uneven coverage may bias our results somewhat if cluster populations vary throughout galaxies. This could happen if they are tidally bound, as the tidal radius would change with galactocentric radius. We defer a detailed examination of this for a future paper.

2.4.2 Mass-Radius Relation

The mass-radius relation shown in Figure 2.11 has a relatively shallow slope and significant intrinsic scatter. Nevertheless, a relation is clearly present, even when splitting by age (Figure 2.12).

In the Milky Way, observations of giant molecular clouds (GMCs) show roughly constant surface densities (Larson, 1981). From this we can expect a mass-radius relation $R = (M/\pi\Sigma)^{1/2} \propto M^{1/2}$. Measurements of clumps have found a range of slopes from 0.3 to 0.6 (Roman-Duval et al., 2010; Urquhart et al., 2018; Mok et al.,

2021). These relations are steeper than the relation we measure for young clusters, which presumably form from such clouds. However, we note that the hierarchical structure of the ISM makes determining the size of a clump more challenging than measuring the radius of a star cluster, so these radii might not be directly comparable (Colombo et al., 2015). We examine cluster density further in Section 2.4.3. The analytic model of Choksi & Kruijssen (2021) also predicts a mass-radius relation of the form $R_{\rm eff} \propto (M/\Sigma_g)^{1/2}$, where Σ_g is the gas surface density. After accounting for the fact that massive clusters can only form in high-density environments, they find a lower slope which is more consistent with this work.

Many previous studies have found inconclusive evidence of a correlation between the mass and radius of young clusters. Zepf et al. (1999) found a shallow mass-luminosity relation, with later studies finding little evidence of a strong mass-radius relation (Bastian et al., 2005; Scheepmaker et al., 2007; Bastian et al., 2012; Ryon et al., 2015, 2017). Some studies have found a mass-radius relation for the most massive clusters above $10^6\,{\rm M}_\odot$ (Kissler-Patig et al., 2006; Bastian et al., 2013), but that mass range is not sampled in our results. The large sample size, uniform LEGUS selection, and uniform fitting procedure presented here allow us to clearly detect this relation. Our result is similar to that of Cuevas-Otahola et al. (2021), who find a power law slope of 0.29. However, the normalization of their relation is higher than ours (see Figure 2.13).

Interestingly, we also find less evolution in the cluster radius with age than seen in other studies. Bastian et al. (2012) fit the cluster radius distribution as a bivariate function of age and mass, finding that age is the stronger driver of cluster radius than mass. Ryon et al. (2015) found a significant age-radius relation. They also found a steepening of the mass-radius relation with time, although their age bins are different from the bins used in this paper. Chandar et al. (2016) found that clusters with age of 100–400 Myr have radii 4 times larger than clusters of similar mass with ages younger

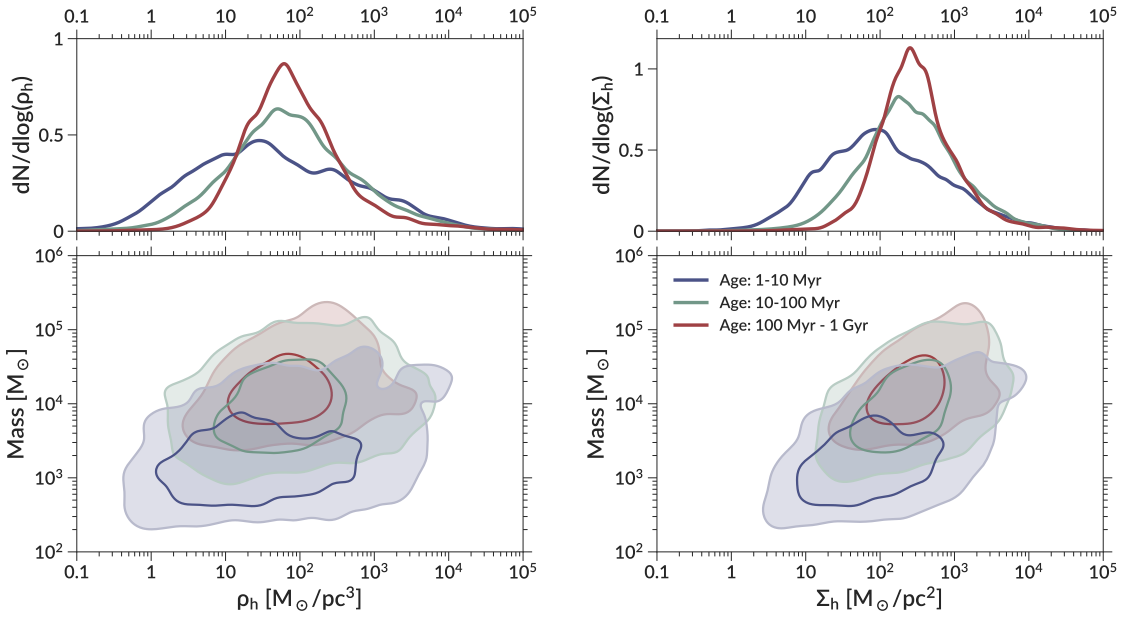

Figure 2.16 Density ρ_h (left panels) and surface density Σ_h (right panels) of clusters within the half-light radius. The top panels shows kernel density estimation of the density distributions, where clusters are smoothed by a Gaussian kernel with a width equal to their measurement error. Each curve is normalized to integrate to the same area. The bottom panels show the dependence of densities on mass. The contours enclose 50 and 90% of the data, and are smoothed by 0.15 dex. In all panels, we split the sample by age.

than 10 Myr. Our results stand in contrast, as we find no significant evolution after 10 Myr in the mass ranges probed.

2.4.3 Density Distribution

Using our measured radii, we calculate the average density and surface density of the LEGUS clusters:

$$\rho_h = \frac{3M}{8\pi R_{\rm eff}^3}, \quad \Sigma_h = \frac{M}{2\pi R_{\rm eff}^2}. \tag{2.22}$$

In this section we will use ρ_h and Σ_h when referring to those quantities respectively, and use the generic term "density" when referring to both of them. In Figure 2.16, the top panels show the distributions of these densities split by cluster age. Younger clusters have wider ranges and extend to lower densities than old clusters. The bottom panels show the distribution of densities as a function of mass.

We find that the distributions shown in the top panels of Figure 2.16 are well described by lognormal functions:

$$\frac{dN}{d\log\rho_h} = \frac{1}{\sqrt{2\pi\sigma_\rho^2}} \exp\left[-\frac{(\log\rho_h - \log\mu_\rho)^2}{2\,\sigma_\rho^2}\right] \tag{2.23}$$

and the equivalent for Σ_h, where μ_ρ and σ_ρ^2 are the mean and variance, respectively. We fit these distributions, and show the parameters in Table 2.3. We also include a fit to the entire distribution without splitting by age.

The decrease in the number of low-density clusters with age is likely to be a combination of selection effects and disruption of low-mass clusters. There appears to be a weak mass-ρ_h relation, and a much stronger relation between mass and Σ_h. As old clusters are massive, they are more likely to have high Σ_h. However, in the narrow mass range where the age distributions overlap (around $5 \times 10^4\,\mathrm{M}_\odot$), the youngest age bin extends to lower density than the older age bins. We also examined the density distributions for clusters in the mass range shared by all age bins, again finding a larger spread at younger ages. This may indicate that disruption is responsible for removing these low-density clusters. At the same time, we cannot rule out that higher mass clusters simply form at higher density.

Observations of GMCs in nearby quiescent galaxies are consistent with a roughly constant surface density $\Sigma_{\mathrm{GMC}} \sim 100\ \mathrm{M}_\odot\,\mathrm{pc}^{-2}$, while in starbursting and high-redshift galaxies the normalization is higher $\Sigma_{\mathrm{GMC}} \sim 2000\ \mathrm{M}_\odot\,\mathrm{pc}^{-2}$ (e.g., Dessauges-Zavadsky et al., 2019). Clumps within resolved Galactic clouds are also consistent with a nearly constant surface density $\Sigma_{\mathrm{GMC}} \gtrsim 1000\ \mathrm{M}_\odot\,\mathrm{pc}^{-2}$ (Urquhart et al., 2018), although fixed column density may be a selection effect.

The density of the LEGUS clusters is similar to that of GMCs in nearby quiescent galaxies. The peak of the Σ_h distribution for young clusters is $\approx 100\ \mathrm{M}_\odot\,\mathrm{pc}^{-2}$. However, there is a wide range in cluster densities, in contrast to the narrower range of

$\Sigma_{\rm GMC}$. This may be due to the fact that there is not a direct connection between the density of GMCs and clusters. Clusters form out of the densest clumps within GMCs, which in the Milky Way typically have $\Sigma_{\rm GMC}$ between 100 and a few 10^4 $\rm M_\odot\,pc^{-2}$ (Urquhart et al., 2018). After stars form out of these clumps, stellar feedback disperses the gas. This causes the cluster to increase in size and decrease in density. We note that we are measuring the radii of clusters at this phase of their evolution, after gas expulsion.

We also note that the LEGUS sample is from many galaxies with a range of star formation rates. This may produce a range of GMC properties that are partially responsible for explaining the scatter in cluster density (Sun et al., 2018). In future work we will examine the dependence of cluster properties on their environment.

Taking full density distributions from the bottom panels of Figure 2.16, we fit power-law relations and obtain $\rho_h \propto M^{0.52\pm0.02}$ and $\Sigma_h \propto M^{0.67\pm0.012}$. The fitted intrinsic scatter in ρ_h is 1.12 ± 0.014 dex, while for Σ_h it is 0.74 ± 0.009 dex. Errors are determined by bootstrapping. As a consistency check, we can compare these fit slopes with those expected based on mass-radius relation fit. For the full LEGUS sample, $R_{\rm eff} \propto M^{0.242}$. By rewriting that relation in terms of densities, the expected relations are $\rho_h \propto M^{0.274}$ and $\Sigma_h \propto M^{0.516}$. These are less steep than the direct fits, especially for ρ_h. Any discrepancy in the slopes may be due to the very large intrinsic scatter in densities. Their dynamic range is larger than the dynamic range in mass. This scatter, along with the lack of a clear relation (particularly for the mass-ρ_h relation), makes it difficult to fit a reliable slope.

2.4.4 Cluster Evolution

Previous literature indicated that clusters may expand with time (e.g. Bastian et al., 2012; Ryon et al., 2017). As stars within clusters lose mass, isolated clusters will slowly expand to maintain virial equilibrium, while at later times two-body relaxation

Table 2.3 Lognormal fits to the density distributions shown in Figure 2.16. The log mean μ and standard deviation σ are given for density ρ_h and surface density Σ_h. Note that the cluster mass ranges given in Table 2.2 apply to these fits too.

Age	$\log \mu_\rho$ ($M_\odot\,pc^{-3}$)	σ_ρ (dex)	$\log \mu_\Sigma$ ($M_\odot\,pc^{-2}$)	σ_Σ (dex)
All	1.80	0.78	2.36	0.61
1–10 Myr	1.56	1.13	1.98	0.82
10–100 Myr	1.83	0.78	2.37	0.60
100 Myr – 1 Gyr	1.82	0.57	2.44	0.43

can also increase the cluster radius (Gieles et al., 2010).

In Figure 2.12 we see a statistically significant evolution in radius with age from the 1–10 Myr bin to the 10–100 Myr bin, with high-mass clusters slightly expanding. However, the magnitude of this increase is quite small. The fit parameters indicate typical clusters at $10^4\,M_\odot$ expand from 2.36 to 2.51 pc, while clusters at $10^5\,M_\odot$ expand from 3.58 to 4.76 pc. Notably, we only see this evolution between our first two age bins. We see no significant differences between the 10–100 Myr and 100 Myr–1 Gyr bins.

To start interpreting these results, we turn to an examination of the Jacobi radius, which sets the radius at which stars belong to a cluster when it is in a tidal field. Clusters that fill a larger fraction of the Jacobi radius are more vulnerable to mass loss. For clusters with mass M in circular orbits with angular frequency ω in a galaxy with a flat rotation curve, the Jacobi radius is defined as

$$r_J = \left(\frac{GM}{2\omega^2}\right)^{1/3} \tag{2.24}$$

We do not directly calculate r_J, as obtaining ω for each cluster is beyond the scope of this paper, and the assumption of a flat rotation curve may not be true for every galaxy. However, we can qualitatively examine how the ratio of the effective radius

to the Jacobi radius scales with cluster mass:

$$\frac{R_{\rm eff}}{r_J} \propto \frac{M^{\beta}}{M^{1/3}} \tag{2.25}$$

where β is the slope of the mass-radius relation (Equation 2.20 and Table 2.2). Note that this assumes no relation between M and ω. For the full sample $\beta = 0.24$, giving $R_{\rm eff}/r_J \propto M^{-0.09}$. In the youngest age bin (1–10 Myr) $\beta = 0.18$, and $R_{\rm eff}/r_J \propto M^{-0.15}$. In either case, high-mass clusters fill less of their Jacobi radii than low-mass clusters.

To examine the evolution of clusters with time, we also use the evolution model from Gieles & Renaud (2016), known hereafter as GR16. This model includes two processes that influence cluster evolution: tidal shocks and two-body relaxation. Tidal shocks increase the energy of the cluster and result in mass loss, while two-body relaxation is assumed to only increase the energy of the cluster without causing mass loss. To show how this model would change the clusters in the sample, we take clusters along the 1–10 Myr best fit relation and evolve them through this model for 300 Myr. Figure 2.17 shows contours of the observed distribution of clusters in the three age bins, their best fit relations, and arrows illustrating how the GR16 model affects clusters in this 300 Myr of evolution.

Figure 2.17 also includes two other toy models of cluster evolution, designed to represent the tidally limited and not tidally limited extremes. When clusters are very small compared to Jacobi radii, any energy injection will increase the radius without causing substantial mass loss. On the other hand, any energy from either two-body relaxation or shocks will cause mass loss in tidally limited clusters, and the cluster's effective radius will decrease along with the tidal radius.

To model these two cases we modify the mass loss prescriptions of GR16. They

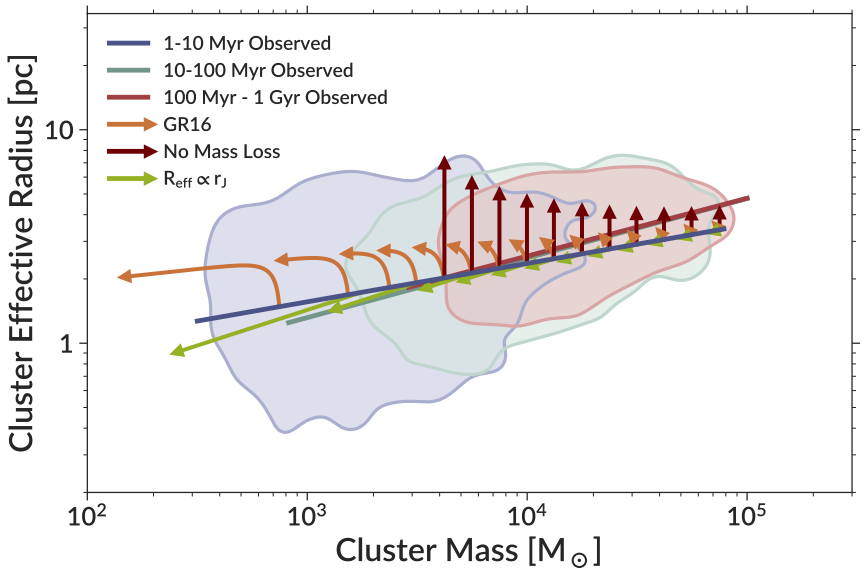

Figure 2.17 A comparison of 3 models of cluster evolution. Contours enclose 75% of the clusters in each age bin, and are smoothed by 0.08 dex. The corresponding solid lines show the fit for these age bins from Figure 2.12, but are restricted to the mass range spanned by the 1–99th percentile of masses in this age bin from Table 2.2. The sets of arrows show how 3 models for cluster evolution as described in Section 2.4.4 would change clusters lying on the 1–10 Myr relation after 300 Myr of evolution.

introduce a parameter f that relates mass loss to energy injection (their Equation 2):

$$\frac{dM}{M} = f\,\frac{dE}{E} \tag{2.26}$$

For shocks, they set $f_{\rm sh} = 3$, while setting $f_{\rm rlx} = 0$ turns off mass loss from two-body relaxation.

We change these f values to produce our two cases. In the case where clusters are not tidally limited and mass loss does not happen, we set both $f_{\rm sh} = f_{\rm rlx} = 0$. We then rederive the model, and its results are shown by the "No Mass Loss" arrows in Figure 2.17. In the tidally limited case, we keep the original $f_{\rm sh} = 3$ but also allow for mass loss from two-body relaxation with $f_{\rm rlx} = 0.2$ (Gieles et al., 2006). As the

cluster loses mass, we require the radius to be proportional to the tidal radius:

$$\frac{R_{\rm eff}}{R_{\rm eff,0}} = \frac{r_J}{r_{J,0}} = \left(\frac{M}{M_0}\right)^{1/3} \tag{2.27}$$

The scaling relation presented above in Equation 2.25 indicates that low-mass clusters are more likely to be tidally limited than high-mass clusters, and Figure 2.17 qualitatively supports this conclusion. If high-mass clusters are not tidally limited, they will lose little mass and expand, matching the observations. While the mass range of old clusters prohibits a detailed examination of low-mass clusters, the mass-radius relation would steepen if the effective radius of low-mass clusters evolves proportional to the tidal radius.

However, it is clear that none of these models do a good job of quantitatively matching the full evolution. The GR16 model pushes clusters towards a mean relation of $R_{\rm eff} \propto M^{1/9}$, making the relation shallower rather than steeper as required by the observations. The toy model with no mass loss can increase the radius of massive clusters, but its effects are weakest for the highest mass clusters where the observations show the largest difference with age. The toy model that assumes the radius changes with the tidal radius may work for low-mass clusters, but for high-mass clusters it has nearly no effect.

Importantly, the time dependence of these models is in strong conflict with the observations. We see a change from the 1–10 Myr age bin to the 10–100 Myr bin, with no significant change afterwards. However, the models change clusters steadily with time, leading to little change in the first ∼30 Myr but large changes after 300 Myr.

In addition, models would need to match the change in scatter with time. As clusters evolve towards the mean relation of the GR16 model, the scatter decreases dramatically. We tested this by taking the full 1–10 Myr sample and putting each cluster through the GR16 model for 300 Myr. At this late time, the distribution of

cluster radii has a much smaller scatter than seen in the observations of clusters at late times. While the observed scatter does decrease with time, it decreases less than this model predicts.

We note that stellar mass loss is not included in the GR16 model and the modified versions presented here. In 1 Gyr, clusters can lose roughly 30% of their mass through stellar evolution alone, and this can cause them to expand (Gieles et al., 2010). In addition, one should be careful comparing the cluster mass from the models to the observed mass. The models treat mass loss from stars leaving the cluster as instantaneous, while in reality stars can remain in clusters for a long time before escaping through the Lagrangian points (Fukushige & Heggie, 2000; Baumgardt, 2001; Claydon et al., 2019). Once stars escape, their low velocity dispersion means that unbound stars can remain near the cluster (Küpper et al., 2008, 2012; Webb et al., 2013). These unbound but nearby stars may still be included in the SED fit and radius fits. The observed mass of the cluster is therefore not necessarily the bound mass of the cluster, which is what the models present.

From all of this, it is clear that more work needs to be done to understand how cluster evolution models can be used to interpret this data set. The scaling relation presented above in Equation 2.25 indicates that high-mass clusters may be less likely to be tidally bound, but beyond that we refrain from drawing definitive conclusions about cluster evolution.

2.5 Conclusions

We implemented a custom pipeline to measure the projected half-light radius $R_{\rm eff}$ of star clusters. This pipeline has several features designed to make it robust to contamination and accurately estimate the local background, producing reliable values of $R_{\rm eff}$ (Section 2.2). We applied this pipeline to the star clusters in 31 LEGUS galaxies, producing a uniformly-measured sample of 7242 star cluster radii. Of these,

we identify 6097 as having reliable radii. This is currently the largest such catalog of star cluster radii available.

We summarize the key results below:

• Most, but not all, galaxies share a common cluster radius distribution, with a peak at around 3 pc (Figure 2.10). The shape of this distribution is asymmetric, with a tail to small radii, and is well described by a Weibull distribution (Equation 2.19).

• We find a clear but shallow mass-radius relation (Figures 2.11, 2.13). This relation takes the form $R_{\rm eff} \propto M^{0.24}$, with an intrinsic scatter of 0.25 dex (Table 2.2).

• This mass-radius relation is present in clusters of all ages probed by LEGUS sample (Figure 2.12). The slope of this relation is shallower at 1–10 Myr than at later times, but the slope does not evolve between the 10–100 Myr and 100 Myr–1 Gyr bins. The intrinsic scatter decreases with time. Selection effects cause the subsamples of different age to span different mass ranges, complicating interpretation (Section 2.4.1).

• The majority of clusters identified in LEGUS are gravitationally bound (Figure 2.14). The majority of unbound clusters are younger than 10 Myr and tend to be less massive (Figure 2.15).

• The distributions of both average density and surface density of LEGUS clusters are well fit by lognormal distributions (Figure 2.16). The width of these distributions is large but decreases with cluster age. The peaks of the distributions for the youngest clusters are $\rho_h \approx 30$ M$_\odot$ pc^{-3} and $\Sigma_h \approx 100$ M$_\odot$ pc^{-2} (Table 2.3).

• While we do not directly calculate the Jacobi radii for the LEGUS clusters, the shallow mass-radius relation implies that high-mass clusters fill less of their Jacobi radius than low-mass clusters (Equation 2.25).

• We create simple toy models of cluster evolution based on the model in Gieles & Renaud (2016) to interpret the trends we see with cluster age (Figure 2.17). None of the models can successfully reproduce all aspects of the observed distributions.

CHAPTER III

Nuclear Star Clusters in Cosmological Simulations

This chapter was published as: ***Brown, G.****, Gnedin, O. Y., & Li, H. 2018, ApJ, 864, 94*

Abstract

We investigate the possible connection between the most massive globular clusters, such as ω Cen and M54, and nuclear star clusters of dwarf galaxies that exhibit similar spreads in age and metallicity. We examine galactic nuclei in cosmological galaxy formation simulations at $z \approx 1.5$ to explore whether their age and metallicity spreads could explain these massive globular clusters. We derive structural properties of these nuclear regions, including mass, size, rotation, and shape. By using theoretical supernova yields to model the supernova enrichment in the simulations, we obtain individual elemental abundances for Fe, O, Na, Mg, and Al. Our nuclei are systematically more metal-rich than their host galaxies, which lie on the expected mass-metallicity relation. Some nuclei have a spread in Fe and age comparable to the massive globular clusters of the Milky Way, lending support to the hypothesis that nuclear star clusters of dwarf galaxies could be the progenitors of these objects. None of our nuclear regions contain the light element abundance spreads that characterize globular clusters, even when a large age spread is present. Our results demonstrate

that extended star formation history within clusters, with metal pollution provided solely by supernova ejecta, is capable of replicating the metallicity spreads of massive globular clusters, but still requires another polluter to produce the light element variations.

3.1 Introduction

Globular clusters have traditionally been seen as having a single-age, single-metallicity stellar population. However, recent evidence has shown that the formation process of globular clusters is more complicated (e.g. Gratton et al., 2012). Color-magnitude diagrams show multiple populations of stars, as evidenced by multiple main sequences, the widening of the red giant branch, and the splitting of the horizontal branch (e.g. Milone et al., 2017). Additionally, spectroscopy has revealed correlations in the abundances of light elements such as N, O, Na, Mg, and Al (e.g. Carretta et al., 2009a). A large fraction of globular cluster stars are affected.

Further deviations from a simple stellar population are present in a few of the most massive clusters, which can have spreads in both metallicity and age. The measurable spread in iron gives these objects the name "anomalous" or "iron-complex" clusters (Marino et al., 2015). They include ω Cen (Johnson et al., 2009), M54 Carretta et al. (2010a), and Terzan 5 (Massari et al., 2014). In clusters with an iron spread, the light element abundances are present within each peak of the metallicity distribution (Marino et al., 2011, 2015). An age spread may be present in these clusters as well. For ω Cen the age spread has several claims in the literature, from around 500 Myr (Tailo et al., 2016) to 1 Gyr (Joo & Lee, 2013) to several Gyr (Villanova et al., 2014). Terzan 5 has a much larger age spread, at around 7.5 Gyr (Ferraro et al., 2016). M54 is still embedded in the stellar population of its host galaxy, making interpretation difficult, but this region has experienced star formation over at least 10 Gyr (Siegel et al., 2007). The extended formation history of these objects may

imply a significantly different origin than typical globular clusters. These anomalous clusters are more similar to the nuclei of dwarf galaxies, and thus may represent transitional objects between star clusters and galaxies.

Most nearby galaxies contain nuclear star clusters (NSCs) (e.g. Leigh et al., 2012). NSCs are more prevalent in smaller galaxies ($M_\star < 10^{10}\,\mathrm{M}_\odot$), but can coexist with supermassive black holes in the mass range around $10^{10}\,\mathrm{M}_\odot$ (Ferrarese et al., 2006; Seth et al., 2010; Georgiev et al., 2016). The Milky Way hosts both a NSC and a central black hole (e.g. Feldmeier-Krause et al., 2017b). Since NSCs sit at the bottom of the galaxy's potential well, they can have inflows of gas that allow multiple star formation episodes, compared to non-nuclear star clusters of the same mass. The Milky Way's NSC shows evidence of a complicated star formation history (Do et al., 2015; Feldmeier-Krause et al., 2017a). Enriched gas processed by the stars in the dense cluster is likely to remain near the cluster, allowing the later populations of stars to have enriched elemental abundances. Of particular interest are the NSCs of satellite galaxies. When a satellite galaxy merges with a larger host, its dense nucleus is likely to survive the tidal interactions. After the galaxy is stripped, the NSC may emerge as one of the host's massive globular clusters. This hypothesis is potentially capable of explaining the age and metallicity spreads of the most massive globular clusters (Freeman, 1993; Böker, 2008). These spreads in metallicity may also result in spreads in light elements, potentially explaining some of the observed spreads in massive clusters.

For this to be a viable scenario, the NSCs of the satellites need to have similar sizes, masses, and elemental abundances as the globular clusters we observe. Here we test this scenario by analyzing the results of recent ultrahigh-resolution simulations of galaxy formation by Li et al. (2017, 2018). These simulations include a novel star formation prescription, where star clusters are the unit of star formation. These star cluster particles form over time until the feedback from their stars termi-

nates their growth. The clusters incorporate continuous enrichment of the interstellar medium and encode some intrinsic metallicity spread. This realistic modeling results in self-consistent cluster masses, age spreads, and metallicity distributions, making the simulations an excellent tool to study this problem.

In this paper we identify nuclear regions in the simulations, then compare their masses and sizes to a sample of observed NSCs (Georgiev et al., 2016). We check the mass-metallicity relation for our galaxies and nuclear regions against the observed relation for dwarf galaxies (Kirby et al., 2013). We use published supernovae yields (Nomoto et al., 2006) to make predictions of the individual elemental abundances for the stars in the nuclear regions. We then examine the age and metallicity spread of the nuclear regions, as a direct comparison to the observed massive globular clusters. Lastly, we compare the spreads in light elements to the observed spreads found in globular clusters.

3.2 Structure of Nuclear Clusters in Simulations

The suite of cosmological simulations by Li et al. (2017, 2018) was run using the Adaptive Refinement Tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999, 2003; Rudd et al., 2008) in a 4 Mpc comoving box, with initial conditions selected to produce one Milky Way-sized galaxy at $z = 0$ along with several satellite galaxies. The ART code solves equations for the gravitational dynamics of the stars, dark matter, and gas, as well as the hydrodynamics of the gas component. The code utilizes Adaptive Mesh Refinement (AMR) to reach very high spatial resolution. The current suite has a maximum resolution of $L_{\rm cell} = 3-6$ physical pc (not comoving), which is high enough to resolve giant molecular clouds. The code calculates run-time transport of UV radiation (Gnedin & Abel, 2001) from both local stellar sources (Gnedin, 2014) and the extragalactic background (Haardt & Madau, 2001). A non-equilibrium chemical network is used to model the various ionization states of hydrogen and helium, and the

formation and destruction of molecular hydrogen (Gnedin & Kravtsov, 2011). Stellar feedback comes from the energy and momentum from supernovae, as well as winds and radiation pressure from young massive stars, which is calibrated by stellar population synthesis models. Chemical enrichment includes contributions from supernovae types II and Ia. The most novel aspect of these simulations is the explicit modeling of stars forming in clusters with a spectrum of masses that matches observations of young star clusters in the nearby universe. Growth of star clusters is terminated by their own feedback, and cluster masses are thus calculated self-consistently. This current suite has completed runs to redshift $z \approx 1.5$.

We use the outputs of several runs with different value of the local star formation efficiency per free-fall time, $\epsilon_{\rm ff}$. All runs start from the same initial conditions and have the same physics, including the supernova momentum boost factor $f_{\rm boost} = 5$. This boost factor accounts for the enhanced momentum feedback of clustered SNe vs isolated SNe, and also compensates for the momentum loss due to advection errors as the SN shell moves across the simulation grid (for details see Li et al., 2018). We take four runs with a constant value of $\epsilon_{\rm ff}$ randing from 10% to 200% (SFE10, SFE50, SFE100, and SFE200), plus one run with turbulence-dependent $\epsilon_{\rm ff}$ (SFEturb, with the median of about 3%). One additional run, SFE50-3SNR with $f_{\rm boost} = 3$, was chosen to test the sensitivity of results to the strength of feedback. Despite the different choices of $\epsilon_{\rm ff}$, all runs with $f_{\rm boost} = 5$ reproduce the galaxy stellar mass and star formation rate as expected from the abundance matching technique; the SFE50-3SNR run has an elevated star formation rate because of weaker feedback. While $\epsilon_{\rm ff}$ has little effect on the global galaxy properties, the properties of individual star cluster particles (such as the cluster mass function, maximum cluster mass, and the cluster formation timescale) depend strongly on this parameter. For this reason, we plot each of these runs separately, although it will be apparent that our results are not affected by the numerical differences between the runs.

3.2.1 NSC Masses and Sizes

We use the ROCKSTAR halo finder (Behroozi et al., 2013a) to identify the dark matter halos in the simulation outputs. Then we find the stellar center of the galaxy by doing an iterative centering process. To account for discreteness due to the small number of stellar particles in the center, we use kernel density estimation (KDE) where we smooth each star particle with a 3D Gaussian kernel, then sum the contributions to create a full stellar density field. We start by smoothing the star particles with a large kernel ($\sigma = 2\,\mathrm{kpc}$). We pick the location with the highest density, then recalculate the stellar density around that point using a kernel that is 3 times smaller. We repeat until we reach the resolution limit of the simulation, given by the cell size at the finest refinement level. The smallest kernel uses $\sigma = L_{\mathrm{cell,min}} = 3\,\mathrm{pc}$. This iterative process has the effect of first picking out large-scale galactic structure, then focusing in on the locations with the highest stellar density within those larger scale dense regions. The location of the absolute highest stellar density sometimes is in a very massive young cluster in the outskirts of the galaxy, but the algorithm avoids these clusters in favor of true nuclear clusters. Using the large scale galactic structure is essential to this process.

Next we define the plane of the galaxy. Following the formalism in Zemp et al. (2011), we use the inertia matrix to calculate the axis ratios of the galaxy. The inertia tensor can be written as

$$I_{ij} = \frac{\sum_k M_k \, \mathbf{r}_{k,i} \, \mathbf{r}_{k,j}}{\sum_k M_k} \tag{3.1}$$

where M_k is the mass of the k-th particle, and $\mathbf{r}_{k,i}$ is the i component of the position of the k-th particle. We use only the star particles to calculate the axis ratios, as we are interested in the stellar component. The eigenvalues of this tensor are $a^2/3$, $b^2/3$, and $c^2/3$, where a, b, and c are the semi-principal axes, and $a \geq b \geq c$. We define the normal vector to the plane of the galaxy to be the eigenvector corresponding to the

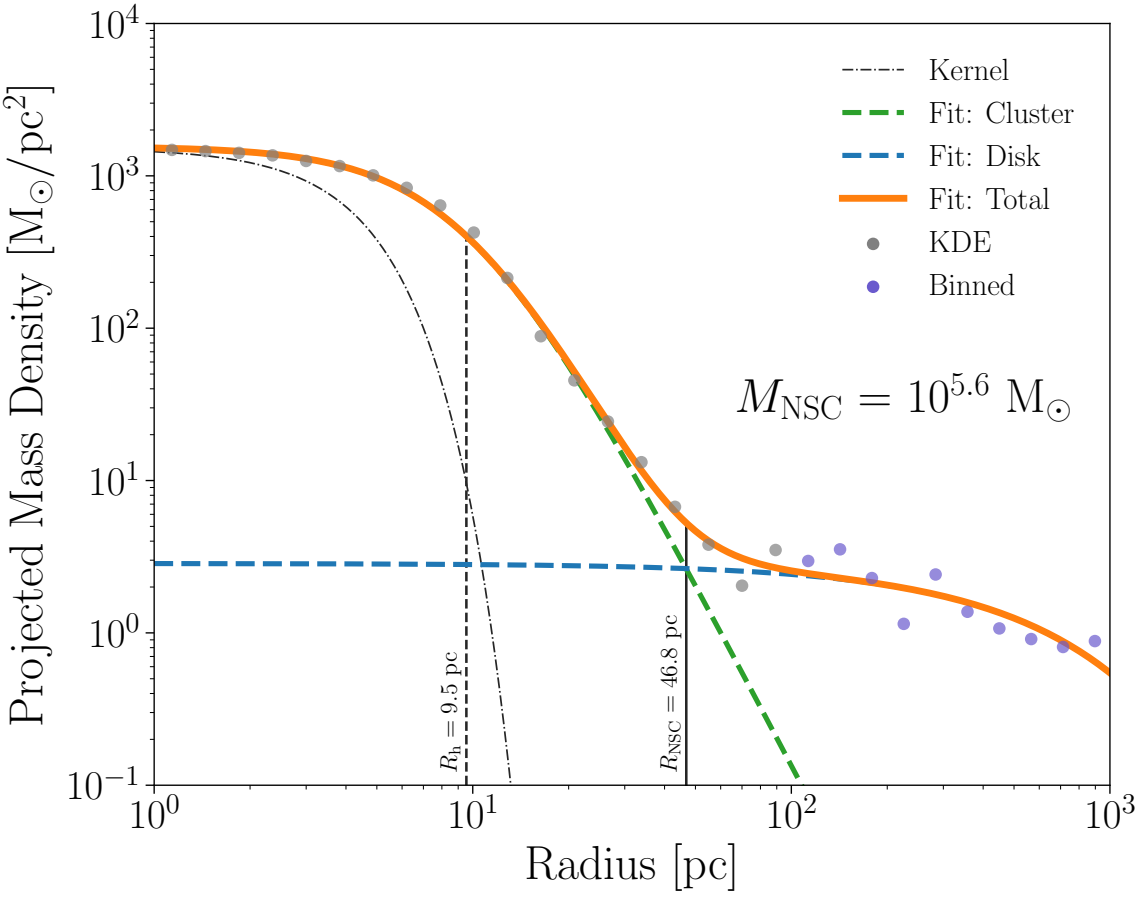

Figure 3.1 An illustration of a two-component fit of the surface density profile of a typical galaxy. The two components of the fit are a nuclear region and an outside disk. Vertical lines show the half-mass radius $R_{\rm h}$ and the full extent of the nuclear cluster $R_{\rm NSC}$. The cluster mass is indicated at the bottom right.

smallest eigenvalue ($c^2/3$).

We project all the star particles onto the plane of the galaxy, then perform a two dimensional KDE to obtain the surface density profile. For star particles within the inner 12 pc, we take the Gaussian kernel with a one sided width of 3 pc ($L_{\rm cell,min}$). Outside of 12 pc, we use a larger 6 pc kernel ($L_{\rm cell,max}$) to reduce counting noise. The surface density is calculated by integrating the KDE density over an annulus to get the mass, then dividing by the area of the annulus. This KDE estimate is used only for the central 100 parsecs, while outside this radius we simply bin the star particles.

To find the nuclear cluster, we perform a two-component fit of the stellar density

profile. Galaxies at these redshifts are expected to be mostly disky with an additional central component (nucleus or bulge), so we use the sum of an exponential disk and a Plummer sphere:

$$\Sigma_{\rm disk}(R) \propto \exp\left(-R/a_d\right), \quad \Sigma_{\rm cluster}(R) \propto [1+(R/a_c)^2]^{-2},$$

where a_d and a_c are the scale radii for the two components (Dejonghe, 1987). We define the size of the nuclear region, $R_{\rm NSC}$, to be the radius where $\Sigma_{\rm cluster}(R_{\rm NSC}) = \Sigma_{\rm disk}(R_{\rm NSC})$. Figure 3.1 shows an example of this fitting process.

To estimate the uncertainty of our determination of $R_{\rm NSC}$ after marginalizing over the other structural parameters, we construct a multivariate Gaussian distribution with dimensions given by the covariance of the parameters of the decomposition (normalization and scale radius of both components). We then sample from this multivariate Gaussian and re-calculate the radius of the nuclear region for each realization. We take the range that encloses 68.3% of these radii to be the error of $R_{\rm NSC}$.

The mass of the nucleus then is the sum of the star particles within this spherical region, $M_{\rm NSC} \equiv \sum_k M_k(r_k < R_{\rm NSC})$. We do not subtract off the disk component within this region. We also calculate the half-mass radius of the nuclear cluster, defined as

$$M(R_{\rm h}) \equiv \frac{1}{2} M(R_{\rm NSC}), \quad M(R) = \int_0^R \Sigma_{\rm KDE}(R)\, 2\pi R\, dR \tag{3.2}$$

where $\Sigma_{\rm KDE}(R)$ is the KDE surface mass density as described above. This half-mass radius is calculated in projection, unlike the mass, for a more direct comparison with observations. The errors on the mass and half-mass radius are estimated by perturbing $R_{\rm NSC}$ according to its error.

Figure 3.2 shows the masses and sizes of the galactic nuclei in the simulations, along with observed NSCs in nearby late-type galaxies of intermediate and low mass

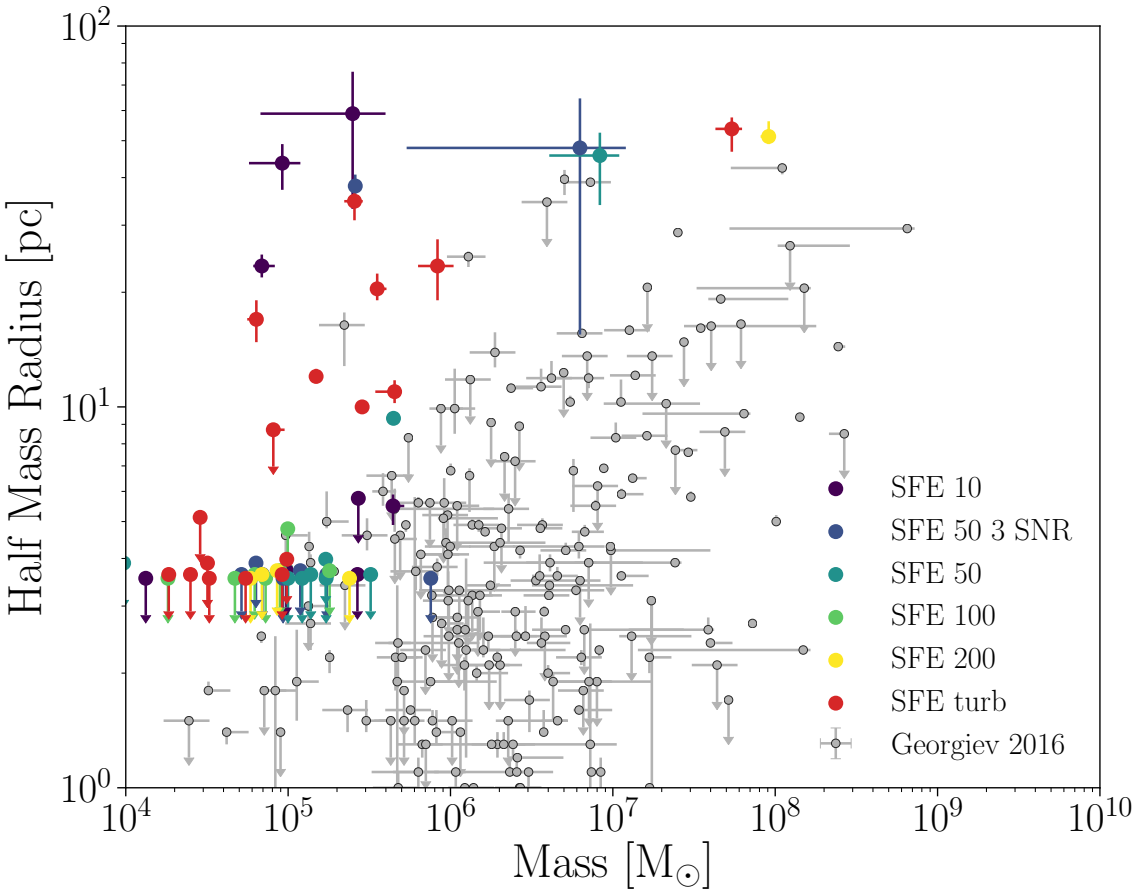

Figure 3.2 Masses vs. half-mass radii of the nuclear regions of our simulated galaxies (colored symbols). Nuclear regions where one star particle contributes more than half of the total mass are shown as having upper limits on $R_{\rm h}$. Additionally, an upper limit is given to one nuclear region with a small density contrast, where the fit parameters were uncertain enough that the 1σ limits on the mass and radius included zero. As a comparison, the sizes and masses of nuclear star clusters in late type galaxies presented in Georgiev et al. (2016) are shown in grey.

($M_\star < 10^{11}$) from Georgiev et al. (2016). We excluded one nuclear region with mass above $10^9\,\mathrm{M}_\odot$, in the SFE50-3SNR run that had insufficiently strong feedback to suppress the formation of a very dense bulge. Simulated nuclear regions with half-mass radii near 3 pc are due to a single star particle dominating the central mass distribution. For these regions, where a single particle contributes more than half of $M_{\rm NSC}$, we take the $R_{\rm h}$ given by Equation 3.2 to be an upper limit because they just reflect the size of the smoothing kernel. At low masses, there are many of these galaxies where a single star particle dominates. These unresolved clusters are consistent with many of the compact clusters in the Georgiev et al. (2016) sample.

However, the nuclei that are resolved in the simulations tend to have $R_{\rm h}$ larger than observed clusters of the same mass. This overestimate is particularly noticeable at low masses, and can be an order of magnitude or more. At high masses the discrepancy is a factor of several.

As discussed above, we select nuclear regions by means of a two-component fitting procedure. While this procedure generally provided sizes larger than the observed radii of the nearby NSCs, any more complicated model fit would be unwarranted due to the irregular structure of our galaxies. Our fit also appears to be robust and nuclear regions clearly defined. The central NSC density exceeds the extrapolated central density of the disk component typically by a factor of 10 or greater.

However, the physical interpretation of these components is not always clear. Strong feedback gives our galaxies an irregular structure at the redshift $z \approx 1.5$ of our last output, making the identification of a center unclear in some cases. Some galaxies had a roughly constant density in the center, with one massive star particle that was chosen by our algorithm to be the center. Others had a more extended central component without a strongly peaked center, where the middle was chosen to be the center. Even in these situations, however, a central component is present and can represent a nuclear region, even if it is too large to be a typical NSC. Visually

inspecting the profiles indicated that the returned sizes were an acceptable description of this central component. These results indicate that higher resolution or improved stellar physics are required to properly model NSCs.

Interestingly, we find that although NSC sizes are generally overestimated in the simulations, the overall galaxy sizes match relatively well with the observed sample of van der Wel et al. (2014). For late-type galaxies with $9.0 < \log M_\star < 9.5$ at redshift $1.5 < z < 2$, the median effective radius is 2.1 kpc, compared to 2.8 kpc for our simulated galaxies in the same range of mass and redshift. The distribution of sizes in the observed sample (68% lie between 1.2 and 3.7 kpc) is broader than in the simulations, which range from 2.5 to 3.3 kpc. The sizes of our simulated galaxies are thus statistically consistent with the available observations.

3.2.2 Ellipticity and Rotation

With the nuclear region defined, we can assess the spatial and kinematic structure of the collection of stars in this region. Using the inertia tensor (Equation 3.1), we calculate the axis ratios of the nuclear regions.

Figure 3.3 shows the axis ratios as a function of the NSC mass. Nuclear regions that have less than 10 star particles are not shown, as they have too few particles to reliably determine the axis ratios. Our nuclear regions are mildly triaxial, but all of them are less flattened than the Milky Way NSC (Feldmeier-Krause et al., 2017b).

To investigate whether the flattening is caused by rotation, we calculate the rotation velocity and the residual velocity dispersion of the NSCs. We define the plane of the galaxy in the same way we did above, based on the inertia tensor. We found that this definition works better than using the angular momentum vector of the nuclear regions, because the latter is relatively small and does not determine the structure of the NSCs. We find that the direction of the nuclear angular momentum is misaligned with the normal to the plane of the galaxy by between 20 and 90 degrees, and has no

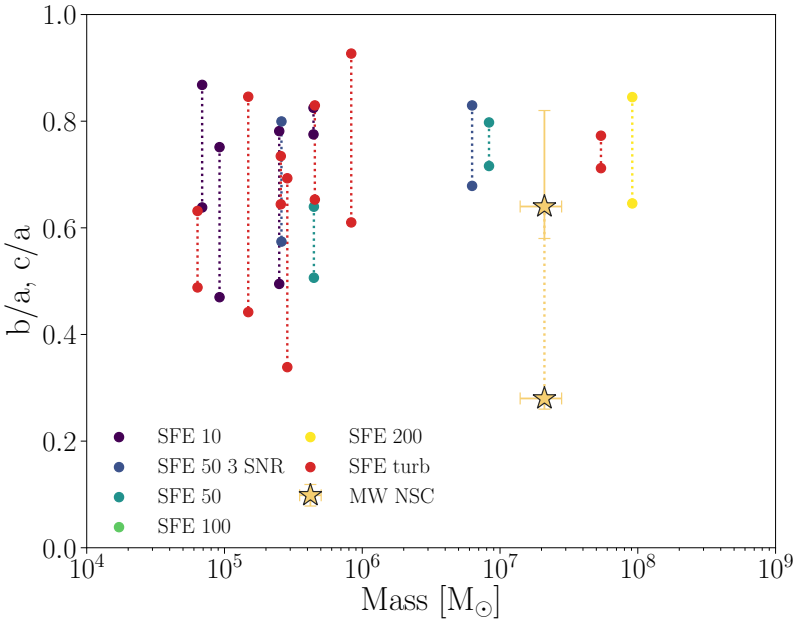

Figure 3.3 Axis ratios of our simulated galaxies as a function of mass of the nuclear region. Each galaxy has two points connected with a line. The upper shows b/a, while the lower shows c/a. Nuclear regions that had too few star particles to properly calculate the axis ratios are not shown. The Milky Way NSC is shown as a star, for comparison (Feldmeier-Krause et al., 2017b).

dependence on cluster mass.

The rotation velocity is calculated as the mass-weighted tangential velocity component of the star particles within the nuclear region. Figure 3.4 shows the ratio of rotational velocity to 3D velocity dispersion as a function of ellipticity, defined as $\epsilon \equiv 1 - c/a$. The black line shows the expected values for a system with an isotropic velocity dispersion that is flattened only by rotation. This relation can be approximated as

$$\frac{v_{\rm rot}}{\sigma_{3D}} \approx \sqrt{\frac{\epsilon}{1-\epsilon}}$$

(e.g., Mo et al., 2010). The simulated nuclei are predominately below this line, indicating that their non-sphericity is an intrinsic shape, not due to rotation. The accretion of particles composing the NSCs must have proceeded anisotropically. Indeed, we find that for nearly all clusters the radial velocity dispersion is larger than

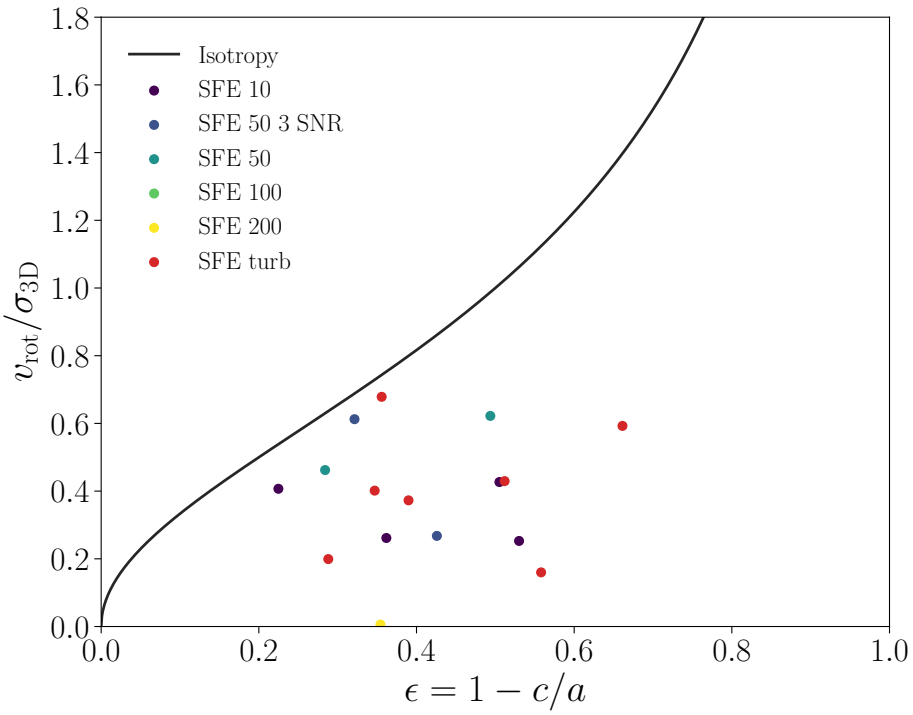

Figure 3.4 The amount of rotational support of the nuclear regions as a function of their ellipticity. The black line shows the expected values for a system with an isotropic velocity dispersion that is flattened only by rotation. Most points lie below this line, indicating that the nuclear regions are intrinsically non-spherical.

the tangential dispersion by up to a factor of two.

3.2.3 Metallicity

As described above, our suite of simulations tracks enrichment from Type Ia and II supernovae, but not AGB winds. We calculate the mass-weighted total metallicity as the sum of all metals divided by the mass of all star particles in the nuclear regions.

To calculate the spread in metallicity, we consider two sources of variance: the internal spread within a star cluster particle due to self-enrichment (recorded in the simulation runtime as the cluster is forming), and the dispersion of final metallicity among star particles.

To calculate the internal spread $\sigma^2_{Z,k}$ within the k-th particle, we transform the spread in overall metallicity from SNII. The details of how this is derived from the

simulation outputs is described in H. Li & O. Gnedin (2018, in preparation), but we summarize the key points here. The star particles in the simulation contain a variable $M^{ZZ} = \sum m_i Z_{i,\mathrm{II}}^2$ that can be used to calculate the metallicity variance of a single star particle without needing its full accretion history. Here we are summing over all accreted mass elements m_i of a specific cluster particle as it is forming at time steps i. The final particle mass is $M_k = \sum m_i$. We only consider the spread in metallicity from Type II supernovae, since the delay time for SNIa is significantly longer than the formation timescale of star clusters and therefore enrichment from Type Ia supernovae will not be relevant for this spread. The variance of the metallicity of a single star particle k is defined as:

$$\sigma_{Z,k}^2 = \frac{\sum_i m_i Z_{i,\mathrm{II}}^2}{\sum_i m_i} - \left(\frac{\sum_i m_i Z_{i,\mathrm{II}}}{\sum_i m_i}\right)^2 = \frac{M_k^{ZZ}}{M_k} - Z_{k,\mathrm{II}}^2. \tag{3.3}$$

We then combine this with the dispersion in metallicity among cluster particles to obtain the total metallicity spread:

$$\Delta Z^2 = \frac{\sum M_k \, \sigma_{Z,k}^2}{\sum M_k} + \frac{\sum M_k \left(Z_k - \overline{Z}\right)^2}{\sum M_k} \tag{3.4}$$

where $\overline{Z}$ is the mass-weighted average metallicity of the nuclear region:

$$\overline{Z} \equiv \frac{\sum M_k Z_k}{\sum M_k}. \tag{3.5}$$

Figure 3.5 shows the metallicities of the nuclear regions plotted against their mass. We represent the metallicity spread with errorbars that cover the interval $\log(\overline{Z} - \Delta Z)$ to $\log(\overline{Z} + \Delta Z)$. Our nuclear regions have a wide spread of metallicities in the range $-2 < [\mathrm{Fe/H}] < -0.25$. Other then the most massive nuclear regions typically having higher metallicity, there is no correlation between metallicity and cluster mass.

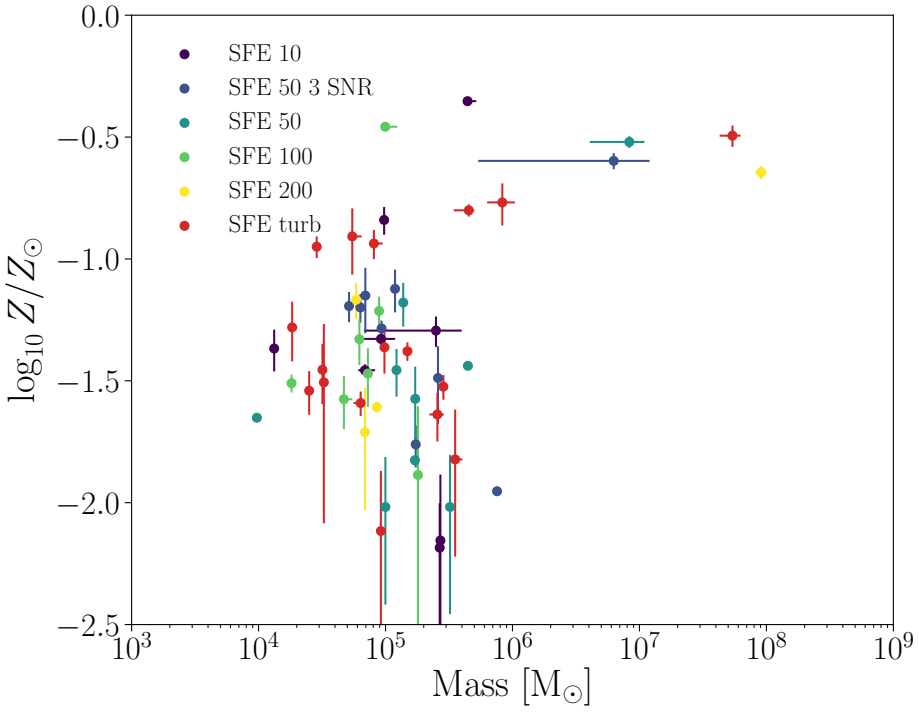

Figure 3.5 Total metallicity of the nuclear regions as a function of their mass. Errorbars cover the interval $\log(\overline{Z} - \Delta Z)$ to $\log(\overline{Z} + \Delta Z)$.

3.3 Elements

3.3.1 Yields

To turn the total metallicity into abundances for individual elements, we use computed supernova yields in the literature. We take the SNII yields of Nomoto et al. (2006) and integrate them over the Kroupa (2001) IMF from 10 to 40 $M_\odot$, using a 50% hypernova fraction for stellar masses where hypernova yields are available. Since the Nomoto et al. (2006) yields are provided at several fixed progenitor metallicities $Z = 0$, 0.001, 0.004, and 0.02, we interpolate between these models to get the yield at any metallicity. We use the W7 model of Iwamoto et al. (1999) for SNIa yields at all metallicities.

To get the elemental abundances of a given star particle, we use the SN yields and scale them to produce the appropriate mass of metals, as set by the $Z_{\rm Ia}$ and $Z_{\rm II}$

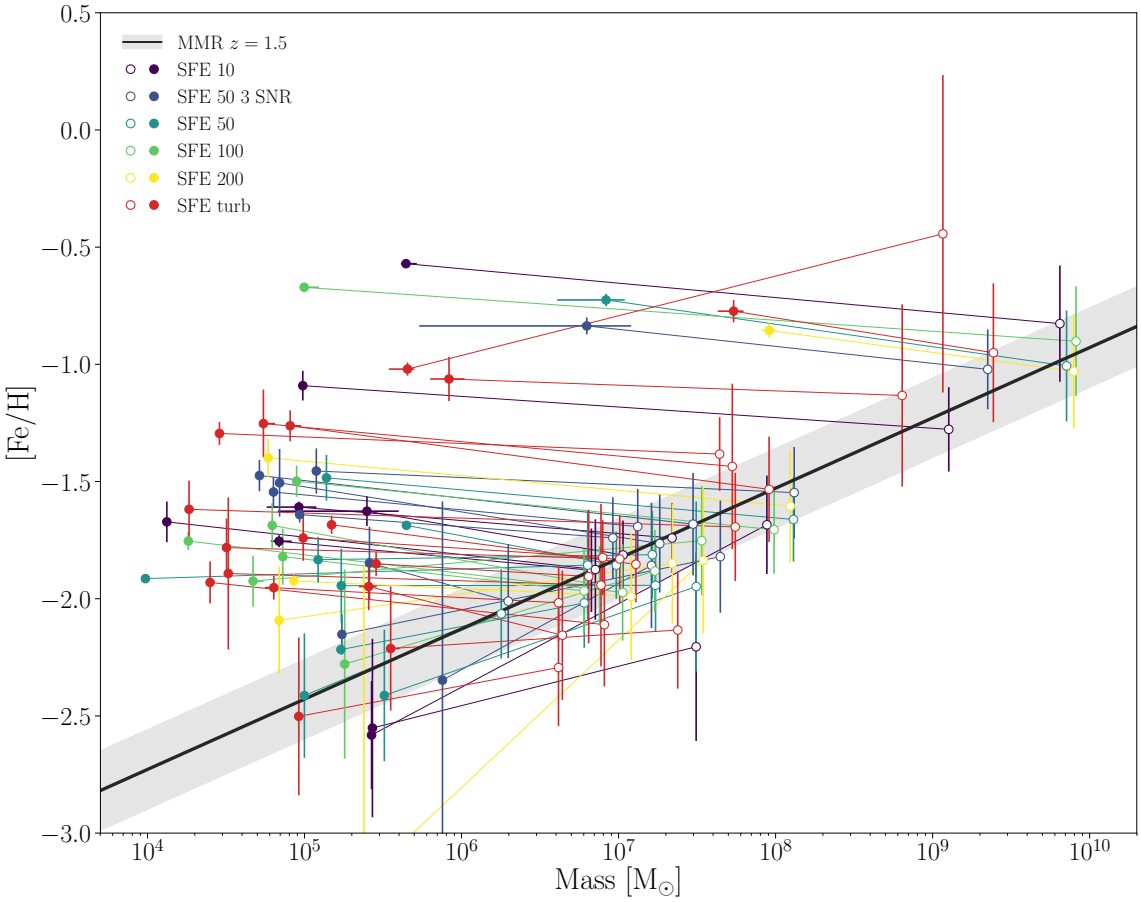

Figure 3.6 Mass-metallicity relation (MMR) for the whole simulated galaxies (open circles) and their nuclear regions (filled circles). The lines connect the nuclei to their host galaxies. Errorbars in [Fe/H] are the spread in [Fe/H] in a nuclear region or galaxy (Equation 3.7). The black line and shaded region show the MMR for nearby dwarf galaxies with its associated RMS scatter of 0.17 dex, shifted down by 0.36 dex to account for evolution in the MMR to $z = 1.5$.

variables in the simulation.

The two sets of published SN yields depend on the metallicity of the supernova progenitor star. Since this information is not recorded in the simulation output, we are using the metallicity of the star particle itself as the metallicity of the progenitor. This may overestimate the progenitor metallicity. However, as we discuss later, we use elements whose yields do not vary strongly with metallicity of the progenitor, reducing the effect on our results.

3.3.2 Mass - Metallicity Relation

By using these yields, we calculate a mass-weighted value of [Fe/H] for both the star particles in the nuclear regions and the stars in the whole galaxy. The average contribution from all star cluster particles is

$$\overline{\mathrm{[Fe/H]}} = \log\left[\frac{\sum_k M_k(Z_{k,Ia} f_{\mathrm{Fe,Ia}} + Z_{k,II} f_{\mathrm{Fe,II}})}{\sum_k M_k X_k}\right] - \log\left[\frac{Z_\odot f_{\mathrm{Fe}\odot}}{X_\odot}\right] \tag{3.6}$$

where f_{Fe} is the mass fraction of the metals that is Fe. The numerator in each fraction is the mass of iron, while the denominator is the mass of hydrogen. For the Sun we use the solar abundances of Grevesse & Sauval (1998), which are $Z_\odot = 0.0169$ and $f_{\mathrm{Fe}} = 0.0757$.

We then transform the metallicity spread (Equation 3.4) into spread in [Fe/H] by using:

$$\Delta\mathrm{[Fe/H]}^2 = \Delta Z^2 \left(\frac{d\mathrm{[Fe/H]}}{dZ_{\mathrm{II}}}\right)^2. \tag{3.7}$$

We calculate this derivative numerically using interpolation of the published SN yields. While this equation is written using [Fe/H], it will also be used to calculate spreads in other elemental abundances.

We calculate the mean and total variance in [Fe/H] for both the nuclear regions and the galaxy as a whole. The results of these calculations are shown in Figure 3.6, along with an empirical mass-metallicity relation (MMR) for dwarf galaxies from Kirby et al. (2013), shifted down by 0.36 dex to account for evolution in the MMR from $z = 0$ to $z = 1.5$, based on the interpolation of the results of Mannucci et al. (2009) for Lyman-break galaxies at $z \approx 3$.

The model galaxies lie near the expected $z = 1.5$ relation, without any fitting, which provides evidence that our simulations are modeling the chemical enrichment of galaxies correctly. The nuclei are systematically more metal rich. This indicates that our galaxies have a metallicity gradient, consistent with observations (e.g. Moustakas

et al., 2010). The predicted iron abundances of NSCs span a wide range $-2.5 <$ [Fe/H] < -0.5, uncorrelated with their mass.

3.3.3 Effects of Extended Star Formation

To examine the effect of an extended star formation history, we first determine the time it took to form the bulk of the stellar mass in each nuclear region. To minimize the effect of few particles born much sooner or later than the bulk of the particles, we pick the smallest time interval in which 90% of the mass was formed. This interval is calculated using only the birth time of the particles, not including the time each star particle took to form. To account for this additional time, we add the maximum duration of star particle formation within the cluster (2 Myr $\lesssim \tau_{\rm dur} \lesssim$ 10 Myr; Li et al. 2017) to the initial time interval.

Figure 3.7 shows that the iron spread in the nuclear regions does not correlate with the length of star formation history. In fact, several NSCs with the longest assembly time have relatively small spread, Δ[Fe/H] < 0.15, most of which is due to different metallicities of its constituent particles, rather than the intrinsic spread within each particle. The largest spread is seen in NSCs that took only 2-3 Myr to form, where all the spread is internal. These regions tend to be low metallicity ([Fe/H] < -1.8) and consist of only one or few star particles.

The largest predicted iron spread is comparable with the spread in three massive globular clusters in the Milky Way where it is clearly detected: ω Cen, M54, and Terzan 5. To make a fair comparison to our models, we define the observed spread as the RMS scatter of the [Fe/H] content of the stars in each cluster. This value is calculated for M54 by Carretta et al. (2010a) and for Terzan 5 by Massari et al. (2014), while we calculate it for ω Cen using the data from Marino et al. (2011) In Figure 3.7 we do not show the observed age spreads, only the metallicity spreads. The age spreads are uncertain, with inconsistencies between different studies of ω Cen (Joo

& Lee, 2013; Villanova et al., 2014; Tailo et al., 2016). In our simulations we are able to detect both age and metallicity spreads significantly below the observational limits, which remain relatively large. However, the simulated assembly timescales are within the inferred age spread of stellar populations in ω Cen of $\sim$ 500 Myr (Tailo et al., 2016), while being shorter than the age spreads in M54 and Terzan 5 (Siegel et al., 2007; Ferraro et al., 2016). As our simulations only reach $z \approx 1.5$ ($t \approx 4.3$ Gyr), we are not able to model the large age spreads present in these two clusters. These results lend support to the hypothesis that stripped NSCs could become progenitors of the anomalous globular clusters. There are still quantitative differences that need to be explored in future work: the nuclei with largest iron spread all have lower metallicity than the three observed anomalous clusters and lie in the lower mass range below $10^6\,\mathrm{M}_\odot$.

3.3.4 Reliability of Light Element Yields

When determining which elements can be modeled reliably by our simulations, we consider several aspects.

(i) Our simulations only model supernovae, not AGB stars or other sources of chemical enrichment, so we select elements without strong AGB contributions.

(ii) As discussed above, we are using the metallicity of the star particle itself as the metallicity of the supernova progenitor star. To reduce the impact on our results, we prefer elements whose yields do not vary strongly with metallicity of the progenitor.

(iii) Since the two sets of yields give generally different results, we prefer the elements with low spread between the different model sets.

(iv) We select the elements whose abundances match observations, whether in direct observation of SN ejecta, observations of low metallicity stars, or chemical evolution models.

Figure 3.8 shows how our two yield sets (Woosley & Weaver, 1995; Nomoto et al.,

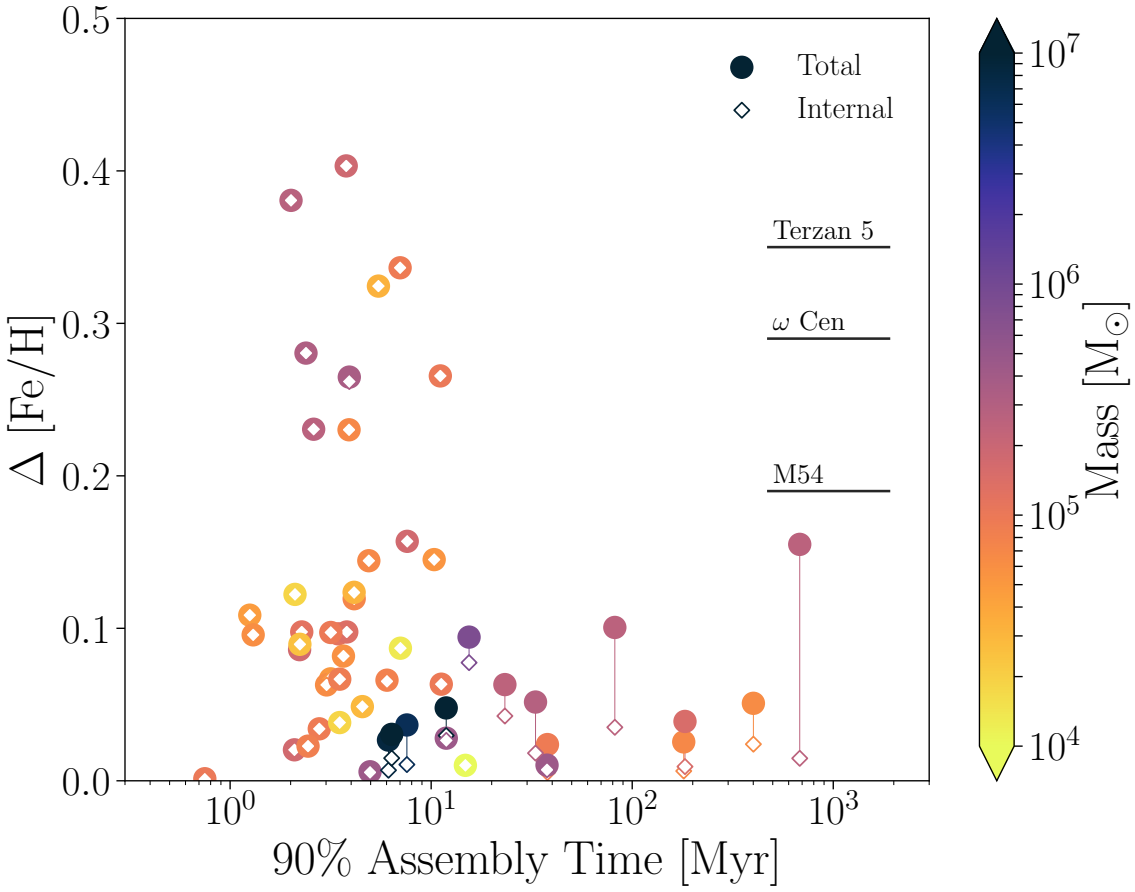

Figure 3.7 Dispersion of Fe in the nuclear regions as a function of the time required to assemble 90% of their mass. The contribution from internal star particle metallicity spread is shown by hollow diamonds and labeled "Internal", while the total dispersion including spread among the multiple star particles in the nuclear region is shown by filled circles and labeled "Total". The internal and total components for each NSC are connected by a vertical line. The points are color-coded by the mass of the nuclear region. The horizontal lines on the right show the observed Fe spreads in massive globular clusters of the Milky Way (see text for references).

2006) vary with metallicity, to address the second and third points. Table 3.1 describes in more detail several elements that are relevant for globular clusters and NSCs, the production source of those elements, and comments on the reliability of the yields. This table includes only elements that we deemed relevant for our analysis. We do not discuss elements such as C and N, which have significant AGB contributions, or other elements that are not strongly involved in the phenomenon of multiple populations.

To evaluate the reliability of our use of supernovae yields to extract elemental abundances for these light elements, we surveyed the literature for studies examining this issue. Romano et al. (2010) and Andrews et al. (2017) put different sets of yields through the same chemical evolution model, then compared the result to observations, showing how different model choices affect the resulting abundances. These papers show both the scatter among different models and the differences between models and data, which are important for assessing reliability of the yields. Andrews et al. (2017) and Pignatari et al. (2016) describe the production sites of the various elements. Wiersma et al. (2009) implement the yields into cosmological smoothed-particle-hydrodynamics simulations and show in their figure A4 the variations caused by different yields.

The synopsis of these studies is that Oxygen is modeled reliably. It matches observations, has a weak metallicity dependence, and has a reasonably small spread between different yield sets. Magnesium is similar to Oxygen. Aluminum and Sodium have much stronger metallicity dependences and wider spreads, making their predicted abundances less certain.

3.3.5 Predicted Correlations of Element Abundances

To test whether our nuclear regions may have properties similar to the observed multiple populations in globular clusters, we examine their detailed elemental abundances. First we compare the overall normalization of the abundances as a function

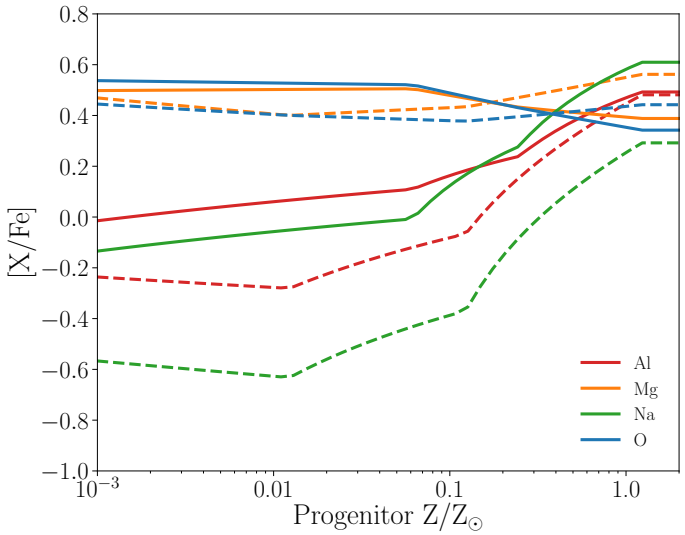

Figure 3.8 Abundance ratios of the ejecta of SN type II as a function of metallicity of the SN progenitor, for two yield models: solid lines are for Nomoto et al. (2006), dashed lines are for Woosley & Weaver (1995). Note that we have adjusted the WW Fe yields down by a factor of 2 and the Mg yields up by a factor of 2, as suggested by other authors (Timmes et al., 1995; Wiersma et al., 2009; Andrews et al., 2017). We aim to study elements that do not vary strongly with Z and have a low discrepancy between the models.

of metallicity. Figure 3.9 and Figure 3.10 show [O/Fe] and [Mg/Fe] for our nuclear regions and compare them to the compilation of 202 red giants from 17 globular clusters presented in Carretta et al. (2009a). Because of the limited sample size of the available measurements, we combine the available data from all clusters and plot the regions enclosing 50% and 90% of all stars. We compare this total range of observed abundances with our model predictions.

In each plot a solid line shows the track the objects would follow if their metallicity came entirely from Type II supernovae. Points below this line indicate some enrichment from Type Ia supernovae, which contribute Fe, but not O or Mg.

Additionally, we examine the elemental anti-correlations (Na-O and Al-Mg) that are key features of the multiple population phenomena in globular clusters. We compute the [Na/Fe] and [O/Fe] ratios for each star particle. Since the underlying metallicity dispersion is responsible for the elemental dispersion of a star particle, the

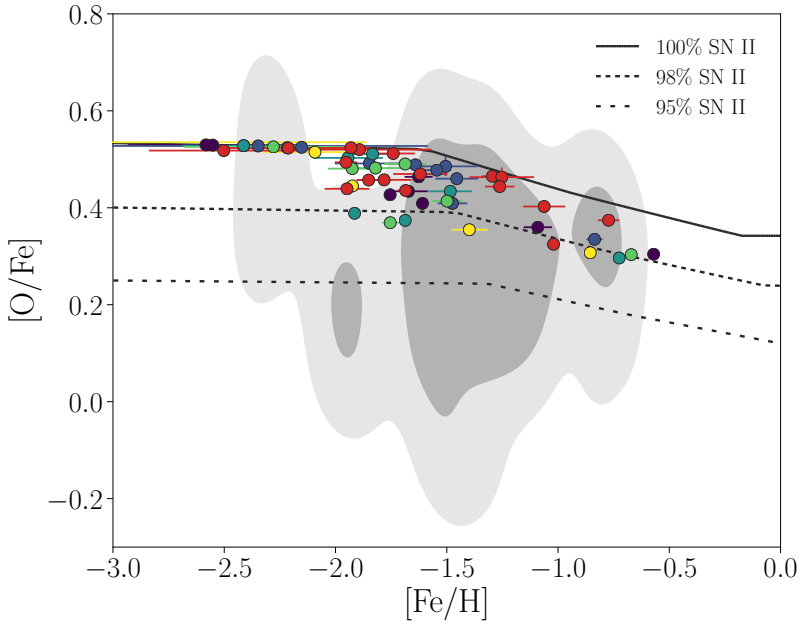

Figure 3.9 Oxygen abundances for model nuclear regions. The dark (light) contours show the location of 50% (90%) of globular cluster stars from Carretta et al. (2009a). Black lines show the elemental ratios of the SN yields, for different percentage contributions of SN type II.

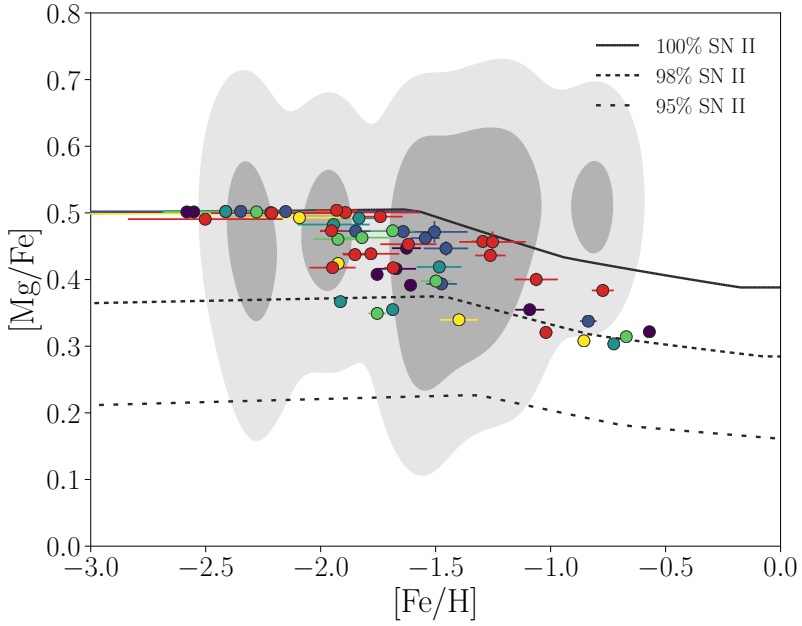

Figure 3.10 Same as Figure 3.9, but for Magnesium.

elemental dispersions will be correlated. A given metallicity spread will cause a spread along a certain direction in the [Na/Fe]-[O/Fe] plane. We smooth each star particle by a Gaussian whose width along this direction will reproduce the elemental spreads (see Equation 3.7). For visual clarity, a lower threshold on this width was chosen to be 0.005 dex, which was also used as the width perpendicular to the direction of the spread. We mass-weight the contributions of all star particles in a cluster, then plot contours that enclose 50% and 90% of the mass for each cluster. Figure 3.11 shows the results of this for the Na-O anticorrelation, while Figure 3.12 shows the Mg-Al anti-correlation. Black lines show the tracks that objects would follow if their metallicity came from a given percentage of Type II supernovae, for metallicities $-3 < [\mathrm{Fe/H}] < 0$. Nuclear regions consisting of a single particle will exhibit spreads along these lines, while regions with multiple particles can have spreads in any direction due to potential variations between particles. It is apparent that nearly all of our simulated nuclear regions do not have a significant spread in any elemental abundance, unlike stars in globular clusters. It is also apparent that the yield lines are not capable of producing the full range of elemental variations, especially for Al, even though they include metallicities up to solar.

As seen in Equation 3.7, elements whose yields change more with metallicity will have larger spreads. Na and Al yields change more than O and Mg (see Figure 3.8), explaining their larger abundance spreads in these figures.

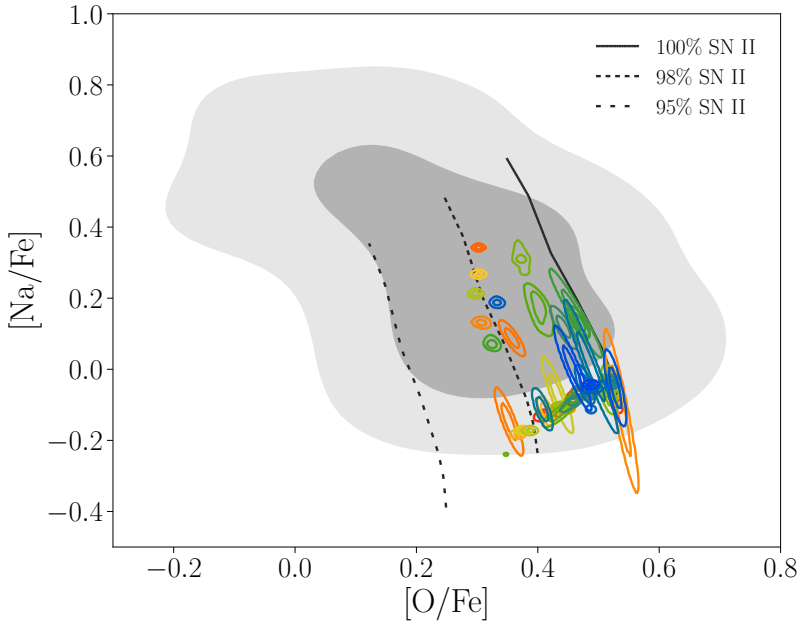

Figure 3.11 Na-O anticorrelation for model nuclear regions (small colored contours) vs. observed globular cluster stars (shaded contours) from Carretta et al. (2009a). The color of the lines simply serves to distinguish different model nuclear regions. Black lines show the elemental ratios of the SN yields, for different percentage contributions of SN type II. [Fe/H] varies along each line from -3 to 0.

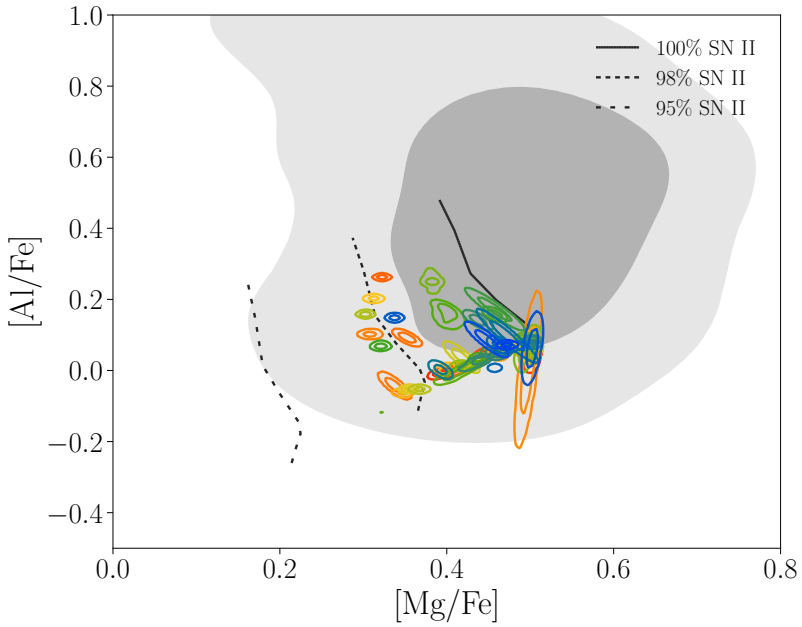

Figure 3.12 Same as Figure 3.11, but for Al-Mg anticorrelation.

Table 3.1 Summary of elements relevant to GCs. Citations are abbreviated as follows – A17: Andrews et al. (2017), C09: Carretta et al. (2009b), FT17: Fernández-Trincado et al. (2017), G12: Gratton et al. (2012), M15: Marino et al. (2015), P16: Pignatari et al. (2016), R10: Romano et al. (2010), R15: Renzini et al. (2015), W09: Wiersma et al. (2009).

Element	Importance & Measurements	Production Source	Reliability of Theoretical Yields
O	Depleted in second-generation (SG) stars as part of Na-O anticorrelation (R15, C09).	Massive stars (A17). Yield not affected by explosive nucleosynthesis (P16).	Good agreement between galactic chemical evolution (GCE) models and observations. Predictions may vary by as much as 0.6 dex at lower metallicities, but less than 0.2 dex at solar metallicity (R10, A17).
Na	Enriched in SG stars as part of Na-O anticorrelation (R15, C09).	Massive stars (A17). Mainly comes from hydrostatic carbon burning, and is consumed in explosions (P16).	Strong metallicity dependence that does not match observations well at low metallicity (R10, A17). Models have large (> 0.5 dex) spread at low metallicity (R10).
Mg	Depleted in SG stars as part of Mg-Al anticorrelation (G12, FT17).	Massive stars (A17). Created both in the pre-SN and explosive nucleosynthesis phases (R10, P16).	Matches observations well (R10, A17). Relatively low spread between models (~ 0.3 dex) for all but the lowest metallicities (W09, R10).
Al	Enriched in SG stars as part of Mg-Al anticorrelation (G12, FT17).	Massive stars (A17). Created both in the pre-SN and explosive nucleosynthesis phases (P16).	Moderate metallicity dependence that does not match observations, but it closer than the Na yield. Spread in models is smaller than for Na, but larger than for O and Mg (R10, A17).
Fe	Some very massive "anomalous" clusters have spread in Fe: ω Cen, M54, Terzan 5 (M15).	SN type Ia and type II (A17).	Models are consistent with each other (W09). The adopted mass cut and details of the explosion are the primary drivers of discrepancy (W09, R10).

3.4 Discussion

The results presented above test the hypothesis that the abundances of nuclear star clusters are consistent with the observed trends in massive globular clusters. Here we discuss these results and the possible presence of dark matter within the NSCs.

3.4.1 Elemental Abundances

As shown in Section 3.3.5, we are unable to reproduce the full spread in light element abundances that are present in globular clusters. For [O/Fe] and [Mg/Fe], the one-sided spread is less than 0.02 dex for all clusters, while for [Na/Fe] and [Al/Fe] the spread is 0.1 dex at most. The observed one-sided spreads are significantly larger: about 0.2 dex for [O/Fe] and [Na/Fe], 0.1 dex for [Mg/Fe], and 0.3 dex for [Al/Fe]. The higher spread in Na and Al in our clusters is due to the stronger variation in the yields of these elements with metallicity (Figure 3.8). Even with this large variation, even the yields themselves are not able to produce as much Al as is present in enriched second-generation globular cluster stars.

The spread we calculate includes both the internal spread within a star particle due to a metallicity variation in the accreted gas during its formation, as well as the spread among different star particles within a given nuclear region. Even with both of these sources, the elemental spreads are too small. Our simulations would be able to reproduce the distinct stellar populations by having multiple star particles with differently enriched material, but we do not see this. Many nuclear regions consist of only one star particle, but those that have multiple do not show the necessary variations. In addition, a substantial age spread cannot be responsible for the elemental spreads, as none of our nuclear regions reach the required light element spread, regardless of their formation timescale.

As described in Section 3.2.1, the sizes of our nuclear regions are an acceptable description of the central component of the central density component, even though

they are typically larger than observed NSCs. This indicates that the particles that are included within each NSC are a reasonable selection. The main conclusion of our calculation is that the extended star formation histories, without internal pollution, are insufficient to reproduce the observed abundance trends in globular clusters. This conclusion would only be strengthened if our selection of NSC particles is an over-estimate. If we reduce the sizes and masses of model NSCs, their abundance spread would only decrease.

This indicates that our simulations do not contain the enrichment sources that are the true cause of the spreads. As supernovae are the only enrichment source we include, these results provide support to the idea that the polluter responsible for globular cluster abundances must lie within the clusters during their formation epoch. A number of models have been proposed to explain these light element abundance variations. Most scenarios call for multiple generations of star formation within the globular cluster's formation epoch, with enriched material from a first generation of stars seeding a second metal enriched generation. Possible versions suggested in the literature include supermassive stars (Denissenkov & Hartwick, 2014), rotating massive stars (Decressin et al., 2007; Krause et al., 2013), AGB stars (Cottrell & Da Costa, 1981; Renzini et al., 2015), and interacting binaries (de Mink et al., 2009). However, all of these current scenarios have some shortcomings in matching the full set of available observations, as discussed by Bastian & Lardo (2018). These models fail to produce correct elemental abundances of the globular clusters, both for the elements listed above and Helium. Additionally, to reproduce the large fraction of enriched stars in a globular cluster, most models require huge amounts of mass loss (of only first generation stars), which does not appear to be realistic (Larsen et al., 2012; Cabrera-Ziri et al., 2015).

One important aspect of our simulations is that they only reach $z \sim 1.5$. As later star formation would make the masses, metallicity spreads, and age spreads increase,

the quantities we report for those properties are best interpreted as lower limits. This complicates the comparison of our NSCs to observed NSCs and GCs, which are all at $z \approx 0$. For example, the later bursts of star formation in Terzan 5 and M54 would not be present in our simulation.

With this caveat in mind, we will discuss the predicted iron spread in model NSCs. The majority of clusters have a spread less than 0.1 dex, which is roughly the limit of the observational sensitivity (e.g., Gratton et al., 2004; Willman & Strader, 2012). However, there are some clusters with [Fe/H] spreads larger than this that could be the progenitors of the iron-complex GCs. These model NSCs are very metal-poor, with the metallicity distribution dominated by internal dispersion. Our simulations include also two nuclear regions that have a longer formation timescale and a [Fe/H] spread dominated by variation among the star particles. These regions may be analogous to clusters like ω Cen or M54, which may share a similar origin (Carretta et al., 2010b). Our objects formed on a timescale consistent with that of ω Cen, which appears to have an age spread of at least 500 Myr (Tailo et al., 2016). The metallicity distribution of the object with the largest spread between particles has a central peak at a metallicity consistent with the peak of the metallicity distribution of ω Cen and M54, but the tail extends to lower metallicities rather than higher. Our objects have lower mass than the observed clusters as well. These quantitative differences need to be explored in future work, but our results lend support to the hypothesis that objects like ω Cen could be made through tidal stripping of NSCs.

If this model is correct, it has implications for the location of the polluter responsible for the light element spreads of the massive GCs. If the NSC formed by in-situ formation (e.g. Hartmann et al., 2011), the polluter must be present for each new burst of star formation. However, if the NSC formed by the inspiral of globular clusters (e.g. Antonini et al., 2012; Gnedin et al., 2014; Tsatsi et al., 2017), the presence of already existing light element spreads in each cluster would provide a natural

explanation for the fact that light element spreads are found within each peak in the metallicity distribution of the iron-complex GCs. No polluter would be needed in the central region of the galaxy. Future observations may be able to test these hypotheses by looking for light element abundance spreads in NSCs. If not detected, both the model of massive GC formation discussed in this paper and the GC inspiral model of NSC formation would need to be discarded. If light element abundance variations are present in NSCs, these models would remain plausible.

3.4.2 Dark Matter

The amount of dark matter contained within the extent of NSCs is important for their ability to represent progenitors of massive globular clusters. Even the largest globular clusters can have only dynamically sub-dominant amount of dark matter within their optical radii (e.g., Conroy et al., 2011; Ibata et al., 2013). Since NSCs form at the bottom of the dark matter potential well, they could retain some of the dark matter even if the rest of the galaxy is tidally stripped.

We find that only one in every 11 NSCs contains any dark matter particles within $R_{\rm NSC}$. This does not necessarily imply that no dark matter is present. The mass and spatial resolution for dark matter is relatively low in our simulations and may prevent reliable estimate of the dark matter content within the compact clusters. The dark matter particle mass is about $10^6\,{\rm M}_\odot$, so that even one particle can be more massive than the whole NSC. Also, the spatial force resolution for dark matter is kept separate from the gas and star particle resolution, at $50-100\,{\rm pc}$, so that the discreteness of massive dark matter particles does not affect the dynamics of gas and stars. This resolution makes it difficult to put strict constraints on the amount of dark matter present, but we can still conclude that dark matter is dynamically sub-dominant within the model nuclear regions. If these central regions turned into an object resembling a massive globular cluster after being tidally stripped by the host

galaxy, the remaining amount of dark matter would likely not violate observational constraints.

3.5 Summary

We have investigated the origin of stellar populations of nuclear star clusters in galaxy formation simulations. We analyzed the outputs of a suite of simulations at redshift $z \approx 1.5$ and identified nuclear regions of galaxies using a two-component fit of the stellar surface density profile. The range of inferred NSC masses is consistent with the observed nuclear regions of nearby galaxies, although the sizes are generally overestimated. We find that the shapes of model NSCs are moderately flattened, but not by rotation.

We examine two of the deviations from a simple stellar population present in massive globular clusters (age and metallicity spreads), and how they affect light element abundances. Age spreads are derived directly from the simulation outputs, while we calculated the abundances of several elements (Fe, O, Na, Mg, Al) by applying theoretical model yields to Type II and Ia supernova ejecta calculated in the simulations.

Our main results can be summarized as follows.

• The nuclear regions are systematically more metal-rich than their host galaxies. The average metallicities of galaxies in the simulations match the observed mass-metallicity relation for galaxies at $z \approx 1.5$.

• We find some nuclear regions with a large spread in Fe that could be progenitors of objects like M54 or ω Cen.

• The predicted spread of light element abundances in NSCs is significantly smaller than that observed in globular clusters, even in clusters with a large age or [Fe/H] spread.

• We find no clear dependence of these trends on the local efficiency of star for-

mation $\epsilon_{\rm ff}$ used in our simulations.

• Our nuclear regions do not contain significant quantities of dark matter.

Our results show that NSCs can plausibly be the progenitors of the massive iron-complex globular clusters. However, these metallicity spreads cannot contribute significantly to the observed light element spreads. The observed abundance spread must involve additional sources, internal to the clusters.

CHAPTER IV

Improving Performance of Zoom-In Cosmological Simulations using Initial Conditions with Customized Grids

This chapter was published as: ***Brown, G.****, & Gnedin, O. Y. 2021, New Astronomy, 84, 101501*

Abstract

We present a method for customizing the root grid of zoom-in initial conditions used for simulations of galaxy formation. Starting from the white noise used to seed the structures of an existing initial condition, we cut out a smaller region of interest and use this trimmed white noise cube to create a new root grid. This new root grid contains similar structures as the original, but allows for a smaller box volume and different grid resolution that can be tuned to best suit a given simulation code. To minimally disturb the zoom region, the dark matter particles and gas cells from the original zoom region are placed within the new root grid, with no modification other than a bulk velocity offset to match the systemic velocity of the corresponding region in the new root grid. We validate this method using a zoom-in initial condition containing a Local Group analog. We run collisionless simulations using the original

and modified initial conditions, finding good agreement. The dark matter halo masses of the two most massive galaxies at $z = 0$ match the original to within 15%. The times and masses of major mergers are reproduced well, as are the full dark matter accretion histories. While we do not reproduce specific satellite galaxies found in the original simulation, we obtain qualitative agreement in the distributions of the maximum circular velocity and the distance from the central galaxy. We also examine the runtime speedup provided by this method for full hydrodynamic simulations with the ART code. We find that reducing the root grid cell size improves performance, but the increased particle and cell numbers can negate some of the gain. We test several realizations, with our best runs achieving a speedup of nearly a factor of two.

4.1 Motivation

Numerical simulations are an indispensable tool for understanding the complex processes that occur during galaxy formation. Various groups, using different numerical approaches, have used simulations to successfully replicate galaxy properties and interpret observations (e.g., Vogelsberger et al., 2020). One approach to simulating individual galaxies within a cosmological context is with so-called "zoom-in" initial conditions (Navarro & White, 1994; Hahn & Abel, 2011). They rely on first simulating a large volume of the universe containing many galaxies with coarse resolution, and selecting a region around a particular galaxy, such as one resembling the Milky Way. Then this comparatively small region is resampled with many more resolution elements while keeping the initial low resolution in the rest of the volume. The zoom-in technique allows simulations to model galaxies specifically picked to have desirable properties at very high resolution, enabling detailed modeling of the relevant physical processes.

Particularly common are zoom-in simulations designed to target the Milky Way or larger Local Group volume. Recent surveys (e.g. Gaia Collaboration et al., 2018;

Majewski et al., 2017) have collected a large amount of data for Milky Way and other Local Group galaxies. To help interpret these observations, theorists have long been simulating isolated galaxies with similar masses to the Milky Way (e.g., Hopkins et al., 2014). However, recent advances both in the knowledge of the Local Group's formation history (Hammer et al., 2007; D'Souza & Bell, 2018) and improved understanding from simulations of how isolated galaxies are different from those in groups (Santistevan et al., 2020), have shown that the Local Group's environment is essential to understanding its formation. As a result, any computational setup aiming to uncover the origin of the Milky Way must be specifically tailored to match the known properties of the full Local Group, including the assembly histories of its galaxies.

As several key galaxy properties such as halo mass and merger history are determined by the initial conditions (ICs) used in a given simulation, these ICs play a critical role in any comparison of results between different simulations. Many galaxy formation simulations have been run by different groups, using different codes and different ICs (e.g., Hopkins et al., 2014; Ceverino et al., 2014; Vogelsberger et al., 2014; Wang et al., 2015; Schaye et al., 2015; Sawala et al., 2016; Wetzel et al., 2016; Grand et al., 2017; Li et al., 2017; Hopkins et al., 2018). These different ICs can lead to different galaxy properties, even before accounting for differences in modeled physics or numerical implementation, making a comparison of results challenging. If groups were to use the same ICs (particularly for zoom-in ICs that replicate the properties of the entire Local Group), any differences caused by different ICs would be eliminated, making it easier to understand how modeling differences affect the results. However, such zoom-in simulations are often computationally expensive, requiring on the order of 10^6 CPU hours (Wetzel et al., 2016; Hopkins et al., 2018). Importantly, the performance of different codes may be affected by the details of the ICs, particularly for zoom-in ICs, where the root grid cells are usually coarse and most of the volume of the

box is uninteresting. This may make simulations using some ICs prohibitively expensive when run using certain codes. Customizing these parameters, while preserving the properties of the zoom region, may improve computational efficiency and lower the cost of these high-resolution simulations of galaxy formation, making a commonly used set of initial conditions more practical and facilitating code comparison.

Here we present a method of customizing the root grid in ICs to improve code performance. This method can be used to reduce the box size and increase the resolution of the root grid. As an example we use the "Thelma & Louise" IC, initially presented in Garrison-Kimmel et al. (2014). It is a zoom-in IC that contains a Local Group analog. In particular, the merger history of the Milky Way and M31 analogs has a qualitative resemblance to the Local Group.

We validate the method using the Adaptive Refinement Tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999, 2003; Rudd et al., 2008). The ART code uses adaptive mesh refinement of the simulation grid to provide high resolution only in interesting regions where it is needed. Specifically, it starts with a uniform grid of root cells that are refined if they meet one of several criteria, such as a density threshold or a comparison to the local Jeans length. The size of the root grid cells is relevant because the load balancing algorithm operates on these root cells. Each root grid cell and all its refined "children" are assigned to a given MPI rank. Very large root grid cells therefore may contain entire structures that are all assigned to a given rank. For example, in a Local Group-like simulation, both main galaxies may each be located in single root grid cells, with very little in any other root grid cells. This decreases efficiency for two reasons. First, it is difficult to evenly divide the needed workload among different computational nodes. Such uneven load balancing may result in one rank taking much longer than the rest, wasting computation time as other processors wait. Second, it is difficult for the ART code to scale to a large number of nodes, as adding more nodes will not bring any benefit if the work cannot be subdivided that

finely. Reducing the root grid cell size will address both of these issues.

We also demonstrate that this method reproduces basic properties of the target galaxies, such as their $z = 0$ dark matter halo masses, growth histories (including times and masses of major mergers), and their populations of satellite galaxies.

4.2 Method

A detailed description of the process of creating our cosmological ICs can be found in Hahn & Abel (2011), but we summarize some of the key points here to provide context for our method. To start, a cube of a chosen cosmic volume is covered by a uniform 3D grid, known as the "root grid." A white noise field is obtained by assigning a random value sampled from a $\mathcal{N}(0, 1)$ distribution to each cell in the root grid. This is then convolved with the matter power spectrum to obtain realistic matter overdensities. The convolution is done using a Fast Fourier Transform, which enforces periodic boundary conditions on the cube. Lagrangian perturbation theory (e.g. Zeldovich, 1970) is used to turn the overdensity field into particle positions and velocities. When used for grid-based codes such as ART, this process creates the gas densities and velocities for root grid cells, as well as the positions and velocities for each cell's corresponding dark matter particle ("root grid particle").

When producing zoom-in ICs, a region of interest is selected, then a finer grid is used in that region. The spacing of the finer grid can be arbitrarily small, only limited by the computational resources required to run a simulation at such high resolution. A buffer of intermediate grid levels is used around the zoom region to transition smoothly to unrefined regions. The process for generating the properties of the dark matter particles and gas cells (referred to as "zoom particles" and "zoom cells") in the zoom region follows the same basic process as the root grid, just on a finer grid and with constraints to ensure consistency with the root grid (see Hahn & Abel 2011 for full details).

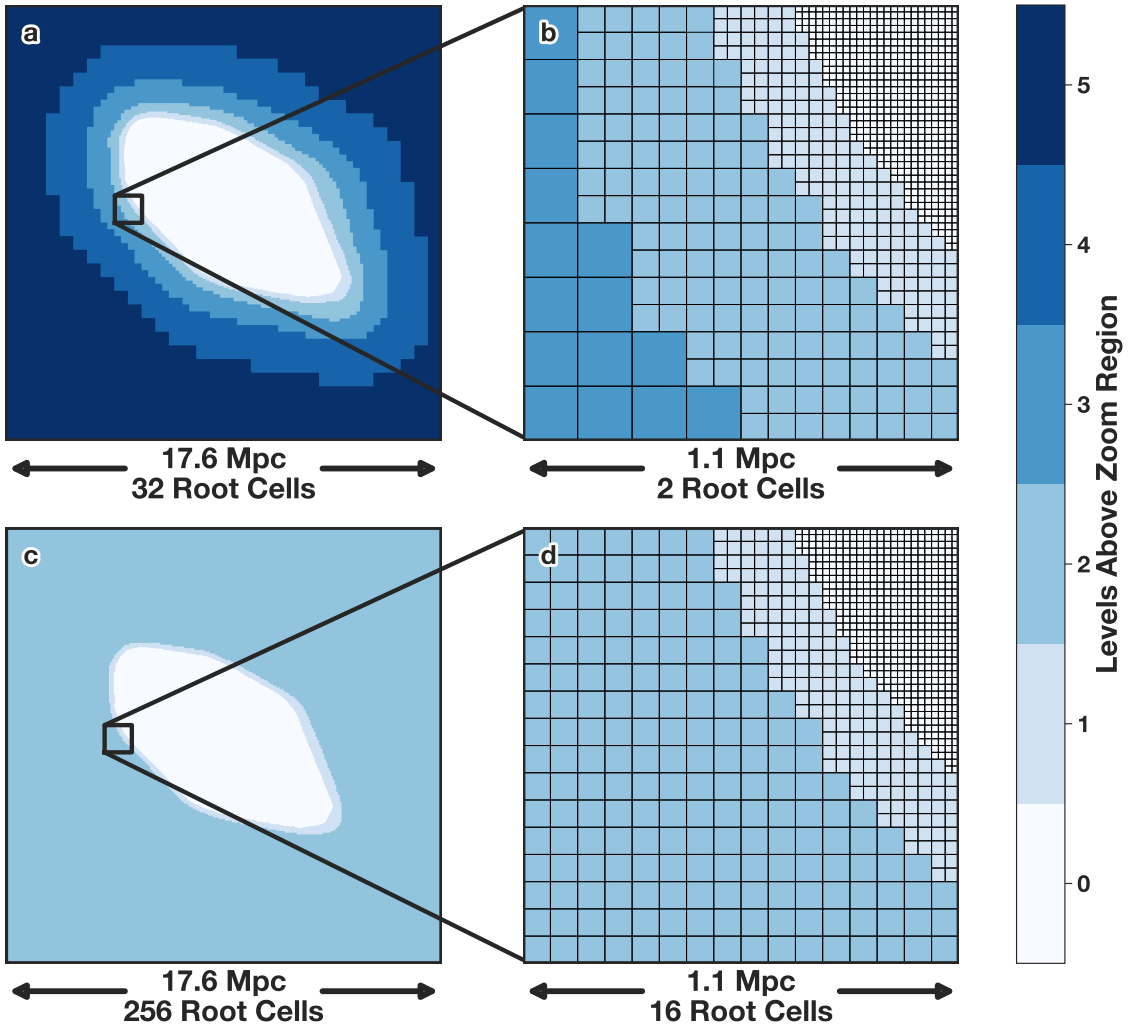

Figure 4.1 An illustration of how the grid structure changes when using our method. Panel **(a)** shows the level of refinement in the original IC. The maximum zoom level contains all the particles that end up in the main galaxies at $z = 0$. Note that this is an inset of the original box, which is 70.4 Mpc across. Panel **(b)** shows a further inset, with cell boundaries outlined. Panels **(c)** and **(d)** are the equivalents from the modified IC with a box length of 17.6 Mpc and a 256^3 root grid. Note how the original root grid and lower-resolution intermediate levels are now replaced by a new higher-resolution root grid, while the original zoom region and high resolution intermediate levels are preserved.

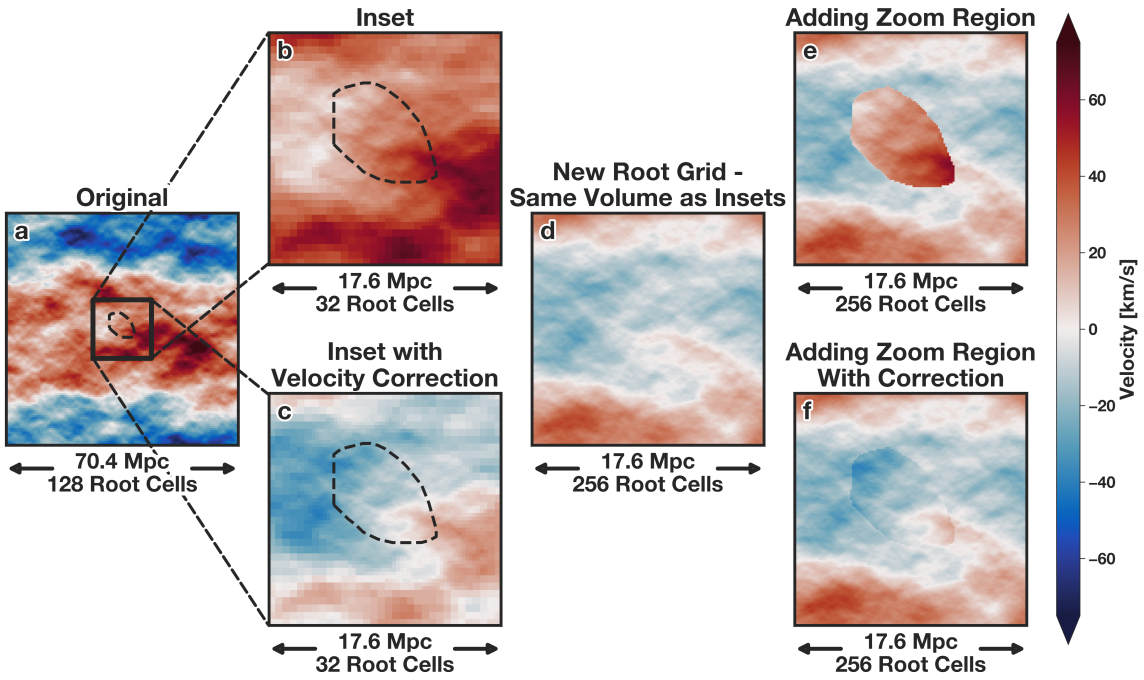

Figure 4.2 An illustration of the process of customizing ICs. All panels show the z-axis component of the gas velocity in the same slice through the simulation volume, with a dashed black line outlining the zoom region. Panel **(a)** shows the original realization, while panel **(b)** shows an inset of the region that will become the new full simulation volume. Note the higher resolution cells in the central refined region, with low resolution cells in the outskirts. This region has a nonzero bulk velocity, so we subtract it off to obtain panel **(c)**. Panel **(d)** shows the new root grid constructed with our method, which represents the same volume of space as the insets **(b, c)**. A comparison with panel **(c)** shows that many of the key features are present in both panels, with differences near the edges. These are required for periodic boundary conditions. Note that in panel **(d)** the velocity at the top boundary matches the bottom, and left matches right, unlike panel **(c)**. The higher resolution root grid cells are also apparent in the new root grid, as the outskirts do not have the pixelation present in panels **(b)** and **(c)**. Panels **(e)** and **(f)** show the result of combining the new root grid with the zoom region from panels **(b)** and **(c)** respectively. The velocity correction is necessary to give a smooth velocity field. Panel **(f)** is used as our final IC.

Our goal is to customize the root grid while preserving the environment of the zoom-in region. To achieve this, we create a hybrid IC, where we produce a new root grid that uses a smaller box size with a higher resolution grid, then embed the zoom region from the original IC. Keeping the original zoom region intact ensures that the galaxies in this region are minimally disturbed.

Creating the new root grid starts by taking the white noise field of the original cube and cutting out a smaller white noise cube. The same process (convolution with the power spectrum and gravity calculations) is used to create a new IC from this trimmed white noise cube. As the white noise cube is what seeds the resulting structures, they will be nearly preserved in the new box. The only changes will be due to the enforcement of periodic boundary conditions by the Fast Fourier Transform. This process results in root grid cells and particles with properties nearly identical to those in the original, with only slight modifications near the boundaries (far from the region of interest) to ensure periodicity. The resulting root grid can be regenerated at any desired resolution. This allows us to have significantly smaller root grid cells, potentially improving performance in codes like ART that use root grid cells for load balancing.

After creating the new root grid, we combine it with the particles and gas cells from the original zoom region. We start by keeping all the particles and cells from the zoom region of the original IC, as we want to minimally disturb the zoom region. We also keep the particles and cells from the intermediate levels that are at the same level or deeper than the new root grid. We then go through all the new root grid particles and remove any that are in the same root grid cell as a previously kept particle. Figure 4.1 shows an example of the new grid structure.

Importantly, we also need to correct the velocities of the zoom particles. The new root grid will have a mean velocity of zero by construction, as the simulation box is not moving in any particular direction, while the equivalent section of the original

Table 4.1 List of initial conditions used to run dark-matter-only simulations. Note that all realizations have the same zoom region that includes 47,517,792 particles, which dominates the total count.

Run	Box Length [Mpc]	Number of Root Grid Cells	Root Grid Cell Size [kpc]	Number of Particles
L1-128	70.4 (Original)	128^3	550.0	51,196,181
L2-256	35.2	256^3	137.5	65,589,360
L4-128	17.6	128^3	137.5	50,908,382
L4-256	17.6	256^3	68.75	64,538,268

volume may have some bulk velocity. We calculate the mean velocity of the root grid particles that are replaced, and correct the mean velocity of the zoom particles to match. This important correction minimizes the discontinuity between the zoom region and the root grid and is illustrated in Figure 4.2.

This simple method, while designed to minimize the discontinuity between the zoom region and root grid, nevertheless cannot fully eliminate the discontinuity. However, this approach allows us to preserve the properties of the zoom region as much as possible (other than its bulk velocity). It results in $z = 0$ galaxy properties most similar to the original. In addition, since there are intermediate levels between the maximally refined zoom region and the root grid, the discontinuities will be far from the galaxies of interest. As the zoom region is typically selected to include all particles that end up within the virial radius of the main galaxies, the particles far outside do not play a strong role in the evolution of the galaxies. This minimizes the affect of velocity discontinuity.

We illustrate our method by generating several modified versions of the Thelma & Louise IC. The original IC is a 70.4 comoving Mpc box with a 128^3 root grid (Garrison-Kimmel et al., 2014). A non-spherical zoom region is approximately 10 comoving Mpc across. This zoom region is 5 levels below the root grid, with intermediate levels nested around the zoom region (see Figure 4.1 for the grid structure in this IC). We create

modified versions where we decrease the box size by factors of 2 and 4. A smaller box should improve the simulation runtime, but may cause too much disruption to the regions near the galaxies of interest, affecting their evolution (Neyrinck et al., 2004). We also use two options for the number of root grid cells, as it is a primary driver of ART's load balancing. Table 4.1 details the suite of ICs we generated, including abbreviated names for each realization, which we use throughout the rest of this paper.

4.3 Comparison of Galaxy Properties

To verify the success of this method, we now evaluate how well the properties of the galaxies at the present time are preserved. To do this, we use all the ICs detailed in Table 4.1 to run dark-matter-onlysimulations. We use the `ROCKSTAR` halo finder (Behroozi et al., 2013a) and the `Consistent Trees` code (Behroozi et al., 2013b) to generate halo catalogs and merger trees.

As a visual representation of the simulations, Figure 4.3 shows the projected dark matter density of the region surrounding the two main galaxies. Their halos are clearly visible in similar positions, with similar large scale filamentary structures.

As the original IC was chosen because of its close match to the observed present-day mass of the Milky Way and Andromeda galaxies, our simulations using modified ICs must reproduce these final masses and growth histories. Figure 4.4 shows the mass assembly history for the analogs of the Milky Way and Andromeda for different simulation runs. The Milky Way analog has a smooth growth history with few mergers (Hammer et al., 2007), while the Andromeda analog has a more violent accretion history (D'Souza & Bell, 2018), qualitatively matching observations and making this IC a good representation of the Local Group. The final halo masses are reproduced well in all simulations. The $z = 0$ halo masses for the Andromeda analog are all within 15% or the original, while the largest difference in the Milky Way analog is

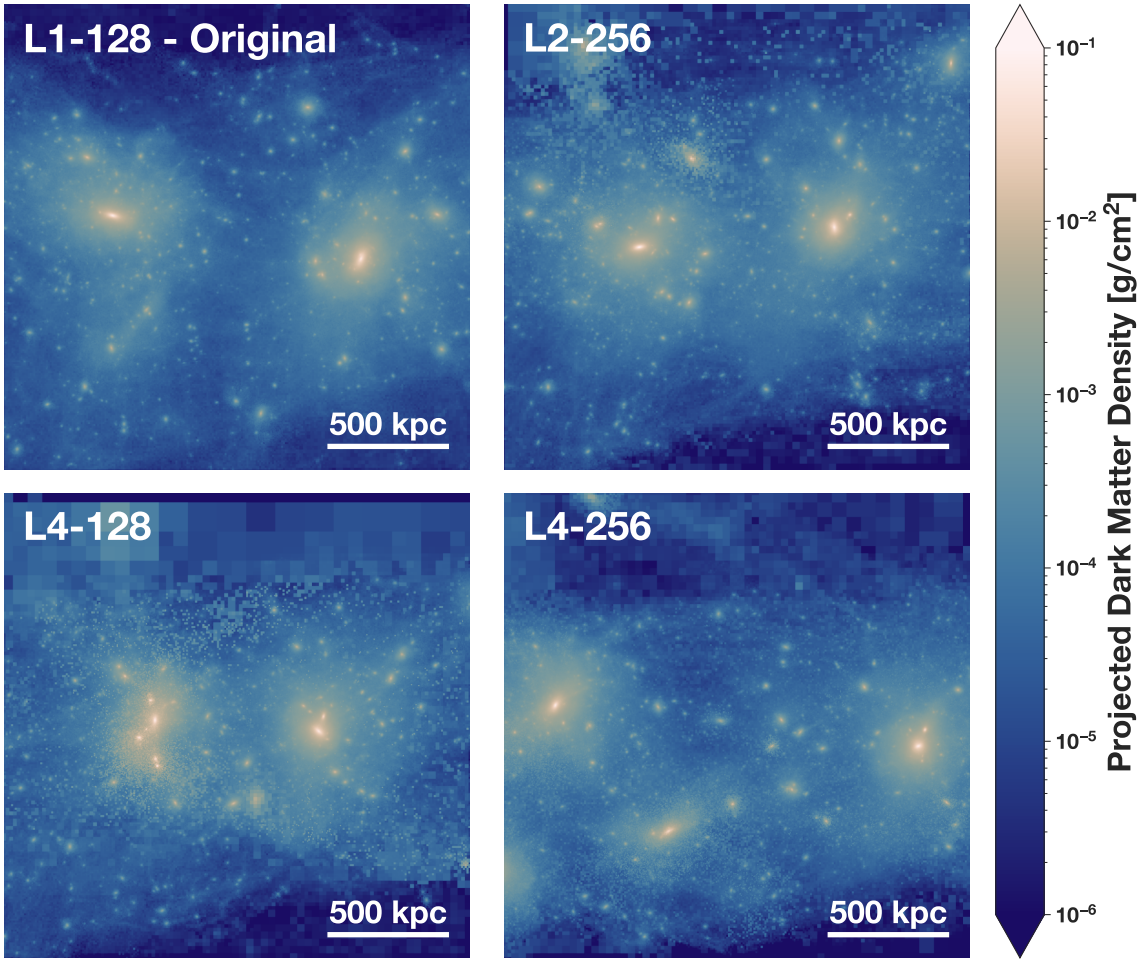

Figure 4.3 Projected dark matter density of the region surrounding the Local Group analogs at $z = 0$ in our dark matter only simulation runs.

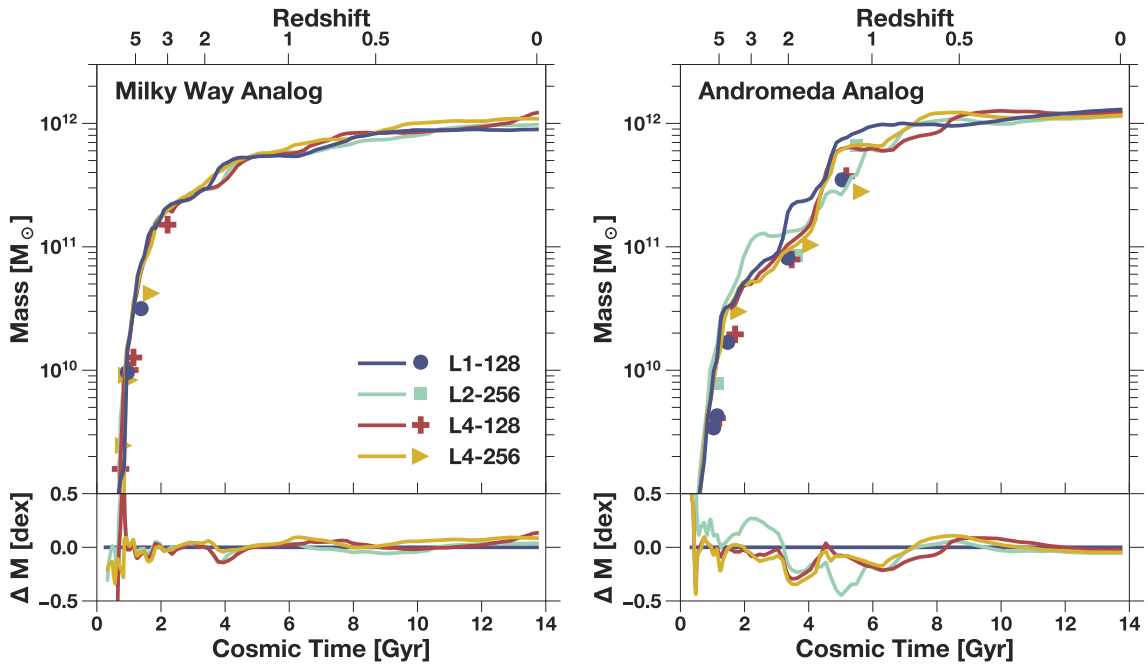

Figure 4.4 The growth of the dark matter halos of the two most massive galaxies in the Local Group. Symbols show when significant galaxy mergers occur, and are placed at the time of the merger and the mass of the infalling satellite. The bottom panels show the differences in mass as a function of time. Note the excellent agreement at all times in the Milky Way analog, and at late times in the Andromeda analog. Larger differences can occur near the times of mergers, but the final agreement is excellent.

30%, due to late growth in the L4-128 run. These halos are all still comparable to the real Milky Way and M31, making them useful for studies of these galaxies.

While galactic mergers are extremely common in the first few billion years of the universe's history, later mergers can make a significant impact on galactic stellar content, and so are more important to reproduce. The simulated Milky Way has no significant mergers after about 2 Gyr in any realization. The Andromeda analog does have a few major mergers, which are reproduced well. Note the clumps of points at around 4 and 6 Gyr in Figure 4.4. These are two mergers that occur at roughly the same time in all simulations.

The merger histories of the two galaxies are also reflected in their growth histories. The Milky Way analog's quieter accretion history results in smooth growth that is

reproduced extremely well, with typical deviations of less than 10%. Some larger deviations are driven by differences in timing, for example the more rapid growth at $z \approx 2$ that occurs later in the L4-128 run than in the original. The Andromeda analog shows larger variations, driven primarily by its more violent accretion history. Large differences in mass (up to nearly a factor of 2) can occur around the times of these major mergers, primarily driven in the differences in timing of rapid growth events. Once the mergers are completed, though, the agreement is excellent, with typical deviations of 10% after $z = 0.5$.

In addition to the central Milky Way and Andromeda galaxies, each has many smaller satellite galaxies. We examine the properties of all dark matter subhalos within 200 kpc (approximately the virial radius) of the two main halos. The left panels of Figure 4.5 show the distribution in the maximum circular velocities ($v_{\rm max}$) of the satellites. This quantity is calculated by `ROCKSTAR` using the dark matter particles identified as belonging to a given satellite. This makes $v_{\rm max}$ subject to numerical discreteness, making exact replication difficult. While the distributions follow similar shapes, the normalization can differ between different runs. One other noteworthy difference is the mass of the most massive satellite, where most modified realizations have a largest satellite significantly more massive than the original run. Interestingly, this larger circular velocity agrees better with the measured circular velocities of the Large Magellanic Cloud around the Milky Way (van der Marel & Kallivayalil, 2014) and M33 around Andromeda (Corbelli, 2003).

The positions of satellites relative to the central galaxies are also important. As the trajectory of a particle is a solution to systems of partial differential equations, small differences separate exponentially over time. This leads to chaotic non-linear orbital dynamics, making it difficult to obtain exactly the same positions relative to the galaxy center. In the Aquarius suite of simulations, positions of satellites were reproduced very well when changing the resolution of the ICs, showing that

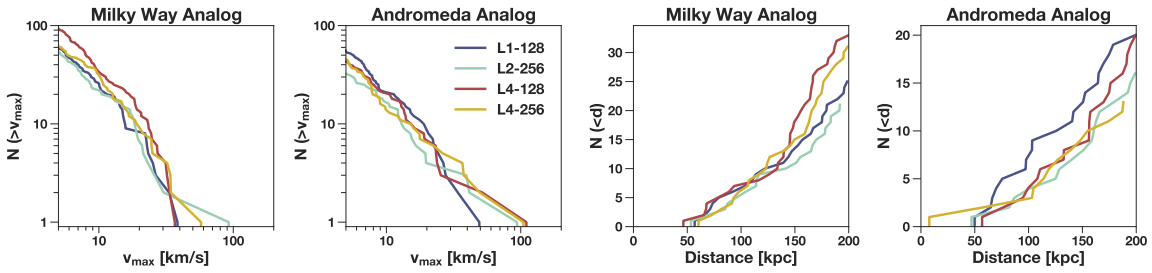

Figure 4.5 Properties of the satellite galaxies within 200 kpc of the two central galaxies at $z = 0$. The left two panels show the cumulative distribution of the maximum circular velocity, while the right two panels show the cumulative radial distribution of satellites with maximum circular velocity above 10 km s^{-1} (most likely to produce observable stars).

positions can be preserved after small changes to the ICs (Springel et al., 2008). Unlike Aquarius, our realizations change the large scale structures around the Local Group analog, resulting in changes to the positions of the main galaxies, as well as significant changes to the positions of the satellites around each central (see the visual differences apparent in Figure 4.3). However, we find that the radial distributions of satellites are similar. The right panels of Figure 4.5 show the cumulative radial distribution of satellite galaxies of the Milky Way and Andromeda analogs with $v_{\rm max} > 10$ km/s (which selects halos likely to produce observable stars). For the Milky Way analog, the agreement is excellent. The radial distributions of satellites agree within 150 kpc. Beyond that, the modified realizations start to diverge, reflecting differences in the total number of satellites with $v_{\rm max} > 10$ km/s. In the Andromeda analog, while the original realization does have significantly more satellites closer in, the shapes of the distributions are quite similar, with differences in normalization again due to differences in the total number of satellites. Note that in this comparison we only considered satellites within 200 kpc, and for the radial distribution plots chose a cut at 10 km/s in circular velocity. Changing these cuts would change the details of all distributions, but does not change the qualitative agreement present in all panels.

As our method subtly modifies the large scale structure around the zoom region, it may change which particles end up in the final $z = 0$ halos. While contamination

from higher mass particles was not identically zero in the runs using modified ICs, in most runs it remained insignificant. Figure 4.6 shows the fraction of mass within the virial radii of the two central galaxies contributed by particles other than the zoom particles. We found higher contamination in our smallest boxes, indicating that moving the periodic boundary conditions closer to the target galaxy modified the structure more strongly. In the L4-128 run, a satellite made entirely of root grid particles flew by both two central galaxies at $z \approx 0.1$, while in the L4-256 run a smaller root grid satellite came near the Milky Way analog. The L2-256 run had less contamination, with no infalling satellites made of root grid particles. However, about 1% of the virial mass of the Milky Way analog came from intermediate level particles. By identifying the contaminating particles in all runs and tracking them back to the IC, we determined that all contaminating particles came from just outside the zoom region. Regenerating the zoom region within the original IC to include the region where the contamination originated, then embedding this larger zoom region within the new root grid would solve this issue.

Similarly, one possible modification to the method presented here would be to use the trimmed white noise cube to generate a new zoom region in addition to the root grid, rather than copying the zoom region from the original IC. This would eliminate the discontinuity between the zoom region and the root grid cells, but would also likely lead to more significant changes in the galaxy properties. The hybrid IC method results in minor changes to the halo growth history and satellite distributions without changing the zoom region of the IC at all. Regenerating the zoom region would likely lead to much larger changes in galaxy properties. A major reason to use a hybrid IC is to facilitate code comparison with other groups using the same IC, and large changes in galaxy properties would eliminate this benefit.

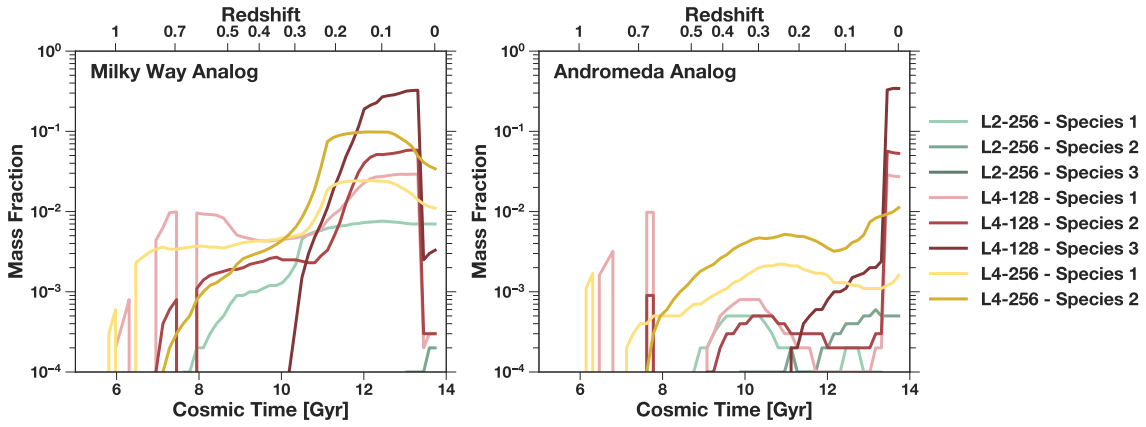

Figure 4.6 The fraction of mass within the virial radii of the two most massive galaxies contributed by particles other than the zoom particles. Not shown is Species 0, the most highly refined zoom particles, which constitute most of the mass. Higher species number corresponds to higher particle mass. The dramatic contamination in L4-128 run comes from several satellites that fly by the Milky Way analog to the Andromeda analog.

4.4 Code Speedup

To test the code speedup provided by this method, we ran full hydrodynamic simulations of each realization of our IC using the latest version of the ART code. The code utilizes adaptive mesh refinement to reach high spatial resolution. It includes radiative transfer of ionizing and UV radiation from both stars and the extragalactic background. Radiative transfer is calculated using an improved version of the Optically Thin Variable Eddington Tensor method that minimizes numerical diffusion (Gnedin, 2014). The ART code includes a non-equilibrium chemistry network that calculates the abundances of all species of hydrogen (H I, H II, H_2) and helium (He I, He II, He III), calibrated using observations in nearby galaxies (Gnedin & Kravtsov, 2011). A subgrid-scale model for numerically unresolved turbulence (Semenov et al., 2016) follows turbulent motions in the interstellar medium generated by stellar feedback. The most novel aspect of these simulations is the time-resolved modeling of star cluster formation (Li et al., 2017, 2018). Growth of star clusters is terminated by their own feedback, allowing a self-consistent calculation of cluster masses. The re-

Table 4.2 The runtime of full hydro simulations to $z \approx 10$ for different realizations of our IC using the ART code. These simulations were all done on Stampede2 with 8 Skylake nodes with identical setups (other than the IC). Here speedup is defined as the walltime of the original divided by the walltime of the other run.

Run	Time [hours] to $z \approx 10$	Speedup	Number of Cells at $z \approx 10$
L1-128	80.0	—	217,326,943
L2-256	57.5	1.39	219,741,152
L4-128	43.3	1.85	191,168,251
L4-256	43.6	1.83	204,306,243

sulting mass function of modeled clusters matches observations of young star clusters in nearby galaxies.

All of these simulations were run on the Stampede2 cluster at the Texas Advanced Computing Center using 8 Skylake nodes. The code version and setup of all runs were identical. The dark matter particle mass in the refined region is $1.57 \times 10^5\,\mathrm{M}_\odot$, with 47.5 million particles in this region. Within the refined region, we use adaptive mesh refinement to reach spatial resolution of ≈ 5 pc (physical, not comoving), which is high enough to resolve giant molecular clouds. We refine cells if either their gas mass or dark matter mass exceeds a given threshold, which acts to keep all cells with roughly the same mass. Additionally, we use a Jeans length criterion, requiring at least 3 cells to resolve the local Jeans length. At $z \approx 10$, most of the cells in the refined region are at the original level or with one additional level of refinement, but densest regions within galaxies reach the maximum resolution of ≈ 5 pc.

Table 4.2 shows the walltime it took each of these simulations to reach $z \approx 10$. All of the modified versions show substantial improvement. The L2-256 run is nearly 1.4 times faster than the original, while both L4 runs are more than 1.8 times faster. To determine the relative influence of different features of these realizations, we examine the particle number, cell number at $z \approx 10$, and the root grid cell size. While the cell number initially equals the particle number, runtime adaptive mesh refinement can

be affected by slightly different growth of structure and stellar feedback.

First, the L2-256 and L4-128 runs have the same root grid cell size, but different particle and cell numbers. The L2-256 run took 1.33 times longer than the L4-128 run, consistent with the higher number of particles (1.28 times) and cells (1.15 times) in L2-256 run. This indicates that the walltime scales roughly with the particle and cell numbers, as expected. Second, we examine the effect of the root grid cell size. The L2-256 run has more cells and particles than the L1-128 run, but also 4 times smaller root grid cell size. It ran 1.4 times faster than the L1-128 run, indicating that smaller cell size did improve computational performance. However, smaller root grid cell sizes do not automatically lead to performance gains. In the L4-256 run the root grid cell size is half that of the L4-128 run, but the finer root grid results in more particles and cells, causing both runs to take nearly the same amount of time. Together, these results indicate that for the ART code, smaller root grid cell size does improve performance, but the increased particle and cell numbers required by finer root grids can negate some of these gains. We expect that load balancing gains enabled by the finer grid would become more important at later cosmic epochs.

4.5 Summary

We have presented a method for customizing the root grid surrounding the region of interest in a zoom-in cosmological simulation. We use the original white noise cube to produce structures that remain similar to the ones in the original IC, while allowing for customized box size and root grid. The modified root grid is combined with the original zoom region to produce an IC that minimally disturbs the galaxies in the zoom region. This method results in galaxies with similar properties to those in an unmodified simulation. The customization of the root grid can be tuned to maximize computational performance for a given code.

By reducing the cost of cosmological zoom-in simulations, this method will allow

for more groups to run simulations using a common well-tested IC. This will enable more detailed code comparisons, as the confounding factor of groups using different ICs will be removed, and will allow the community to determine which aspects of galaxy formation are modeled most robustly.

CHAPTER V

Testing Feedback from Star Clusters in Simulations of the Milky Way Formation

This paper has been submitted to MNRAS with authors ***Brown, G.****, & Gnedin, O. Y. As of March 28, 2022, it is awaiting the referee report.*

Abstract

We present a suite of galaxy formation simulations that directly model star cluster formation and disruption. Starting from a model previously developed by our group, here we introduce several improvements to the prescriptions for cluster formation and feedback, then test these updates using a large suite of cosmological simulations of Milky Way mass galaxies. We perform a differential analysis with the goal of understanding how each of the updates affects star cluster populations. Two key parameters are the momentum boost of supernova feedback $f_{\rm boost}$ and star formation efficiency per freefall time $\epsilon_{\rm ff}$. We find that $f_{\rm boost}$ has a strong influence on the galactic star formation rate, with higher values leading to less star formation. The efficiency $\epsilon_{\rm ff}$ does not have a significant impact on the global star formation rate, but dramatically changes cluster properties, with increasing $\epsilon_{\rm ff}$ leading to a higher maximum cluster mass, shorter age spread of stars within clusters, and higher integrated star

formation efficiencies. We also explore the redshift evolution of the observable cluster mass function, finding that most massive clusters have formed at high redshift $z > 4$. Extrapolation of cluster disruption to $z = 0$ produces good agreement with both the Galactic globular cluster mass function and age-metallicity relation. Our results emphasize the importance of using small-scale properties of galaxies to calibrate subgrid models of star cluster formation and feedback.

5.1 Introduction

Most stars form in clustered environments (Lada & Lada, 2003), and young massive clusters (YMCs) are found in the Milky Way and other star-forming galaxies. The properties of young clusters are sensitive tracers of the star formation process. Young clusters show a well-defined mass function typically described as a Schechter (1976) function with a low-mass slope of -2 (Bastian, 2008; Portegies Zwart et al., 2010). The cutoff mass scales with the star formation rate of the host galaxy, as does the maximum cluster mass (Larsen, 2002).

Globular clusters (GCs) are also ubiquitous within galaxies, as they are found in all nearby galaxies with stellar masses above $10^9\,\mathrm{M}_\odot$ (Brodie & Strader, 2006). GCs are typically old, with ages above 10 Gyr (Puzia et al., 2005; Strader et al., 2005), and have sizes of a few parsecs that are consistent with YMCs found in the local universe (Brown & Gnedin, 2021b). This naturally leads to the hypothesis that GCs are the surviving subset of a larger population of YMCs that formed at high redshift. However, the mass function of GCs is well characterized by a lognormal distribution with a peak mass of around $2 \times 10^5\,\mathrm{M}_\odot$ (Harris, 1991; Jordán et al., 2007), in contrast to the Schechter (1976) function commonly used to describe YMCs. This transformation of the mass function over cosmic time requires a preferential destruction of low-mass clusters (Fall & Zhang, 2001; Vesperini et al., 2003; Prieto & Gnedin, 2008; Elmegreen, 2010; Kruijssen, 2015).

The dynamical evolution of clusters results from a combination of stellar evolution, two-body relaxation, tidal truncation, and tidal shocks (Spitzer, 1958; Elmegreen, 2010; Gnedin & Ostriker, 1997; Gieles & Renaud, 2016). While stellar evolution and internal two-body relaxation can be well understood by studying isolated clusters, dynamical evolution depends on the tidal field and requires more detailed modeling. Throughout their lifetime, clusters experience tidal interactions with their natal giant molecular clouds (GMCs), the galactic structure, and other GMCs, leading to complex evolution that is not well-approximated by idealized models (Meng & Gnedin, 2022).

Simulations of galaxy formation are well-suited for a detailed investigation of cluster formation and disruption (Renaud et al., 2017; Pfeffer et al., 2018). By situating clusters within their galactic context, their formation and evolution can be realistically tracked. However, few cosmological simulations have the resolution required to directly resolve cluster formation or disruption, so they must rely on subgrid models (although see Kim et al. 2018; Lahén et al. 2019; Benincasa et al. 2020; Ma et al. 2020; Hislop et al. 2022). As cluster formation is terminated by feedback from the newly-formed stars, simulations must self-consistently determine this feedback to obtain reliable properties of star clusters. Prescriptions for stellar feedback, particularly supernova feedback, have undergone many revisions over the years as they are calibrated against observations (Katz, 1992; Springel & Hernquist, 2003; Stinson et al., 2006; Agertz et al., 2011, 2013; Hopkins et al., 2014; Keller et al., 2014; Hopkins et al., 2018). However, these feedback models are often only tested against galaxy-scale properties, such as the global star formation rate or Kennicutt-Schmidt relation (Schmidt, 1959; Kennicutt, 1998). To properly model star cluster formation, these feedback prescriptions must be calibrated on smaller scales.

In Li et al. (2017) and Li et al. (2018) (hereafter L18), our group introduced a suite of cosmological simulations that directly models star cluster formation and disruption. The high spatial resolution of these simulations (3-6 pc) allows us to resolve

GMCs where star cluster formation occurs. Star particles are seeded within GMCs and accrete material from their surroundings until feedback from the newly-formed stars stops further accretion. The final masses of these star particles are set self-consistently and represent the masses of individual star clusters. These simulations were able to reproduce many aspects of the observed young cluster populations, including the shape of the initial cluster mass function, the total mass of stars contained in bound clusters, the relationship between the maximum cluster mass and the star formation rate surface density, and the formation timescales of star clusters. Some of the central clusters in satellite galaxies have properties consistent with nuclear star clusters in the local universe (Brown et al., 2018). Star formation sites in the modeled galaxies at high redshift are temporarily contained in giant clumps, which dissolve after $\sim$100 Myr (Meng & Gnedin, 2020). In addition, these simulations resolve dense irregular structures within the thick galactic disk (Meng & Gnedin, 2021), allowing for an accurate calculation of the tidal field and therefore the tidal disruption of clusters (Li & Gnedin, 2019; Meng & Gnedin, 2022).

While these simulations advanced our modeling of star cluster formation in cosmological simulations, they still had several limitations. First, they reached only redshift $z = 1.5$. This precludes a direct comparison to the GCs of the Milky Way (MW), as the disruption up to $z = 0$ must be estimated. Second, these simulations include only one MW-mass galaxy and its satellites, decreasing the statistical power of the results and potentially making the results dependent on the specific initial condition (IC) used.

In this paper, we present the next generation of simulations based on the prescriptions of L18. These simulations use two Local Group-like ICs, with the goal of reaching $z = 0$ with four MW-mass galaxies. In Section 5.2 we describe improvements to the formation and feedback schemes, then describe the new suite of simulations. This suite includes nine runs using Local Group-like ICs and 20 using the Milky

Table 5.1 Key parameters of the star formation and feedback prescriptions with the values used in this paper.

Parameter	Value(s)
Molecular fraction threshold for cluster creation	50%
Maximum virial parameter for cluster creation	10
Density threshold for cluster creation and growth	1000 cm^{-3}
Star formation efficiency per freefall time (ϵ_{ff})	1%, 10%, 100%
Radius of GMC	5 pc
Clumping factor (C_ρ)	3, 10, 30
Stellar IMF range	0.08–50 $\mathrm{M}_\odot$
Stellar mass range for SNII	8–50 $\mathrm{M}_\odot$
Stellar mass range for HN	20–50 $\mathrm{M}_\odot$
Initial hypernova fraction ($f_{\mathrm{HN},0}$)	0%, 5%, 20%, 50%
SNII momentum boost (f_{boost})	1, 2, 3, 5
Stellar mass range for AGB	0.08–8 $\mathrm{M}_\odot$
Number of SNIa per unit stellar mass	$1.6 \times 10^{-3}\,\mathrm{M}_\odot^{-1}$

Way-like IC from L18. These runs vary a wide range of feedback and cluster formation parameters, allowing us to explore how different prescriptions affect the resulting cluster properties in Section 5.3. We perform a differential analysis, systematically exploring each of the parameters we vary. In Section 5.4 we present an application of these simulations by presenting the redshift evolution of the observable cluster mass function. We discuss remaining uncertainties and compare our results with observations in Section 5.5, then summarize our results in Section 5.6.

5.2 Simulation Code and Setup

In this section we describe the ART code and the properties of the simulations. Throughout this section we introduce several parameters of the code, which we list for convenience in Table 5.1.

5.2.1 The ART Code

For our simulations we use the Adaptive Refinement Tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999; Rudd et al., 2008; Li et al., 2017, 2018). The ART code

includes many physical processes that are important for modeling the formation of galaxies. Radiative transfer is calculated using an improved version of the Optically Thin Variable Eddington Tensor method (Gnedin & Abel, 2001), which has been revised to minimize numerical diffusion (Gnedin, 2014). Radiation from both stars and the extragalactic background (Haardt & Madau, 2001) are included. A non-equilibrium chemistry network of molecular hydrogen is used to identify star-forming regions within GMCs. It was calibrated using observations in nearby galaxies (Gnedin & Kravtsov, 2011) and updated to include line overlap in computing self-shielding of molecular hydrogen (Gnedin & Draine, 2014). This chemical network also calculates the ionization states of hydrogen and helium. This model uses the local abundance of all these species to calculate the heating and cooling functions self-consistently, without any assumptions of photoionization equilibrium or collisional equilibrium. The ART code also includes a subgrid-scale (SGS) model for numerically unresolved turbulence developed by Semenov et al. (2016), which follows the results of the MHD simulations of Padoan et al. (2012).

A particularly novel aspect of the ART code is the direct modeling of time-resolved star cluster formation (Li et al., 2017, 2018; Li & Gnedin, 2019). Star cluster particles are seeded in dense gas, and accrete gas from a surrounding region until feedback from the new cluster terminates gas accretion. This region, which we refer to as the "GMC," has a radius of 5 pc and is fixed in physical size at all cosmic epochs. With the maximum spatial resolution of our simulations being set in the range of 3-6 pc, the GMC can extend past the central cell, allowing the cluster to accrete gas from neighbor cells. Specifically, the growth rate of a given cluster is

$$\dot{M} = \frac{\epsilon_{\rm ff}}{t_{\rm ff}} \sum_{\rm cell} f_{\rm GMC}\, V_{\rm cell}\, f_{\rm H_2}\, \rho_{\rm gas} \tag{5.1}$$

where $\epsilon_{\rm ff}$ is the local star formation efficiency per freefall time $t_{\rm ff}$, $f_{\rm GMC}$ is the fraction

of cell volume $V_{\rm cell}$ included within the GMC sphere, $f_{\rm H_2}$ is the local mass fraction of molecular gas, and $\rho_{\rm gas}$ is the local total gas density. This mass growth is accumulated at each local timestep, which is typically in range of $10^2 - 10^3$ years. As long as the local gas density is above the threshold, clusters can continue accreting gas. This accretion stops either when it has accreted no material in the last 1 Myr or when it has reached an age of 15 Myr.

To avoid the spurious creation of many small clusters, we impose a threshold such that clusters must have an expected mass (defined as the initial $\dot{M}$ times the maximum allowed formation time of 15 Myr) of at least 6000 $M_\odot$. As clusters typically form over a few Myr, rather than the full 15 Myr, this results in the elimination of small clusters below about 1000 $M_\odot$.

Due to the complex dynamical evolution that occurs throughout the process of cluster formation, not all stars in a given star-forming region will be bound to the fully-formed cluster. To model this, star cluster particles include a variable tracking the fraction of mass that is gravitationally bound. This is set at cluster formation (see Section 5.2.2) and is updated as clusters undergo dynamical disruption throughout their lifetime.

5.2.2 Updates to the Cluster Formation Modeling

We implement several updates to the ART code to improve the star cluster formation algorithm. In the implementation of L18, a cluster particle is created if the gas density in a cell reaches $n_H > 1000\,{\rm cm}^{-3}$ and the local H_2 mass fraction is larger than 0.5, meaning the cell contains mostly dense molecular gas. Here we introduce an additional criterion based on the local virial parameter of the gas, intended to select gravitationally bound gas. Generally, the virial parameter is

$$\alpha_{\rm vir} = \frac{5\sigma^2 R}{3GM} \tag{5.2}$$

where σ is the local gas velocity dispersion, R is the radius of the sphere we consider, and M is the mass within this sphere. We calculate this locally in any cell meeting the other star formation criteria, assuming a sphere with a diameter equal to the size of the cell ($l = 2R$), giving

$$\alpha_{\rm vir} = \frac{5\sigma^2}{\pi G \rho_{\rm gas} l^2}. \tag{5.3}$$

We use both the turbulent velocity and sound speed when calculating the velocity dispersion ($\sigma^2 = v_{\rm turb}^2 + c_s^2$), but do not include cell-to-cell velocity differences. We require $\alpha_{\rm vir} < 10$ to seed star clusters. This threshold is near the typical value for observed GMCs in the Milky Way (Miville-Deschênes et al., 2017). Star formation is allowed on the four finest refinement levels.

We also use a new prescription for the initial bound fraction of star clusters, as determined by Li et al. (2019). These authors performed simulations of 80 isolated molecular clouds with a range of mass, size, velocity configuration, and feedback strength. After feedback terminates star formation, they calculate the integrated star formation efficiency $\epsilon_{\rm int}$, which is the fraction of the initial gas mass that formed stars, as well as the fraction of stars that are bound to the final cluster $f_{\rm bound}$. They then determine the relation between these two parameters:

$$f_{\rm bound} = \left[\mathrm{erf}\left(\sqrt{\frac{3\epsilon_{\rm int}}{\alpha_\star}} \right) - \sqrt{\frac{12\epsilon_{\rm int}}{\pi \alpha_\star}} \exp\left(-\frac{3\epsilon_{\rm int}}{\alpha_\star} \right) \right] f_{\rm sat} \tag{5.4}$$

where $\alpha_\star = 0.48$ and $f_{\rm sat} = 0.94$ are free parameters the authors fitted. Determining $\epsilon_{\rm int}$ in our simulations is not trivial. The initial gas mass when the cluster was seeded is not an accurate representation of the available gas mass, as GMCs accrete material over time. To account for this, we define $\epsilon_{\rm int}$ as the ratio of the final stellar mass to the maximum value of the stellar mass plus gas mass at any time during cluster formation:

$$\epsilon_{\rm int} = \frac{M_{\star,\rm final}}{\max\left(M_\star(t) + M_g(t)\right)} \tag{5.5}$$

We then use this directly in Equation 5.4 to calculate the initial bound fraction for each star cluster.

5.2.3 Cluster Disruption Modeling

Our model for cluster disruption is unchanged from that described in detail in Li & Gnedin (2019), but we summarize the key points here. At each global timestep of the simulation (the length of the global timestep is typically a few Myr, with a maximum of 50 Myr), we calculate the tidal tensor around all fully formed clusters using the second-order finite difference of the gravitational potential across a $3 \times 3 \times 3$ cell cube centered on the star particle. To determine cluster disruption in runtime, we calculate the three eigenvalues of the tidal tensor $\lambda_1 > \lambda_2 > \lambda_3$, which describe the strength of the tidal field in the direction of their corresponding eigenvectors. We use the maximum of the absolute value of the eigenvalues to determine the dynamical timescale within the Roche lobe of the cluster:

$$\Omega^2_{\rm tid}(t) = \frac{\lambda_m}{3} \tag{5.6}$$

where

$$\lambda_m \equiv \max_i |\lambda_i| \tag{5.7}$$

We then use it to determine the cluster disruption timescale:

$$t_{\rm tid} = 10 \text{ Gyr} \left(\frac{M(t)}{2 \times 10^5 \, \mathrm{M}_\odot} \right)^{2/3} \frac{100 \text{ Gyr}^{-1}}{\Omega_{\rm tid}(t)} \tag{5.8}$$

Finally, we use this cluster disruption timescale to decrease the mass bound to each cluster. We track it with the variable $f_{\rm dyn}$, which describes the fraction of cluster mass lost due to dynamical disruption. At the n-th global timestep of length dt_n, we

update this fraction as follows:

$$f_{\rm dyn}^{n+1} = \exp\left(-dt_n/t_{\rm tid}\right) f_{\rm dyn,}^{n} \tag{5.9}$$

We also output the full tidal tensor for each star particle at each global timestep, allowing us to postprocess star cluster disruption and explore how different prescriptions for tidal disruption, including capturing tidal shocks, may change cluster properties.

5.2.4 Updates to the Stellar Feedback Modeling

5.2.4.1 Abundances of Individual Elements

We have implemented runtime tracking of most important individual elements (C, N, O, Mg, S, Ca, Fe) and ejecta of AGB stars. This gives 10 total fields tracking chemical enrichment (C, N, O, Mg, S, Ca, Fe, $Z_{\rm SNIa}$, $Z_{\rm SNII}$, and $Z_{\rm AGB}$) in both gas and stars. These elements are some of the most abundant in the universe, have reliable yields, and enable comparisons with both gas-phase and stellar abundance measurements at a variety of redshifts. N, O, and S are commonly used to measure gas-phase metallicity (e.g. Kewley & Dopita, 2002; Maiolino & Mannucci, 2019). Fe, Mg, and Ca are commonly measured in stellar spectra, with Fe representing total metallicity and Mg and Ca being representative α elements (Gallazzi et al., 2005; Kirby et al., 2013; Hayden et al., 2015).

5.2.4.2 Discrete Supernova Events

We have updated the supernova (SN) feedback prescriptions in the ART code to include discrete SN explosions at rate calculated from the stellar lifetimes, IMF, and total stellar mass of the particle. Conceptually, we use the stellar lifetimes to calculate the mass range of stars leaving the main sequence during a given timestep, then

integrate the IMF over this range to determine the total number of stars leaving the main sequence. We explode an integer number of these as SN, leaving any fractional SN to accumulate to the next timestep. This leads to only an integer number of SN exploding in a given timestep while also appropriately conserving the total number of SN over the life of the stellar population.

We calculate the number of SN in a given timestep:

$$N_{\rm SN}(\tau)+N_{\rm SN,leftover}(\tau+dt) = M_\star(\tau)\int_{\mathcal{M}(\tau)}^{\mathcal{M}(\tau+dt)}\Phi(\mathcal{M})d\mathcal{M} + N_{\rm SN,leftover}(\tau) \quad (5.10)$$

where τ is the age of the stellar population (discussed in more detail in Section 5.2.4.7), dt is the length of the current timestep, $M_\star$ is the total mass of the cluster particle, $\mathcal{M}(\tau)$ is the mass of the star leaving the main sequence at age τ, $\Phi(\mathcal{M})$ is the IMF normalized such that $M_\star = \int \mathcal{M}\Phi(\mathcal{M})d\mathcal{M}$, and $N_{\rm SN,leftover}$ is the fractional number of SN not exploded in the previous timestep. $N_{\rm SN}$ is always an integer value, and $0 \leq N_{\rm SN,leftover} < 1$. We use a Kroupa (2001) IMF with a mass range of 0.08 to 50 $\rm M_\odot$, and use 8 $\rm M_\odot$ as the minimum mass to explode as a SN. We use the metallicity-dependent analytic stellar lifetimes from Raiteri et al. (1996).

When SN explode, we inject energy and mass into the surroundings. The mass of different elements is taken directly from the stellar yield tables of Kobayashi et al. (2006). We use the yield for a star of mass $\mathcal{M} = 0.5\,(\mathcal{M}(\tau) + \mathcal{M}(\tau+dt))$, and use the metallicity of the star particle. We linearly interpolate the yield tables in both mass and metallicity to determine the yields at arbitrary stellar masses and metallicities.

5.2.4.3 Introduction of Hypernovae

Hypernovae (HN) are SN explosions with significantly more energy than a typical SN, and may be associated with gamma ray bursts (e.g. Iwamoto et al., 1998). The

Kobayashi et al. (2006) yield tables include stellar yields and energies for HN so we include them in our feedback model. We model both the energy and yields from HN self consistently. SN with progenitor stellar masses above $20\,M_\odot$ are eligible to explode as HN. Each explosion is randomly assigned to be either HN or SN, depending on a metallicity-dependent HN fraction. We use the functional form proposed by Grimmett et al. (2020):

$$f_{\rm HN} = \max\left(f_{\rm HN,0}\exp\left(-\frac{Z}{0.001}\right), 0.001\right) \tag{5.11}$$

These authors suggest that $f_{\rm HN,0} = 0.5$, but we leave it as a free parameter to test how varying it affects galaxy properties. SN explosions always inject $E_{51} \equiv 10^{51}$ ergs of energy, while for HN we use the mass-energy relation from Kobayashi et al. (2006), where the energy ranges from 10 to 30 E_{51}, with high mass stars releasing the most energy. We linearly interpolate the energy released by HN for stellar masses between those given in Kobayashi et al. (2006). Increasing $f_{\rm HN}$ significantly changes the energy injected into the simulation. Figure 5.1 shows the cumulative energy injected from SN as a function of cluster age. Different lines show different metallicity and therefore different $f_{\rm HN}$. As HN are only active for stars with masses above $20\,M_\odot$, the difference in $f_{\rm HN}$ is apparent at early times, while at later times SN energy injection is the same. As our stellar lifetimes are metallicity-dependent, the age of the onset of SN and the age at which HN end changes as well. Of note, the Raiteri et al. (1996) lifetimes give an onset of SN in this new prescription that is always later than the constant 3 Myr onset adopted by L18.

5.2.4.4 Momentum Boost

To model SN feedback, we use the prescriptions from Martizzi et al. (2015). They used simulations of inhomogeneous turbulent medium to parametrize the partition of the SN remnant energy into the thermal, kinetic, and turbulent components. The resulting energy and momentum input depend on the ambient gas density and spatial

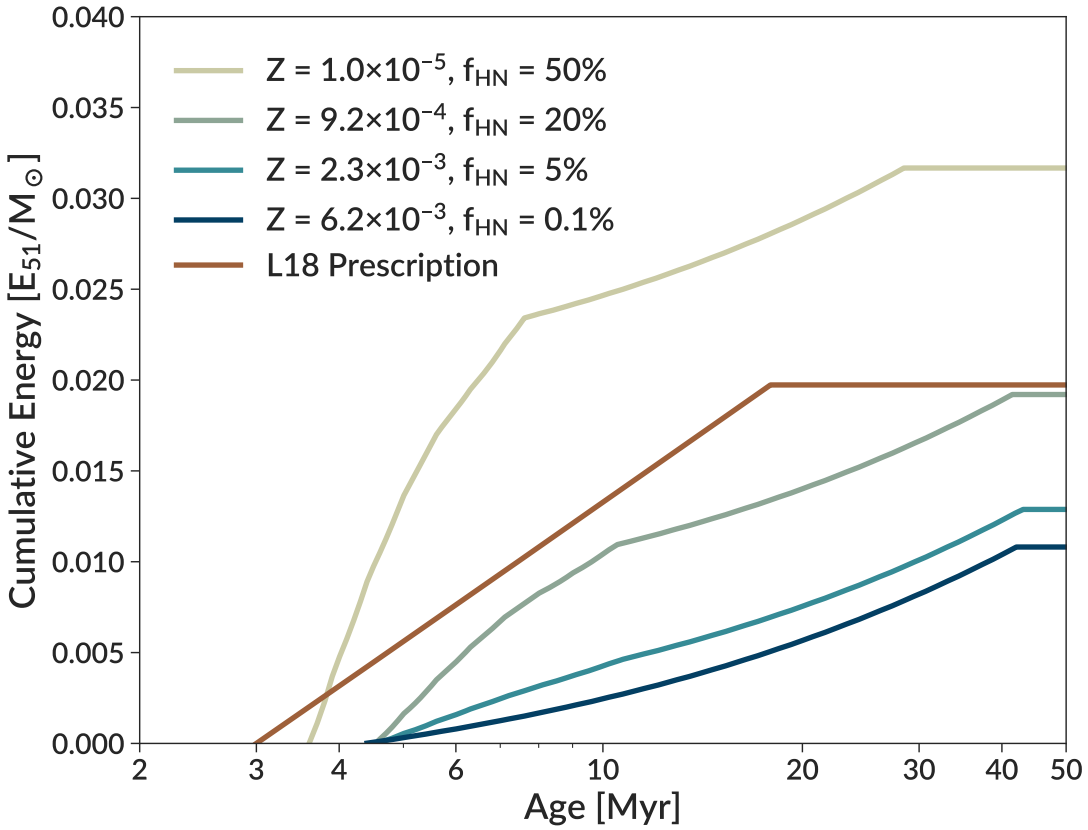

Figure 5.1 Cumulative energy injected by SN per unit stellar mass in units of 10^{51} erg $M_{\odot}^{-1}$ as a function of time since beginning of star formation for different prescriptions. Four lines show the model used in this set of simulations, while the last shows that used by L18. The new prescription is plotted at several metallicities, with HN fractions following Equation 5.11 with $f_{\rm HN,0} = 0.5$. The metallicity-dependent stellar lifetimes also change the time of the onset of SN. The line with $f_{\rm HN} = 0.1\%$ is visually indistinguishable from a line with $f_{\rm HN} = 0$.

resolution of the simulation. However, their simulations of isolated SN explosions underestimate the effect for star clusters. Cluster-forming regions usually produce a large number of massive stars that undergo simultaneous SN explosion. Gentry et al. (2017) found that such clustering of SN can enhance momentum feedback by an order of magnitude relative to that delivered by an isolated SN. L18 tested a boost to the momentum feedback from SN remnants by a factor $f_{\rm boost} = 3 - 10$ and found that the value $f_{\rm boost} = 5$ can reproduce the galactic star formation history expected from the abundance matching technique. As $f_{\rm boost}$ is a key parameter of our feedback model, we explore its ideal value in our new simulations below in Section 5.3.2. The momentum created by stellar particles is distributed spherically to 26 nearest neighbor cells surrounding the parent cell of the particle, as in Li et al. (2017).

5.2.4.5 Supernovae Type Ia

We have updated the SNIa feedback prescription, implementing discrete SN and a new delay time distribution (DTD). We use the power-law DTD for field galaxies from Maoz & Graur (2017):

$$\frac{dN_{\rm SNIa}}{dt} \propto \tau^{-1.13} \tag{5.12}$$

normalized to produce 1.6×10^{-3} SNIa per $\rm M_\odot$ of stellar mass. Similarly to how we integrate over the IMF to produce the number of SNII, we integrate over the DTD to produce the number of SNIa. We model these as discrete events as we do for SNII, and use the yields from Nomoto & Leung (2018). The feedback from SNIa is modeled simply as an injection of $2E_{51}$ of thermal energy.

5.2.4.6 AGB Feedback

Our final addition to the feedback prescription is chemical enrichment from AGB stars, defined to be the last stages of evolution of stars with masses below $8\,\rm M_\odot$. The prescription for AGB stars is analogous to that for SNII as described by Equation 5.10.

However, we abandon the requirement for integer numbers and simply use the full integral in Equation 5.10. This is justified by the fact that this phase of stellar evolution is not instantaneous like a SN. We use the yields from Ritter et al. (2018). We only inject mass from AGB feedback. We do not inject the energy or momentum, as their wind velocities are small and have little impact on the total feedback budget (Goldman et al., 2017; Hopkins et al., 2018).

We also include two other sources of feedback, which are unchanged from the implementation of L18: radiation pressure from massive stars using the analytical fit by Gnedin (2014), and momentum from stellar winds as an analytical fit to the results of Leitherer et al. (1992).

5.2.4.7 Timing of Cluster Feedback

Since our star cluster particles accrete material over time, defining a single age to use in the above feedback prescriptions is not trivial. Without storing the full cluster growth histories, which are prohibitively large, we must make some assumptions. One choice would be to simply use the time t since the star particle was seeded: $\tau_{\rm birth}(t) = t$. We refer to this as the "birth approach", since it treats all stars as forming at the same time as the first one in the cluster. This prescription is problematic if the cluster has significant star formation after the onset of SN at about 4 Myr. For example, consider some stars formed 6 Myr after the birth of that cluster particle. The birth approach assigns all stars in the cluster an age of 6 Myr, including these newly formed stars with a true age of zero. As these newly formed stars never had an age in the 0–6 Myr range, the feedback they should contribute during that age range is skipped (particularly SN feedback from 4–6 Myr). This prescription also gets the timing of feedback wrong, as the assumption that all the mass of the cluster formed at the initial time is incorrect.

An alternative is to adjust the age based on the mass-averaged time of cluster

formation: $\tau_{\rm ave}(t) = t - t_{\rm ave}(t)$. This average time for cluster formation is calculated in runtime as

$$t_{\rm ave}(t) \equiv \frac{\int_0^t t\,\dot{M}(t)dt}{\int_0^t \dot{M}(t)dt} \tag{5.13}$$

where $\dot{M}$ is the cluster star formation rate at time t (L18). This approach, which we refer to as the "average approach", does a much better job of reproducing the total amount of feedback. However, this approach pushes back the onset of SN feedback, allowing some clusters (particularly massive ones) to have unphysically long formation timescales before their growth is terminated by feedback.

To solve this problem, we introduce a hybrid approach, where we allocate a fraction of cluster feedback to use the birth approach and the rest to use the average approach. Denoting the amount of feedback generally as $\mathcal{F}$, we set

$$\mathcal{F}_{\rm tot}(t) = f_{\rm birth}(t)\ \mathcal{F}\left(\tau_{\rm birth}(t)\right) + [1 - f_{\rm birth}(t)]\mathcal{F}\left(\tau_{\rm ave}(t)\right) \tag{5.14}$$

such that $f_{\rm birth}$ is the fraction of the cluster mass assigned to the birth approach. This hybrid approach gives the best of both worlds, as it gives the correct delay before the first SN explodes while also accurately reproducing the total amount of feedback. Using idealized test cases, we find that clusters with a larger age spread require a larger $f_{\rm birth}$. Conceptually, this is because clusters with a large age spread have a larger fraction of their feedback that comes from stars formed away from the mean cluster age. We use the following parametrization:

$$f_{\rm birth}(t) = \frac{\tau_{\rm spread}(t)}{20\ {\rm Myr}} \tag{5.15}$$

where the 20 Myr scale parameter was determined from idealized test cases, and

$\tau_{\rm spread}$ is the cluster age spread calculated in runtime as

$$\tau_{\rm spread}(t) \equiv \frac{M(t)}{\langle \dot{M} \rangle} = \frac{M^2(t)}{\int_0^t \dot{M}^2(t) dt} \tag{5.16}$$

As cluster age spreads are typically a few Myr, this gives no more than 20% of the feedback coming early, with the majority using the average age. Figure 5.2 shows an example of this prescription for the feedback from a toy cluster consisting of a 4 Myr period of constant star formation rate. The delayed onset of SN when using the average approach is clear, as is the increased energy output when assuming all stars formed at the birth of the cluster. As this hybrid approach is a weighted sum of the two other approaches, there is a break in the hybrid approach between 6 and 7 Myr due to the onset of SN in the average approach.

5.2.5 Hydrodynamics

When updating the code from the version used in L18 to ART version 2.0, we updated the modeling of the energy equation that governs how thermal energy is calculated in the presence of subgrid turbulence. This update more accurately tracks thermal energy in shocks. It has little effect in the disc of the galaxy, as the thermal energy generated by shocks is subdominant to other process that govern energy balance such as heating, cooling, and stellar feedback. However, we find that the circumgalactic medium is affected by this update. In our new runs, there is significantly more hot gas in the halo. This in turn leads to less cold gas accreting onto the galaxy, leading to less star formation. The decrease in the amount of cold gas requires changes to the parameters governing star formation and feedback as we describe below. We describe the update to the hydrodynamics in more detail in Appendix B. In our suite of simulations, we used both this updated energy-based approach and the new entropy-conserving scheme of Semenov et al. (2021). These authors found that

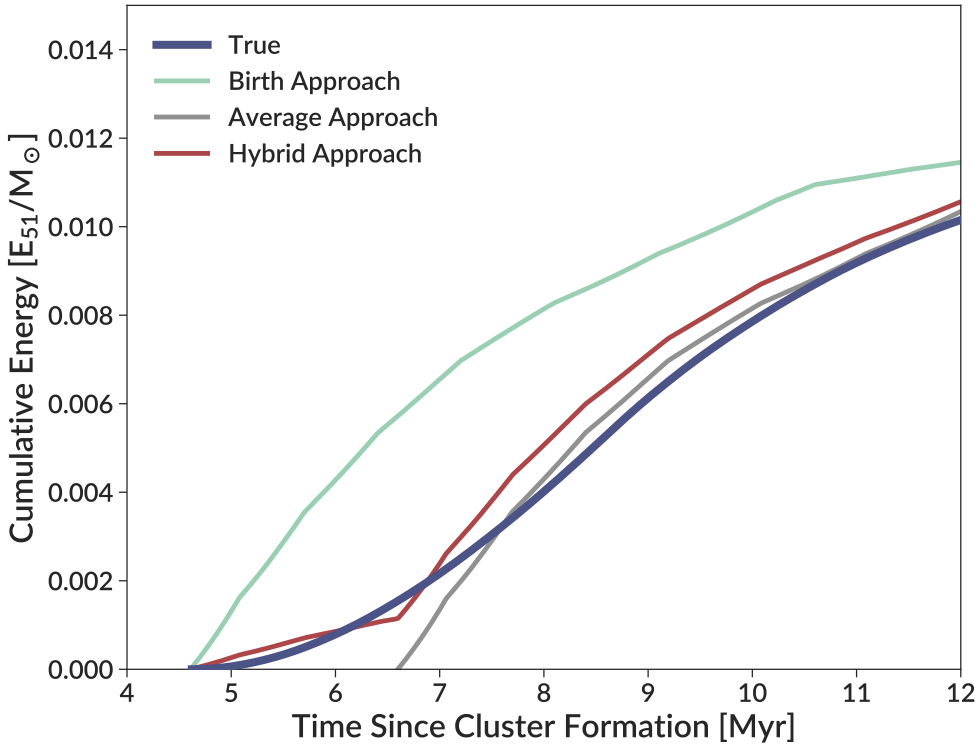

Figure 5.2 The cumulative energy injected by SN per unit stellar mass as a function of time since beginning of cluster formation for three approaches for the timing of SN. The input star formation history is a constant star formation rate for 4 Myr. The "True" line shows the actual energy injection produced by this stellar population, while the other lines show the energy injection for different ways of treating this star formation history as a simple stellar population, as described in the text. The hybrid approach is a weighted combination of the other two.

the entropy-conserving scheme is able to more accurately evolve nonthermal energy components. They ran simulations of an $L_\star$ galaxy and found differences between the energy-based and entropy-based schemes. However, these differences are much smaller than those we find between the energy-based schemes of L18 and the updated version that is the default in the ART code version 2.0.

5.2.6 Initial Conditions

In this work we use three different ICs. One is the IC used by L18, a periodic comoving box of size 4 Mpc that contains a single central galaxy with a total mass of 10^{12} $M_\odot$ at $z = 0$, which we refer to as `Isolated MW`. We also use two zoom-in ICs from the ELVIS project (Garrison-Kimmel et al., 2014): `Thelma & Louise` and `Romeo & Juliet`. Both of these ICs contain a Local Group analog with two Milky Way-mass galaxies, which we describe in more detail below. The `Isolated MW` box is much less computationally expensive to run than the zoom-in runs, so we use it to explore a broader range of parameter space.

`Thelma & Louise` is a desirable IC as it has qualitative agreement with the accretion histories of the MW and M31. The less massive (MW-like) halo has a quieter accretion history (Hammer et al., 2007), with no significant mergers after $z \approx 5$, while the more massive (M31-like) halo has more mergers at later times as expected from observations (D'Souza & Bell, 2018). `Romeo & Juliet` has two galaxies with much quieter merger histories. Including two different sets of ICs allows us to explore how our results vary with galaxy merger histories.

To improve computational performance with the ART code, we modify these zoom-in ICs following the prescription of Brown & Gnedin (2021a). Our initial conditions have a small zoom region in a large box (50-100 Mpc). This large box size with a small zoom region is difficult for the ART code to parallelize well, so our method decreases the box size and increases the resolution of the root grid. Starting from

the white-noise cube used to seed the structures present in the original IC, we cut out a smaller region of interest, then use this trimmed cube to regenerate the cosmic structure using the `MUSIC` software (Hahn & Abel, 2011). To avoid disturbing the zoom region, the particles from this region are transplanted into the new box with a velocity offset to match the systemic velocity of this region in the new box. We find that these modifications improve performance while minimally changing central galaxy properties. Table 5.2 details some key properties of these ICs, and Figure 5.3 shows the halo mass growth of these galaxies in collisionless runs.

We run an initial suite of nine simulations with these zoom-in ICs, varying $\epsilon_{\rm ff}$, $f_{\rm HN,0}$, and $f_{\rm boost}$. Table 5.3 details the parameters of these runs. We also ran a large suite of 20 simulations on the `Isolated MW` initial condition varying many aspects of stellar feedback, which are detailed in Table 5.4.

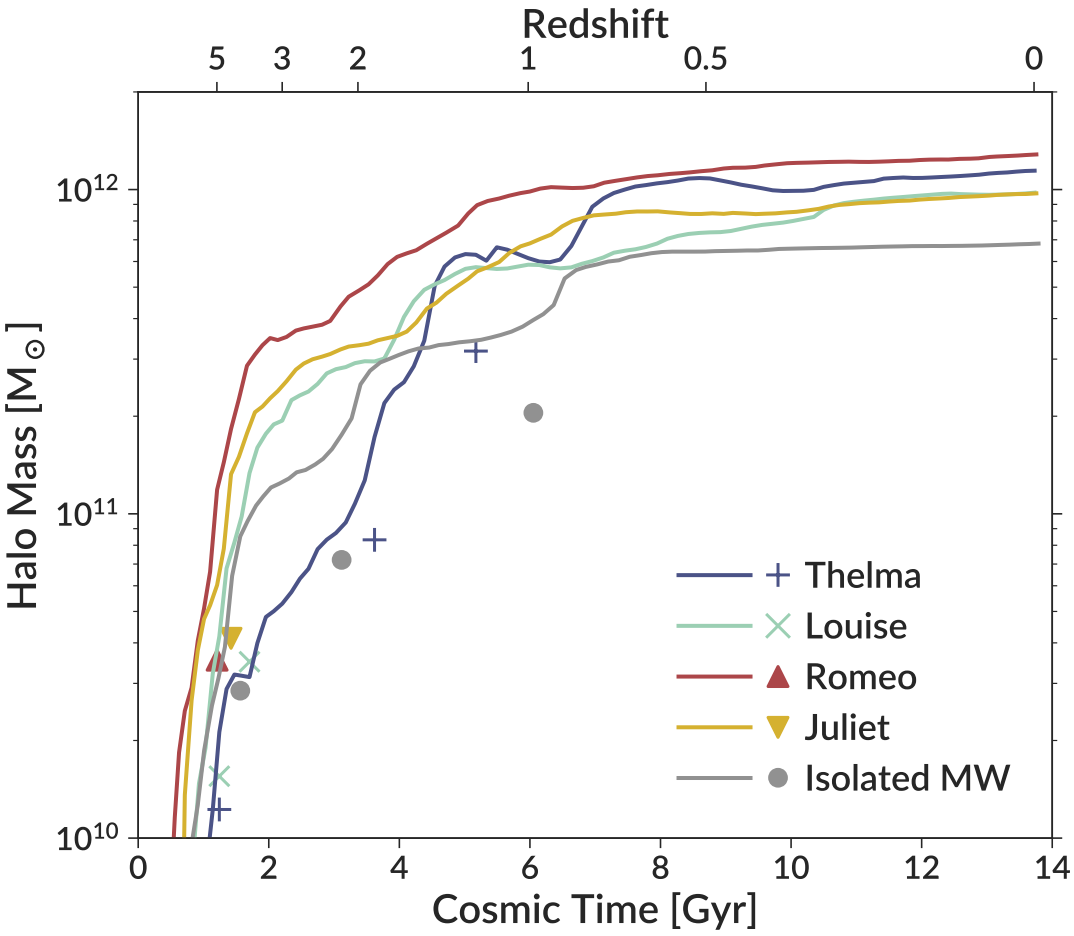

Figure 5.3 Mass growth of the central halos from collisionless runs with three initial conditions. Markers show major mergers with a mass ratio less than 4:1, and are placed at the maximum virial mass of the satellite and the time at which it reached this maximum mass before merging with the central galaxy. Note that `Thelma` and `Isolated MW` have major mergers at $z < 2$, while the other three galaxies have quiet merger histories.

Table 5.2 Description of key properties of the initial conditions used here. For zoom-in ICs, the mass resolution quantities refer to the zoom region.

Initial Condition	Box Size (comoving h^{-1} Mpc)	Dark Matter Particle Mass ($M_\odot$)	Typical Baryonic Cell Mass ($M_\odot$)	Ω_m	Ω_Λ	Ω_b	h
`Thelma & Louise`	25.0	1.57×10^5	2×10^4	0.266	0.734	0.0449	0.71
`Romeo & Juliet`	23.12	1.53×10^5	2×10^4	0.31	0.69	0.048	0.68
`Isolated MW`	4.0	1.0×10^6	4×10^4	0.304	0.696	0.0479	0.681

Table 5.3 The runs using the Local Group initial conditions included in this simulation suite. $z_{\rm last}$ is the redshift of the last output of each run. All runs use average approach for SN timing and the energy-based hydrodynamics scheme that is the default in the ART code version 2.0.

Initial Condition	$\epsilon_{\rm ff}$	$f_{\rm boost}$	$f_{\rm HN,0}$	$z_{\rm last}$
`Thelma & Louise`	1%	5	20%	3.32
`Thelma & Louise`	10%	5	20%	2.36
`Thelma & Louise`	100%	1	0%	3.17
`Thelma & Louise`	100%	3	0%	2.80
`Thelma & Louise`	100%	5	0%	1.83
`Thelma & Louise`	100%	5	5%	1.86
`Thelma & Louise`	100%	5	20%	2.66
`Romeo & Juliet`	10%	5	20%	2.78
`Romeo & Juliet`	100%	5	20%	1.87

Table 5.4 The runs using the `Isolated MW` initial conditions included in this simulation suite. In the "Hydro Scheme" column, "S21" refers to the entropy-based scheme of Semenov et al. (2021), "Energy" refers to the energy-based scheme that is the default in the ART code version 2.0, and "L18" is the hydro scheme used in L18. The schemes mentioned in the "SN Timing" column are described in Section 5.2.4.7. Simulations are roughly grouped by the attribute that is varied. All runs progressed to $z = 1.5$ except for the two runs with $\epsilon_{\rm ff} < 100\%$ and $f_{\rm boost} = 2$, which reached $z \approx 2$.

$\epsilon_{\rm ff}$	$f_{\rm boost}$	$f_{\rm HN,0}$	SN Timing	Hydro Scheme	Other Attributes
100%	5	0	Average	Energy	Analogous to Local Group runs
100%	5	0	Average	Energy	Feedback scheme of L18
100%	5	0	Average	L18	
100%	1	0	Average	S21	Continuous energy injection from SN
100%	1	0	Average	S21	No virial parameter criterion for star formation
100%	1	0	Average	S21	$C_\rho = 3$
100%	1	0	Average	S21	$C_\rho = 3$, changed shielding to Gnedin & Kravtsov (2011)
100%	1	0	Average	S21	$C_\rho = 30$
1%	1	0	Average	S21	
10%	1	0	Average	S21	
1%	2	0	Average	S21	
10%	2	0	Average	S21	
100%	1	50%	Average	S21	
100%	1	0	Average	S21	
100%	2	0	Average	S21	
100%	3	0	Average	S21	
100%	5	0	Average	S21	
100%	1	0	Hybrid	S21	
100%	1	0	Average	S21	
100%	1	0	Birth	S21	

5.2.7 Run Setup

We keep the spatial resolution of the finest grid level between 3–6 physical pc at all times. To accomplish this, we add refinement levels as the simulation progresses. The specific levels and when they are added depend on the initial condition. In the `Isolated MW` box, we start with 9 levels of refinement on the 128^3 root grid, then add levels at $z = 9$, 4, and 1.5. For `Thelma & Louise`, we allow 11 levels of refinement on the 256^3 root grid, then add additional levels at $z \approx 10.2, 4.6, 1.8$, and 0.41. `Romeo & Juliet` also starts with 11 levels, but its slightly different box size requires adding levels at $z \approx 9.8, 4.4, 1.7$, and 0.35.

We use three criteria to determine when to refine the grid. In this section we will illustrate the refinement criteria using specific values from the `Thelma & Louise` IC, but the principles are the same for all ICs. First, we use Lagrangian refinement for both gas and dark matter. Cells are refined when their gas mass exceeds approximately $1.6 \times 10^5\,\mathrm{M}_\odot$ or dark matter exceeds $3.9 \times 10^6\,\mathrm{M}_\odot$. The gas refinement is active on all levels, while the dark matter criterion is not active on the four finest levels. We also increase the dark matter mass refinement threshold above that from the simple baryon fraction scaling. These changes are for two reasons. First, the discrete dark matter particles (of mass $1.5 \times 10^5\,\mathrm{M}_\odot$) do not allow their mass to be distributed evenly, so their distribution cannot be trusted on small scales. Second, we find that there are times when the dark matter criterion will prevent a cell with very small gas mass from derefining. If stellar momentum feedback is imparted on this cell, it will acquire very high velocities due to its small mass, leading to small timesteps and a slower runtime of the simulation. Restricting the levels on which the dark matter Lagrangian criterion is active and increasing the mass threshold for dark matter-triggered refinement mitigates this situation. The final refinement criterion uses a local Jeans length. Cells are refined if their size exceeds twice the Jeans length. This criterion is applied only on the four deepest levels. We find that with these re-

finement criteria, cell gas masses remain around $2\times10^4\,M_\odot$. Table 5.2 includes the typical baryonic cell masses for all ICs.

While we do not record the level on which a star is formed in runtime, we post-process the outputs to see the levels on which stars can form. In the runs using the Local Group ICs, we find that 15% of the cells that satisfy the star formation criteria are on the highest refinement level with sizes of 3–6 pc, 60% have sizes in the 6–12 pc range, 25% are within 12–24 pc, and a very small fraction are on the fourth level with sizes of 24–48 pc. The lower mass resolution of the `Isolated MW` runs results in the corresponding fractions of 10%, 35%, 50%, and 5%, respectively.

The ART code uses adaptive time stepping, such that the finest levels have much shorter timesteps than the coarse root grid. For the `Thelma & Louise` runs with $\epsilon_{\rm ff} = 100\%$, the global timestep of the root grid is restricted to be less than 10 Myr. We write outputs at each global timestep. For all other runs, the output spacing is allowed to be at most 50 Myr. The timestep for the finest level is similar for all runs, typically between 100–1000 years.

5.3 Effects of Cluster Formation and Feedback Modeling

In this section, we analyze the large suite of simulations laid out in Tables 5.3 and 5.4 to test the implementation of code updates and explore how parameter variation affects our results. We will primarily focus on the galaxy star formation rate, cluster mass function, and the timescales of cluster formation. In this section we exclusively use the particle mass at the end of its star formation episode, which does not account for the initial bound fraction, stellar evolution, or dynamical disruption of a star cluster represented by that particle. We explore those quantities and the observable cluster mass function in Section 5.4. We also note that when examining star cluster populations, we include all clusters from the central galaxies in the simulations (the one MW-mass galaxy in `Isolated MW`, and the two galaxies in the Local Group-like

environments of `Thelma & Louise` and `Romeo & Juliet`). When plotting the star formation rate of these galaxies we plot the two central galaxies in the Local Group-like IC separately, but when plotting cluster properties of a given run we group these two galaxies together.

5.3.1 Timing of Supernova Feedback

In Section 5.2.4.7 we describe how the finite length of cluster formation makes it difficult to create an accurate prescription for the timing of stellar feedback. We ran simulations with the birth approach, the average approach, and the hybrid approach. We also compared these to the feedback model of L18, which has SNe that start earlier (see Figure 5.1). We found no significant differences in any galaxy-scale properties between these prescriptions. However, we did find that the cluster formation lifetimes were different between these prescriptions. In particular, the average approach gave significantly longer timescales for massive clusters. Figure 5.4 shows the cumulative distribution of the length of star formation within clusters formed using different timing choices, for the local efficiency $\epsilon_{\rm ff} = 100\%$. Note that the quantity we plot here is the duration of star formation, defined as the age difference between the birth of the cluster and its last accretion event. This is not $t_{\rm ave}$ or $\tau_{\rm spread}$ as defined in Section 5.2.4.7. We use this quantity as it clearly demarcates when feedback ends cluster formation.

With all the approaches to SN feedback, the majority of low-mass clusters have finished their accretion before the onset of SNe at 3–4 Myr, leading to little difference in the durations between our approaches. Such short durations indicate that the other sources of feedback are able to terminate cluster formation before the start of SN feedback (Kruijssen et al., 2019b; Grudić et al., 2022). SN feedback remains more relevant for massive clusters.

We do see a difference in the high mass clusters. The feedback prescriptions of

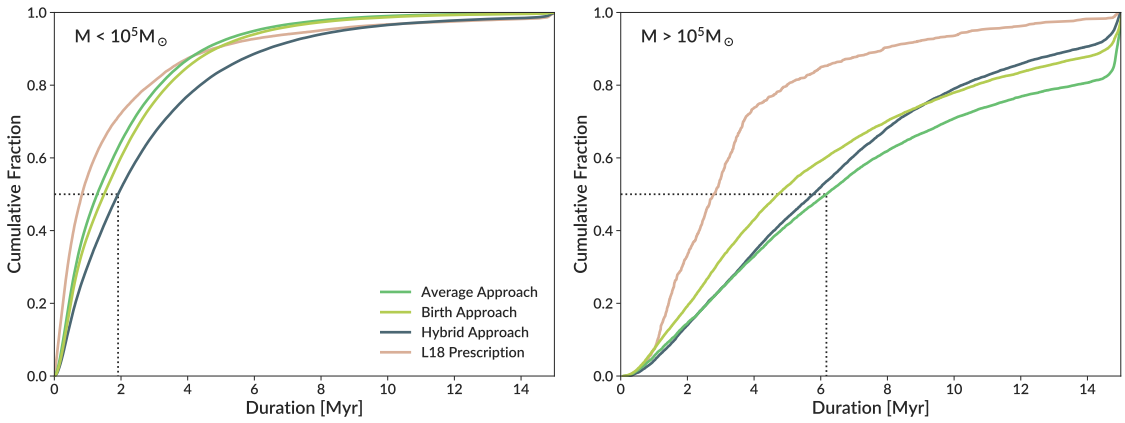

Figure 5.4 The cumulative distribution of the duration of cluster formation for different approaches to determining the timing of SN feedback, as described in Section 5.2.4.7. The left panel shows clusters less massive than $10^5\,M_\odot$, while the right panel shows clusters more massive than $10^5\,M_\odot$. The dotted line shows the longest median duration of cluster formation. Cluster growth is algorithmically truncated at 15 Myr. Note that here we use a new run with the L18 feedback model, not the L18 simulations themselves. The L18 prescription uses $f_{\rm boost} = 5$, while all other runs use $f_{\rm boost} = 1$. All runs use the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

L18 produce the shortest durations of star formation. Among the three new models of determining the timing of SN feedback, the average approach produces clusters with the longest duration, the birth approach gives clusters with the shortest duration, and the hybrid approach is in the middle. As the birth approach has the most early feedback and the average approach has the least early feedback, these results indicate that delaying the start of SN feedback tends to increase the time over which massive clusters can accrete material. This matches what we see in the L18 model, which allows SN feedback begin earlier and stop cluster growth earlier.

These trends are also reflected in the integrated star formation efficiency $\epsilon_{\rm int}$, defined in Equation 5.5. Figure 5.5 shows the distribution of $\epsilon_{\rm int}$ for the runs with variations in the timing of SN feedback. The L18 feedback model has the earliest SN feedback and the lowest mean value of $\epsilon_{\rm int}$ (15%), while the average approach has the latest SN feedback and the highest mean value of $\epsilon_{\rm int}$ (35%). Interestingly, the

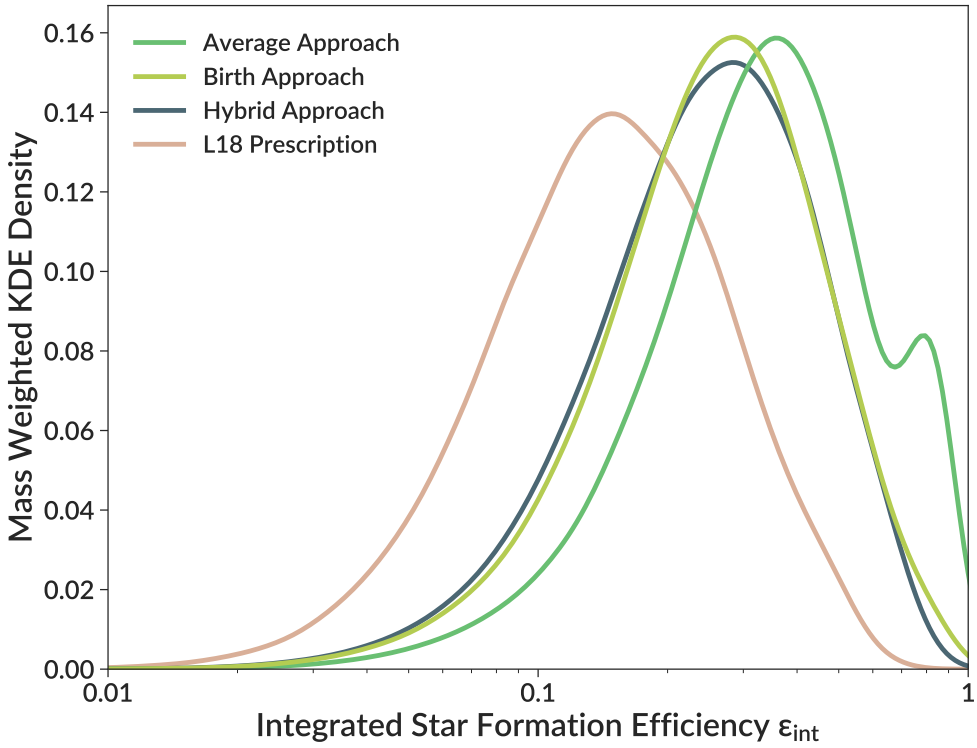

Figure 5.5 Kernel density estimation for the distribution of integrated star formation efficiency for clusters in the runs with variations in the timing of SN feedback, as described in Section 5.2.4.7. We use a Gaussian kernel with a width of 0.05 dex. Each curve is normalized to the same area for comparison purposes. Note that here we use a new run with the L18 feedback model, not the L18 simulations themselves. The L18 prescription uses $f_{\rm boost} = 5$, while all other runs use $f_{\rm boost} = 1$. All runs use the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

hybrid approach and birth approach are very similar, with mean values at $\epsilon_{\rm ff} \approx 25\%$. This may be because early SN feedback (present in both variations to some extent) is important for dispersing gas before it can be accreted by the cluster. While the simulation with the L18 model uses $f_{\rm boost} = 5$ instead of $f_{\rm boost} = 1$, other runs varying $f_{\rm boost}$ show no little difference in either the duration of star formation or $\epsilon_{\rm int}$, indicating that the SN timing is responsible.

Despite these differences in the duration of star formation and $\epsilon_{\rm int}$, we see no significant differences in the star particle mass functions.

5.3.1.1 Discreteness of Supernova

In addition to multiple runs with different prescriptions for SN feedback, we also ran one simulation with continuous energy injection from SN. This run uses the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$. The number of SN still follows the IMF integral as in Equation 5.10, but with the modification that we do not require there to be an integer number of SN in each timestep. We find that this change makes little difference to galaxy properties. The star formation rate was not affected, and neither were star cluster properties, including their mass function and age spread.

We note that the similarity between these two runs is despite real differences in how the energy is injected over time. The SN rate changes with time, but is within the range of $(2-6)\times10^{-10}\,N_{\rm SN}\,{\rm M}_\odot^{-1}\,{\rm yr}^{-1}$. Our typical timesteps on the highest refinement levels are below 10^3 yr, so even massive clusters with $M = 10^6\,{\rm M}_\odot$ do not have a SN every timestep. Clusters of mass $M = 10^3\,{\rm M}_\odot$ have only 10 SN over the ~ 40 Myr timescale for SN feedback, resulting in significant gaps between SNe. The onset of SN can also be delayed in low-mass clusters, as the decrease in the normalization of the IMF means we need to integrate to lower stellar masses to reach one star (Equation 5.10). These results indicate that the total injected energy and the timing of the onset of SN cause larger differences than does discretizing SN events.

To summarize, we find that different prescriptions to change the onset of SN (without changing the total energy injection) do not affect any galaxy-scale properties, but do affect the properties of star clusters. When SN feedback is delayed, massive clusters have longer formation timescales, and all clusters have higher $\epsilon_{\rm int}$. When comparing disretized SN to continuous energy injection, we find no significant differences.

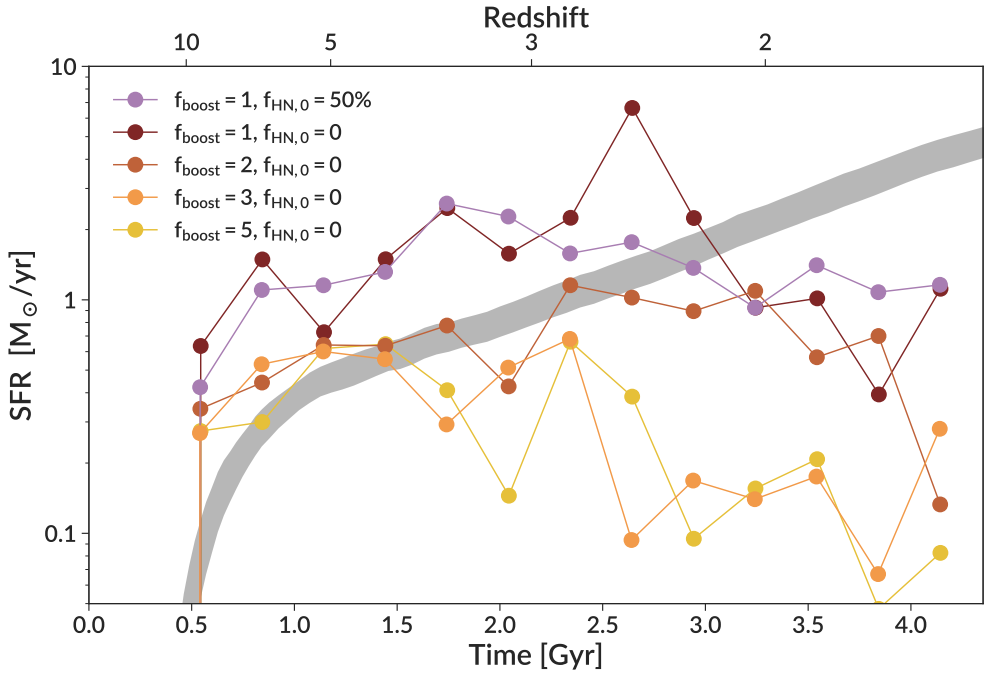

Figure 5.6 A comparison of the star formation history for the central galaxy in the `Isolated MW` IC when varying $f_{\rm boost}$ and $f_{\rm HN,0}$. The shaded region shows the expected star formation history as given by UniverseMachine. All runs use $\epsilon_{\rm ff} = 100\%$.

5.3.2 Strength of Supernova Feedback

Our simulations have two main parameters to control the strength of SN feedback: $f_{\rm boost}$ and $f_{\rm HN,0}$. In this section, we explore how those parameters affect our results.

In Figures 5.6 and 5.7 we show the impact of these two parameters on the star formation history of the main galaxies. In Figure 5.6 we show the star formation history of the single central galaxy of the `Isolated MW` IC, while in Figure 5.7 we show two lines for each run representing the two main galaxies in a Local Group-like environment. We also show the expected star formation history as given by UniverseMachine (Behroozi et al., 2019). However, we note that the MW assembly history may be atypical for halos of its mass, as both the ancient merger of Gaia-Enceladus Sausage and the current infall of the LMC influence its evolution (Evans et al., 2020).

First, we find that $f_{\rm boost}$ has a strong impact on the global star formation rate.

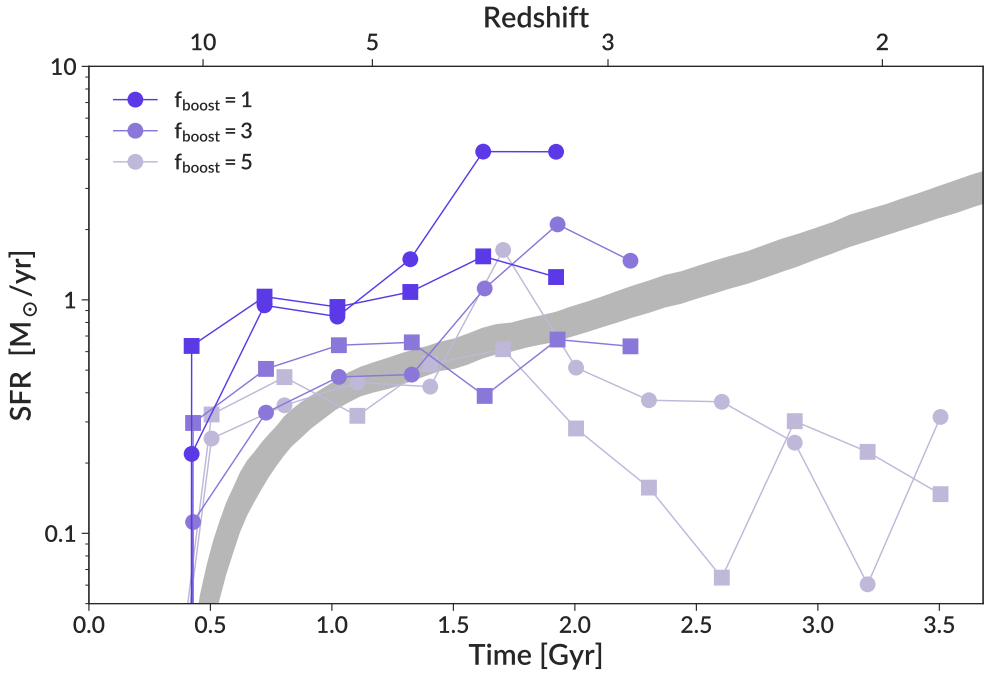

Figure 5.7 Same as Figure 5.6, but for the `Thelma & Louise` IC and only showing variations in $f_{\rm boost}$. There are two main galaxies in each run. Circles represent the MW analog, with squares representing M31. All runs use $\epsilon_{\rm ff} = 100\%$ and $f_{\rm HN,0} = 0$.

Higher values of $f_{\rm boost}$ result in generally lower star formation rates. In the `Isolated MW` runs shown in Figure 5.6, the run with $f_{\rm boost} = 5$ matches the UniverseMachine prediction well until roughly $z \approx 4$, at which point the star formation rates start to decline significantly. This is similar to what we see for the $f_{\rm boost} = 5$ run in `Thelma & Louise` in Figure 5.7. The star formation rate is reasonable until $z \approx 4$, with a significant decline afterwards. A value of $f_{\rm boost} = 3$ matches UniverseMachine more closely in both ICs, but in the `Isolated MW` IC the star formation rate drops off significantly after $z = 3$. The $f_{\rm boost} = 3$ run using `Thelma & Louise` has only progressed to $z = 2.8$ at the time of writing, so it remains possible that its star formation rate will drop as it did in the `Isolated MW` run. However, we must be careful making direct comparisons between different ICs, as it is likely that they will have different star formation histories. In particular, Santistevan et al. (2020) found that Local Group-like galaxies form earlier than isolated galaxies. They conclude that

the denser environment of Local Group-like pairs causes the initial collapse of halos to happen earlier (Gallart et al., 2015). This leads to more mass forming earlier, and this buildup of stellar mass may affect how feedback affects the galaxy at later times.

A slightly lower value of $f_{\rm boost} = 2$ matches UniverseMachine well up to $z \approx 2$ before decreasing greatly. Finally, runs with $f_{\rm boost} = 1$ have the highest levels of star formation in both initial conditions. This high level has persisted in `Thelma & Louise` until the last available output, but in `Isolated MW` the star formation rate dropped dramatically starting at $z = 2$. Even this low value of $f_{\rm boost}$ is not able to produce reasonable galactic star formation histories over the full time range spanned by these simulations.

L18 calibrated $f_{\rm boost}$ in their simulations, finding a preferred value of $f_{\rm boost} = 5$. The difference in our result is due to the changes in hydrodynamics. As described above in Section 5.2.5, that change led to a decrease in the amount of cold gas that reaches the galaxy. This requires changes to the feedback modeling to compensate. Without decreasing $f_{\rm boost}$, the galaxies have lower total gas mass and less cold gas, which leads to less molecular gas. Since molecular gas is required by our star formation prescription, this decrease leads to less star formation.

While we find that $f_{\rm boost}$ has a strong impact on the star formation rate, we find that $f_{\rm HN,0}$ does not. In Figure 5.6, runs with $f_{\rm boost} = 1$ have similar star formation histories, regardless of the value of $f_{\rm HN,0}$. While we do not show runs varying $f_{\rm HN,0}$ in Figure 5.7 for clarity, runs with $f_{\rm HN,0} = 0, 5\%$, and 20% all show similar star formation rates (all using $f_{\rm boost} = 5$).

This is likely due to the metallicity dependence of the hypernova fraction $f_{\rm HN}$ (see Equation 5.11). The value of $f_{\rm HN}$ is highest at low metallicity, but decreases rather quickly with metallicity. Figure 5.8 shows the metallicity of stars forming at different times and their $f_{\rm HN}$. This plot uses the run on the `Isolated MW` IC with maximum $f_{\rm HN,0} = 50\%$, yet the quick enrichment means that the bulk of clusters have

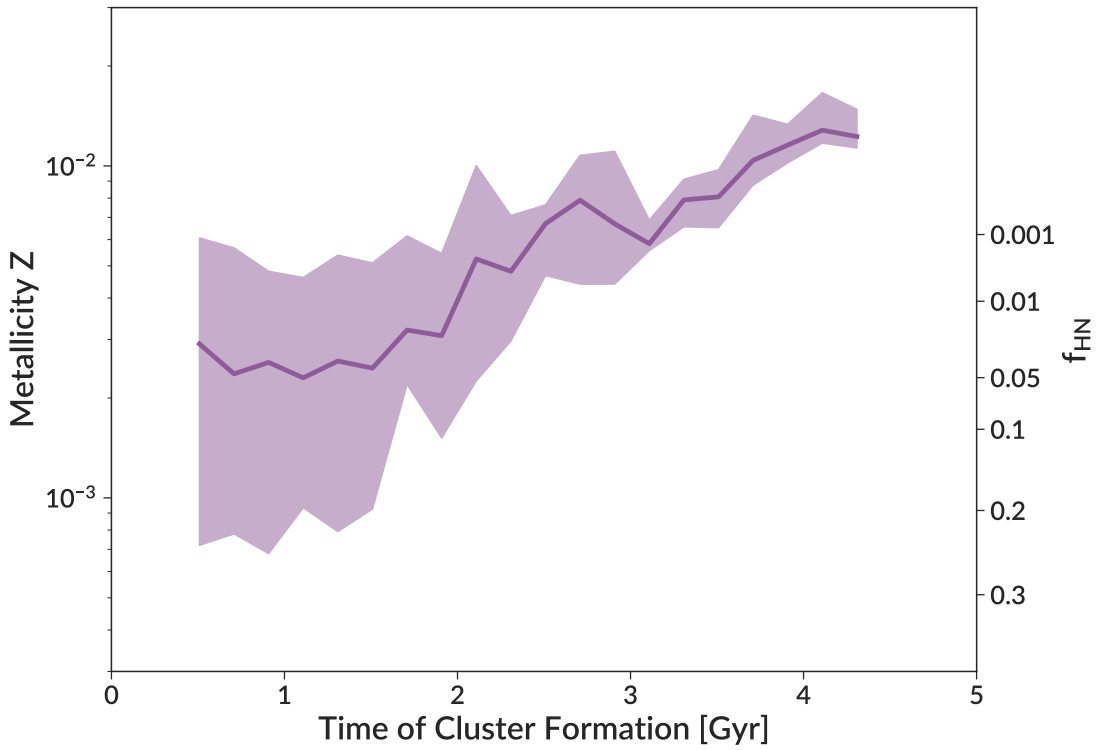

Figure 5.8 The stellar metallicity Z (not scaled to solar metallicity) of clusters forming across cosmic time in the run using the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 50\%$. The shaded region shows the interquartile range at a given age, while the solid line shows the median. The right axis labels shown the hypernova fraction at a given metallicity.

$f_{\rm HN} < 10\%$. As shown in Figure 5.1, this small $f_{\rm HN}$ produces energy injection rates not too dissimilar from $f_{\rm HN} = 0$. This small change is in contrast to the large changes in momentum feedback that come from varying $f_{\rm boost}$ by a factor of 5, explaining why $f_{\rm boost}$ has a strong impact on galactic properties while $f_{\rm HN,0}$ does not.

While changes in $f_{\rm boost}$ lead to dramatic changes in global galaxy properties, the changes to the cluster mass function are more subtle. Figure 5.9 shows the initial cluster mass function for the `Isolated MW` IC when varying $f_{\rm boost}$ and $f_{\rm HN,0}$. We show all clusters formed before $z = 4$, as this higher redshift reduces the differences caused by variations in the star formation rate and includes a higher fraction of low-metallicity clusters where $f_{\rm HN}$ could potentially make a difference. The normalization

changes reflect the change in total stellar mass. Interestingly, the high-mass end is less affected by $f_{\rm boost}$ than the low-mass end. A lower $f_{\rm boost}$ serves to increase the number of low-mass particles without systematically increasing the number of massive clusters or the maximum cluster mass. While not shown in Figure 5.9, we see the same trends when examining the runs using the Local Group ICs.

We find little difference in the cluster mass function when changing $f_{\rm HN,0}$. Figure 5.9 shows little difference between $f_{\rm HN,0} = 50\%$ and $f_{\rm HN,0} = 0$ for $f_{\rm boost} = 1$ for masses below $10^6\,{\rm M}_\odot$. However, the run with $f_{\rm HN,0} = 0$ has several clusters with masses above $10^6\,{\rm M}_\odot$, while the run with $f_{\rm HN,0} = 50\%$ does not. There are very few clusters in these mass ranges, so stochasticity may play a role in these results. We also examined the low-metallicity clusters separately, again finding no difference. This is true as well of the runs with the Local Group ICs.

To quantitatively evaluate the shape of the mass functions, we fit them with a power-law. As our mass functions do not show a power-law behavior down to low masses, we restrict our fit to masses above $10^5\,{\rm M}_\odot$ where it is approximately a power-law. Again we note that we are using the particle masses without including the bound fraction, so these results are not directly comparable to observations. Including the bound fraction generally makes the mass function shallower, as high mass clusters have a higher bound fraction (see Section 5.4). For $f_{\rm HN,0} = 0$, we find slopes of -2.94, -2.48, -2.16, and -2.31 for $f_{\rm boost} = 1, 2, 3$ and 5, respectively. For $f_{\rm HN,0} = 50\%$ and $f_{\rm boost} = 1$, we find -2.78. Lower values of $f_{\rm boost}$ tend to have steeper slopes due to the higher number of low mass clusters. The $f_{\rm boost} = 3$ run has the shallowest slope due to the large number of clusters at $\approx 5 \times 10^5\,{\rm M}_\odot$ that deviate from a pure power law fit and draw the fit toward a shallower slope. This feature becomes less prominent at $z = 1.5$ as more clusters form and fill out the mass function more evenly. We see similar trends in the Local Group runs, where the slope takes values of -2.62, -2.40, and -2.22 for $f_{\rm boost} = 1, 3$, and 5 respectively.

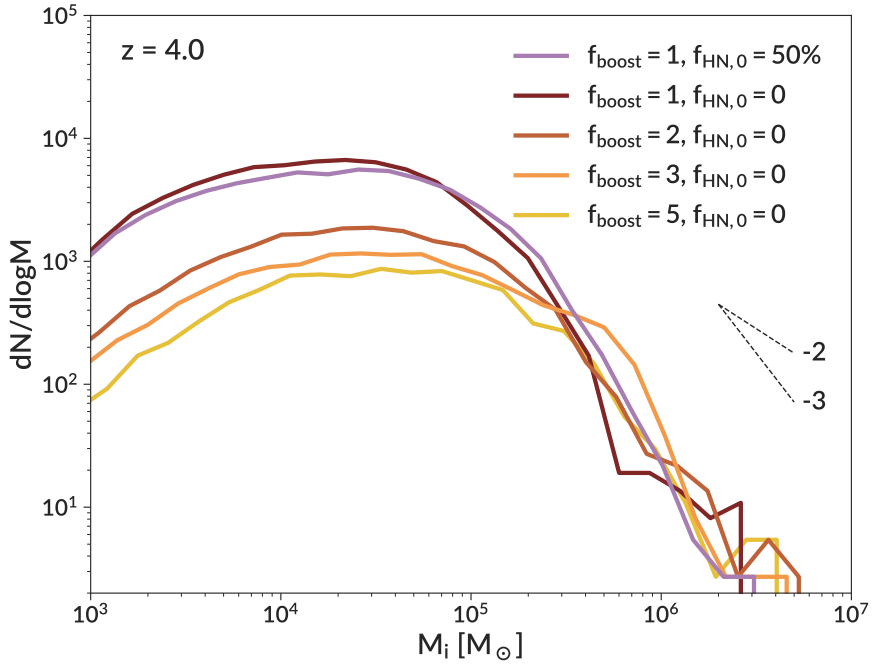

Figure 5.9 A comparison of the initial cluster mass function for runs with varied $f_{\rm boost}$ and $f_{\rm HN,0}$ at $z = 4$. Black dashed lines indicate power-low slopes of -2 and -3. The lower limit of the plot corresponds to one cluster. All runs use the `Isolated MW` IC and $\epsilon_{\rm ff} = 100\%$.

Lastly, we examined the visual appearance of the gas distribution in these galaxies. L18 found that reducing $f_{\rm boost}$ to 3 led to a dramatic increase in star formation and the formation of an axisymmetric disc, while runs with $f_{\rm boost} = 5$ produced very irregular galaxies (Meng et al., 2019). Here, we find that all of our runs produce irregular galaxies, even with $f_{\rm boost} = 1$.

In summary, we find that higher values of $f_{\rm boost}$ can greatly decrease the galactic star formation rate by decreasing the number of low mass clusters that form, without changing the number of massive clusters. Increasing the initial hypernova fraction $f_{\rm HN,0}$ has little effect on galaxy properties. The fraction quickly approaches zero as metallicity increases, leading to little difference in the injected energy and momentum.

5.3.3 Molecular Gas Prescription

A key ingredient in modeling star formation in our simulations is the amount of molecular gas, as we require a given cell to have a mass fraction of molecular gas greater than 50% to seed a cluster particle.

As discussed in Gnedin & Draine (2014) and Appendix A7 of Gnedin & Kravtsov (2011), the clumping factor C_ρ is one of the tunable parameters of the molecular gas model. This factor accounts for the fact that gas is clustered on scales that are not resolved in a given simulation, so H_2 formation would be missed. Larger values of the clumping factor produce more molecular gas at a given surface density. Numerical simulations of turbulent molecular clouds find lognormal density distributions with widths that imply $C_\rho \approx 3-10$ (McKee & Ostriker, 2007). Gnedin & Kravtsov (2011) and Gnedin & Draine (2014) calibrated the clumping factor in the ART code based on simulations, finding that values in the range of 10 to 30 work well. However, those simulations had lower resolution than our runs. This would imply that our runs should prefer a lower clumping factor, because they are resolving more substructure and leaving less on subgrid scales.

Motivated by the disagreement between our simulations and the observed global galactic star formation history (e.g. Figure 5.6), we explored a range of molecular gas prescriptions. We ran simulations with a range of clumping factors, using $C_\rho = 3, 10,$ and 30. L18 used 10, as do all other runs presented in this paper. We also used one run with $C_\rho = 3$ where we changed the prescription for shielding from that of Gnedin & Draine (2014) to that of Gnedin & Kravtsov (2011). The Gnedin & Draine (2014) model includes the effects of line overlap in the Lyman-Werner bands, increasing self-shielding, which is particularly relevant in low metallicity environments with less dust shielding. However, both models for self-shielding were calibrated using runs with lower resolution than our runs ($>$ 50 pc compared to 3–6 pc) and with a different feedback model. These differences in simulation setup can affect the performance of

the H_2 formation model, so we decided to explore both shielding prescriptions. All runs used $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and $\epsilon_{\rm ff} = 100\%$.

As expected, the only significant differences caused by C_ρ were in the amount of molecular gas. While the mass of molecular gas in each run varies greatly with time, we find a general trend that larger values of C_ρ produce more molecular gas. We see little change in molecular gas masses when changing the shielding prescription. These differences in the amount of molecular gas when changing C_ρ led to some differences in star formation histories. The total stellar mass at $z = 1.5$ for the run with $C_\rho = 3$ is $3 \times 10^9\,{\rm M}_\odot$, while the mass for the run with $C_\rho = 30$ is $6 \times 10^9\,{\rm M}_\odot$. In particular, a higher clumping factor leads to more late-time star formation.

5.3.4 Star Formation Efficiency

The local star formation efficiency per freefall time $\epsilon_{\rm ff}$ is a key parameter of our model (see Equation 5.1). As L18 showed, this parameter strongly influences many star cluster properties, particularly the mass function, while not strongly affecting the global galaxy properties. We continue that exploration here.

As $\epsilon_{\rm ff}$ controls how fast star particles accrete material, we expect it to be reflected in the duration of cluster formation episodes. We find that to be the case. In particular, we find that runs with low values of $\epsilon_{\rm ff}$ often fail to finish forming massive clusters before the algorithmic end to a star formation episode at 15 Myr. For example, in the run using the `Isolated MW` IC, $\epsilon_{\rm ff} = 1\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$, only 20% of clusters with masses above $10^5\,{\rm M}_\odot$ finished their formation before it was automatically stopped.

When this time cap is imposed, cluster formation ends even when gas is available to continue accreting onto the cluster. Therefore, we cannot interpret these particles as the end-products of cluster formation. Their masses are not self-consistently determined by their feedback. The masses we obtain are lower limits to the true masses

that would have formed over longer timescales. However, as we will discuss more in Section 5.5.3, such long age spreads of stars within a single cluster are ruled out by observations. We define runs as having failed cluster formation if more than 50% of clusters with masses above $10^5\,M_\odot$ have durations longer than 14 Myr. This applies to all runs with $\epsilon_{\rm ff} = 1\%$ and the run using the `Isolated MW` IC, $\epsilon_{\rm ff} = 10\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$. While we still include these runs in plots in this section, we indicate the cluster mass ranges where they are unreliable using dashed lines. We defer a full investigation of this failed cluster formation to Section 5.5.4.

To illustrate the difference in the timescale of cluster formation, Figure 5.10 shows the cumulative distribution of age spread $\tau_{\rm spread}$ for runs using the `Isolated MW` IC. The dependence on $\epsilon_{\rm ff}$ is clear. For massive clusters, the median age spread is 8.6 Myr for $\epsilon_{\rm ff} = 1\%$, while it is 2.4 Myr for $\epsilon_{\rm ff} = 10\%$ and 0.9 Myr for $\epsilon_{\rm ff} = 100\%$. For $\epsilon_{\rm ff} = 1\%$ many clusters have unphysically long age spreads, some longer than 15 Myr. We note that the age spread can be longer than the duration of star formation in some cases, as it is a measure of the variance in the star formation rate rather than simply its length. Atypical star formation histories, such as one with bursts of star formation at early and late times, can lead to large values of $\tau_{\rm spread}$. There is also a clear mass dependence. Clusters with masses below $10^4\,M_\odot$ and $\epsilon_{\rm ff} \geq 10\%$ have median age spreads less than 0.2 Myr, with all low-mass clusters from those runs having age spreads less than 2 Myr. However, for $\epsilon_{\rm ff} = 1\%$ there is a clear tail to long age spreads even among low-mass clusters, with some clusters having age spreads as long as 10 Myr.

We next investigate the effect of $\epsilon_{\rm ff}$ on the cluster mass functions. Figure 5.11 shows the initial cluster mass function for runs using the `Isolated MW` IC, and Figure 5.12 shows the same for the Local Group runs. Similar trends are seen in both plots. Higher values of $\epsilon_{\rm ff}$ lead to more massive clusters and a higher maximum cluster mass, while lower values of $\epsilon_{\rm ff}$ produce more low-mass clusters. The exception to

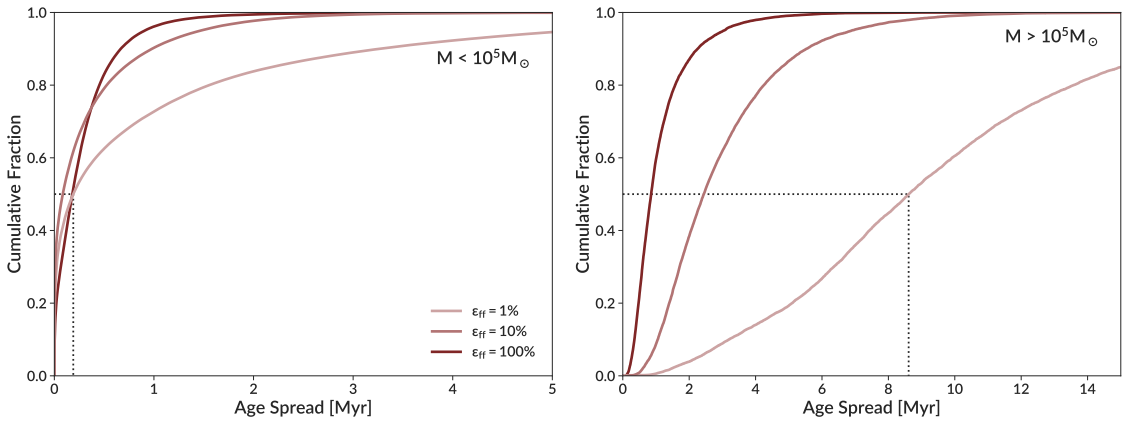

Figure 5.10 The cumulative distribution of the star particle internal age spread $\tau_{\rm spread}$ for the `Isolated MW` IC runs with varied $\epsilon_{\rm ff}$. Note that this is not the duration of star formation as plotted in Figure 5.4, it is the age spread as defined in Equation 5.16 evaluated at the end of cluster formation. The left panel shows clusters less massive than $10^5\,{\rm M}_\odot$, while the right panel shows clusters more massive than $10^5\,{\rm M}_\odot$. Note the different range spanned by the two panels. The dotted line shows the longest median age spread. All runs use $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

this is a handful of very massive clusters that formed in the `Isolated MW` $\epsilon_{\rm ff} = 1\%$ run, leading to a separate hump in the high-mass end of the mass function. The indicates that even with low values of $\epsilon_{\rm ff}$, massive clusters are still possible, although typically rare. We note that we do not see such hump in the `Thelma & Louise` run with $\epsilon_{\rm ff} = 1\%$.

The slope of the high-mass end of the mass function varies with $\epsilon_{\rm ff}$, with the mass function being shallower for higher values of $\epsilon_{\rm ff}$. As with all calculations of the mass function slope, we restrict our fit to clusters above $10^5\,{\rm M}_\odot$. For $\epsilon_{\rm ff} = 1\%$, the slope is between -3.78 and -4.41 for runs on the different ICs, while for $\epsilon_{\rm ff} = 10\%$ it is between -2.94 and -3.42, and for $\epsilon_{\rm ff} = 100\%$ it is between -2.25 and -2.60.

The exact shape of the mass function is somewhat different between the runs that use the `Isolated MW` IC and those that use the Local Group ICs, with the Local Group runs having fewer low-mass particles. These Local Group runs used $f_{\rm boost} = 5$, which decreases the number of low-mass clusters compared to lower values of $f_{\rm boost}$

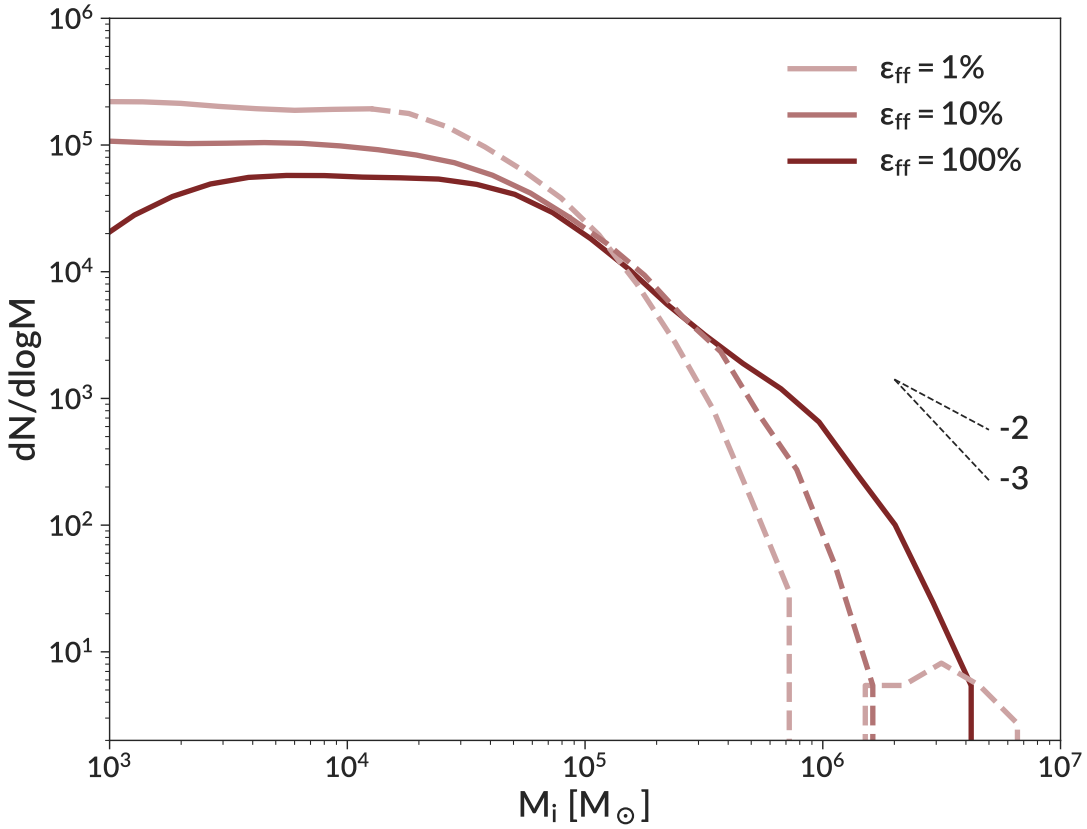

Figure 5.11 A comparison of the star particle initial mass function for the `Isolated MW` IC runs with varied $\epsilon_{\rm ff}$. For runs with failed cluster formation, dashed lines indicate the range where more than 50% of clusters have formation durations longer than 14 Myr, indicating that their masses are likely underestimated. Black dashed lines indicate power-low slopes of -2 and -3. The lower limit of the plot corresponds to one cluster. All runs use $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

(see Figure 5.9). The different redshift of these runs also likely contributes. We find that the majority of massive particles form at very high redshift or in galactic mergers, when the star formation rate is high. This matches what was seen in L18, and agrees with both observations and theoretical expectations (Portegies Zwart et al., 2010; Kruijssen, 2014). In more quiescent epochs, high-mass particles do not form, giving proportionally more low-mass particles. As time progresses, more low-mass clusters are likely to form in the Local Group runs, possibly making their mass functions more similar to those seen in the `Isolated MW` IC.

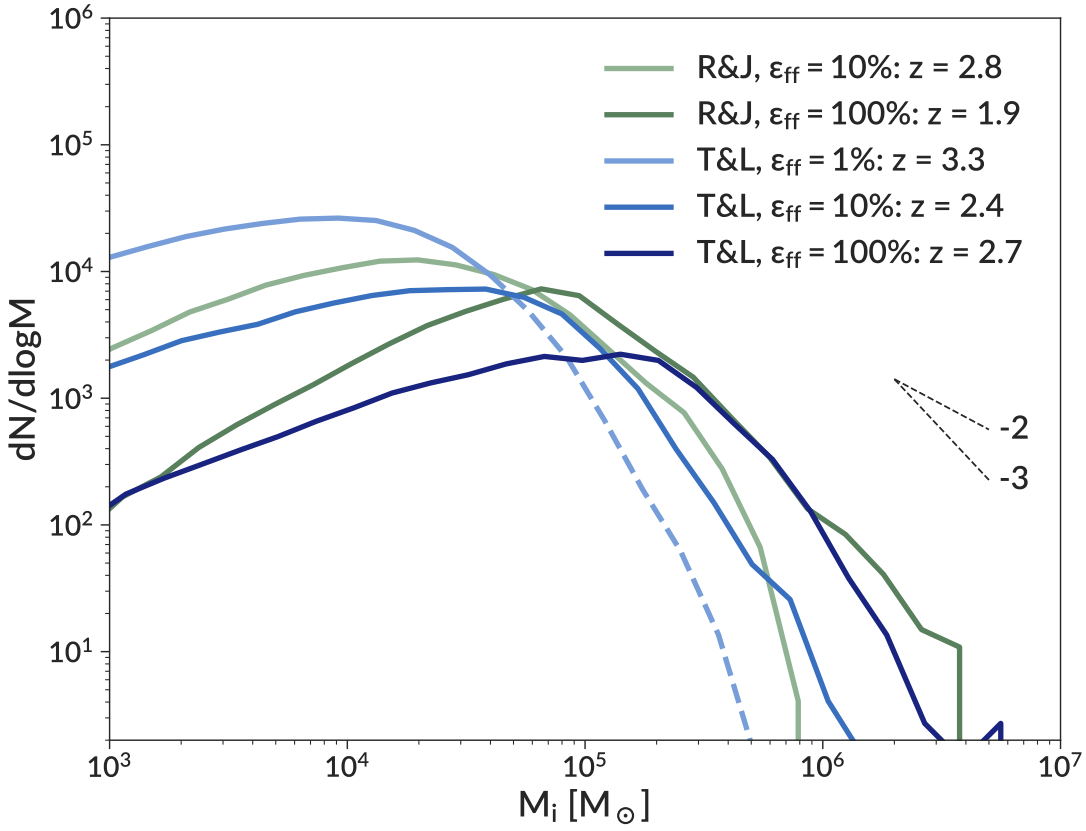

Figure 5.12 A comparison of the star particle initial mass function for the Local Group ICs with varied $\epsilon_{\rm ff}$. For the $\epsilon_{\rm ff} = 1\%$ run with failed cluster formation, dashed lines indicate the range where cluster masses are likely underestimated. Black dashed lines indicate power-low slopes of -2 and -3. The lower limit of the plot corresponds to one cluster. All runs use $f_{\rm boost} = 5$, $f_{\rm HN,0} = 20\%$, and show all clusters formed before the redshift listed in the legend.

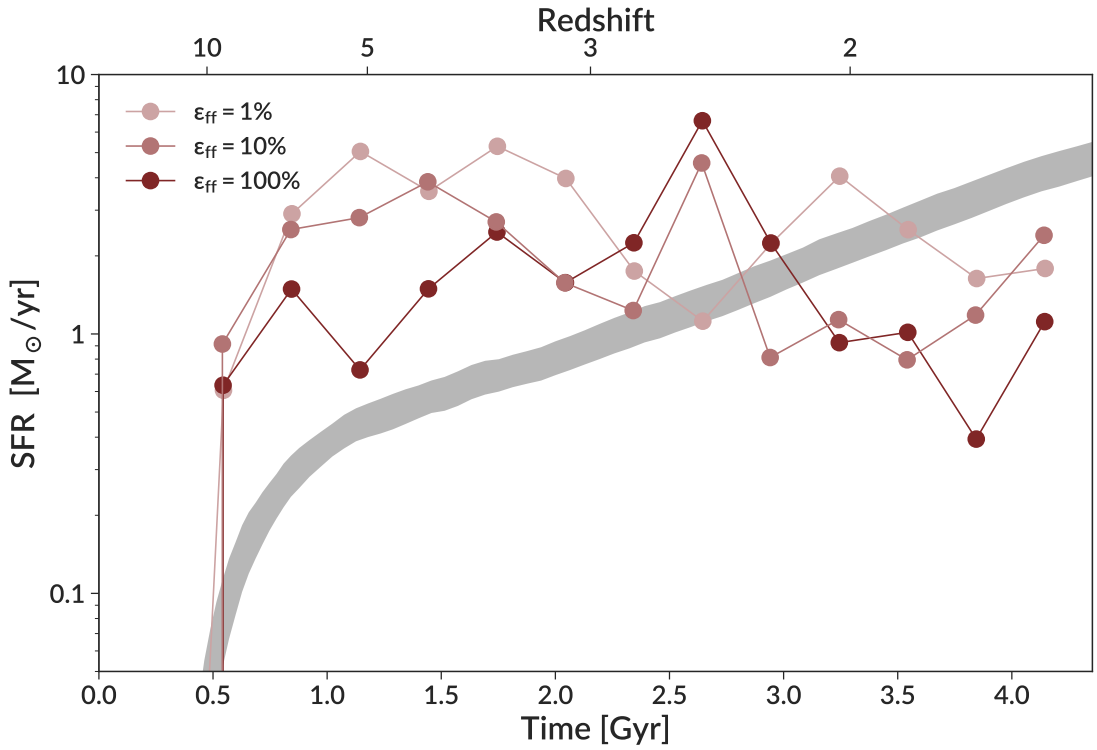

Figure 5.13 A comparison of the star formation rate of the central galaxy in the `Isolated MW` IC when varying $\epsilon_{\rm ff}$. All runs use $f_{\rm boost} = 1$ and $f_{\rm HN,0} = 0$.

While $\epsilon_{\rm ff}$ significantly affects cluster properties, it does not change the galactic star formation rate appreciably. Figure 5.13 shows the star formation histories of runs when varying $\epsilon_{\rm ff}$ while holding $f_{\rm boost} = 1$ and $f_{\rm HN,0} = 0$ constant. Here we find that lower values of $\epsilon_{\rm ff}$ lead to somewhat higher star formation rates at early times. These star formation rates at $z \approx 5$ are significantly higher than predicted by UniverseMachine, and tend to decline with time rather than increase. However, we find opposite trends during the major merger at $z \approx 2.6$, when the high $\epsilon_{\rm ff}$ runs show a stronger burst. In the runs using the Local Group analogs, the star formation history does not change significantly with $\epsilon_{\rm ff}$.

In summary, we find that $\epsilon_{\rm ff}$ does not have a significant impact on the galactic star formation rate, but does strongly influence star cluster properties. In particular, higher values of $\epsilon_{\rm ff}$ lead to more massive star clusters and shorter timescales for cluster

formation. These results confirm those found in L18, indicating that they are robust predictions of our simulations.

5.3.5 Virial Criterion

One of the other changes to our star cluster formation prescription was the addition of a criterion restricting star-forming gas to be gravitationally bound (see the beginning of Section 5.2.2). To investigate the difference this makes in cluster properties, we ran one simulation with the virial criterion turned off. While we find no significant differences in large scale galactic properties, we find differences in the star cluster populations. Figure 5.14 shows the mass function for runs with and without the virial criterion. The addition of the virial criterion leads to more high-mass clusters and fewer low-mass clusters. Quantitatively, the power-law slopes of the mass functions for clusters above $10^5\,M_\odot$ are -2.60 for the run with the virial criterion and -3.30 for the run without it. While the maximum cluster mass is similar between the two runs, there are significantly more clusters with masses above $10^6\,M_\odot$ when the virial criterion is enabled.

The increase in the number of high-mass clusters is expected, as Equation 5.3 shows lower gas densities lead to higher virial parameters. The cut on the virial parameter prevents these lower density GMCs from forming stars until they accrete more gas and collapse to higher density, leading to more total mass available for star formation. The later onset of star formation also delays stellar feedback, allowing more gas to accrete onto the cluster. These processes shift many low-mass clusters to higher masses, explaining the decrease in the number of low-mass clusters. In addition, as the virial criterion allows more gas accretion onto the GMC, its larger mass becomes more difficult to disperse with feedback, leading to longer durations of star formation. As a consequence of these effects, clusters have higher values of $\epsilon_{\rm int}$ when the virial criterion is enabled. In Figure 5.15 we show the distribution of

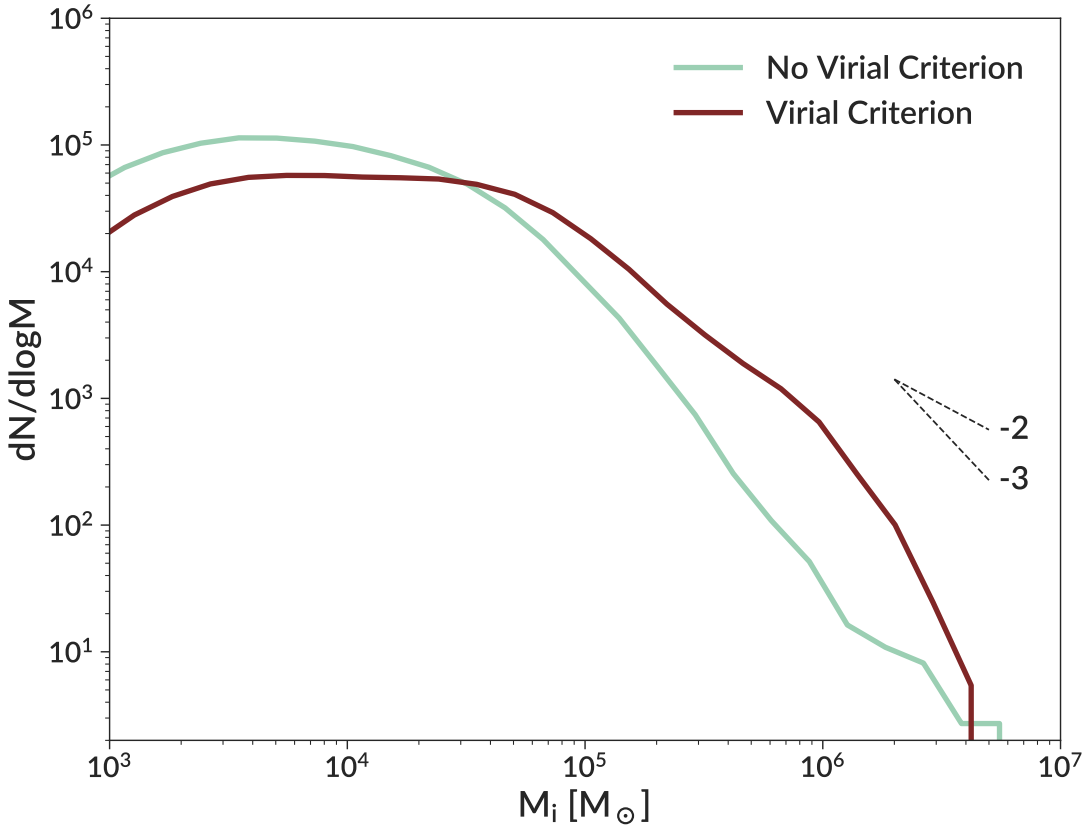

Figure 5.14 A comparison of the initial cluster mass function for runs with and without the virial criterion for seeding star formation. Black dashed lines indicate power-low slopes of -2 and -3. Both runs used the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

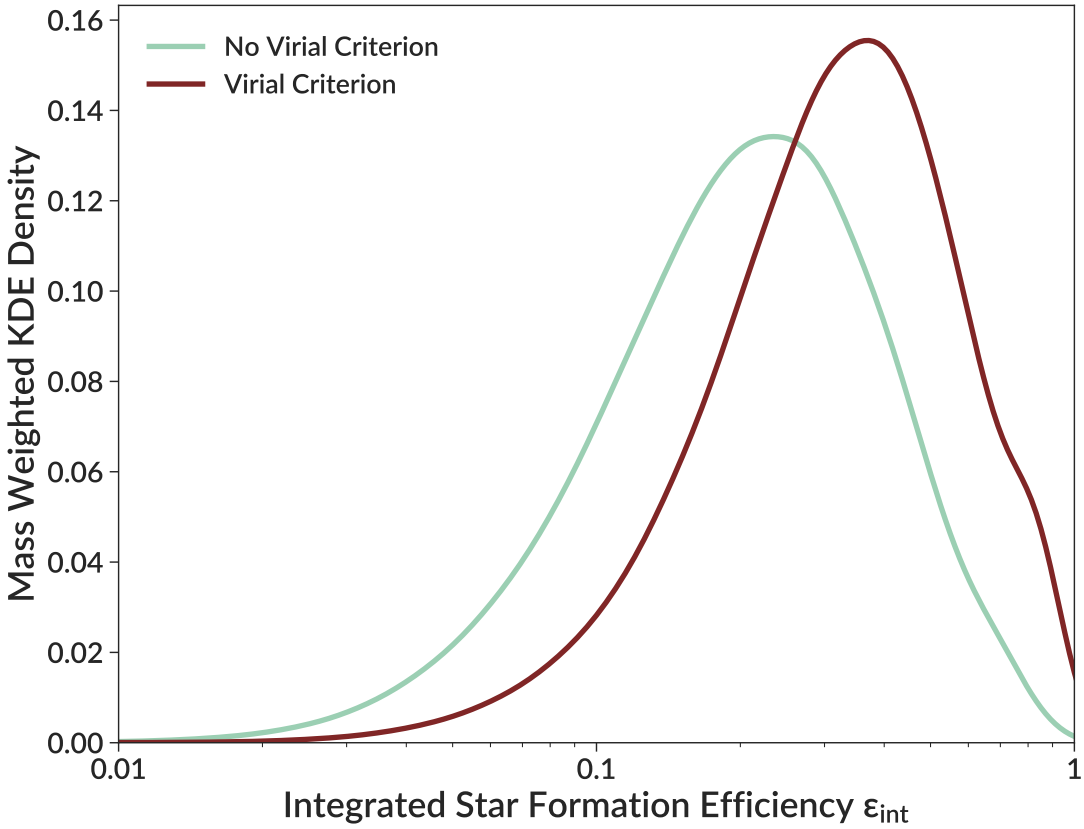

Figure 5.15 Kernel density estimation for the distribution of integrated star formation efficiency for clusters in the runs with and without the virial criterion. We use a Gaussian kernel with a width of 0.05 dex. Each curve is normalized to the same area for comparison purposes. Both runs used the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and show all clusters formed before $z = 1.5$.

$\epsilon_{\rm int}$ with and without the virial criterion. Both distributions have widths ≈ 0.25 dex, but the mean value for the run with the virial criterion is significantly higher (35% compared to 21%).

In the run where we did not impose the virial criterion, we output the virial parameter $\alpha_{\rm vir}$ of each cluster as it formed. Using this information, we can postprocess the results to see if there are any correlations between the virial parameter and the resulting cluster properties. We find that clusters with $\alpha_{\rm vir} < 10$ tend to have higher initial masses, higher $\epsilon_{\rm int}$, and higher initial bound fractions than those with $\alpha_{\rm vir} > 10$. The virial criterion acts in a biased fashion to allow star formation to happen in regions

that preferentially lead to higher mass clusters. Additionally, regions with $\alpha_{\rm vir} > 10$ are able to accrete more material over time until they pass the $\alpha_{\rm vir} < 10$ threshold, increasing the cluster mass that formed out of a given GMC.

In summary, we find that adding the requirement that star-forming gas have a virial parameter $\alpha_{\rm vir} < 10$ increases the number of massive clusters, gives clusters a longer formation timescale, and leads to higher values of $\epsilon_{\rm int}$.

5.4 Evolution of the Cluster Mass Function

In the previous section we exclusively used the masses of the star particles at the end of their formation process. As not all stars are gravitationally bound to the newly formed cluster, we must incorporate the initial bound fraction to obtain the observable cluster masses. In addition, the plots in the previous section showed the distributions of initial masses for all clusters formed over the full time spanned by the simulation. This is not observable. In this section we include the cluster bound fraction and present the instantaneous cluster mass function at a given redshift to allow for more direct comparison with observations. While these are not true mock observations, the results shown here accurately represent the existing cluster populations at a given redshift in our simulations.

We start by examining the cluster initial bound fraction, which is needed to turn raw particle masses into bound cluster masses. Figure 5.16 shows the initial bound fraction of clusters as a function of mass. As in L18, we see the trend of higher mass clusters having higher bound fraction. Additionally, runs with higher $\epsilon_{\rm ff}$ have higher bound fractions at a given particle mass.

Our prescription for the initial bound fraction (Equation 5.4) makes it solely dependent on the integrated star formation efficiency $\epsilon_{\rm int}$. In Figure 5.17 we show the distributions of $\epsilon_{\rm int}$. Runs with lower $\epsilon_{\rm ff}$ have lower $\epsilon_{\rm int}$. For a given run, the spread is due to trends with mass, where high-mass clusters have higher $\epsilon_{\rm int}$ than

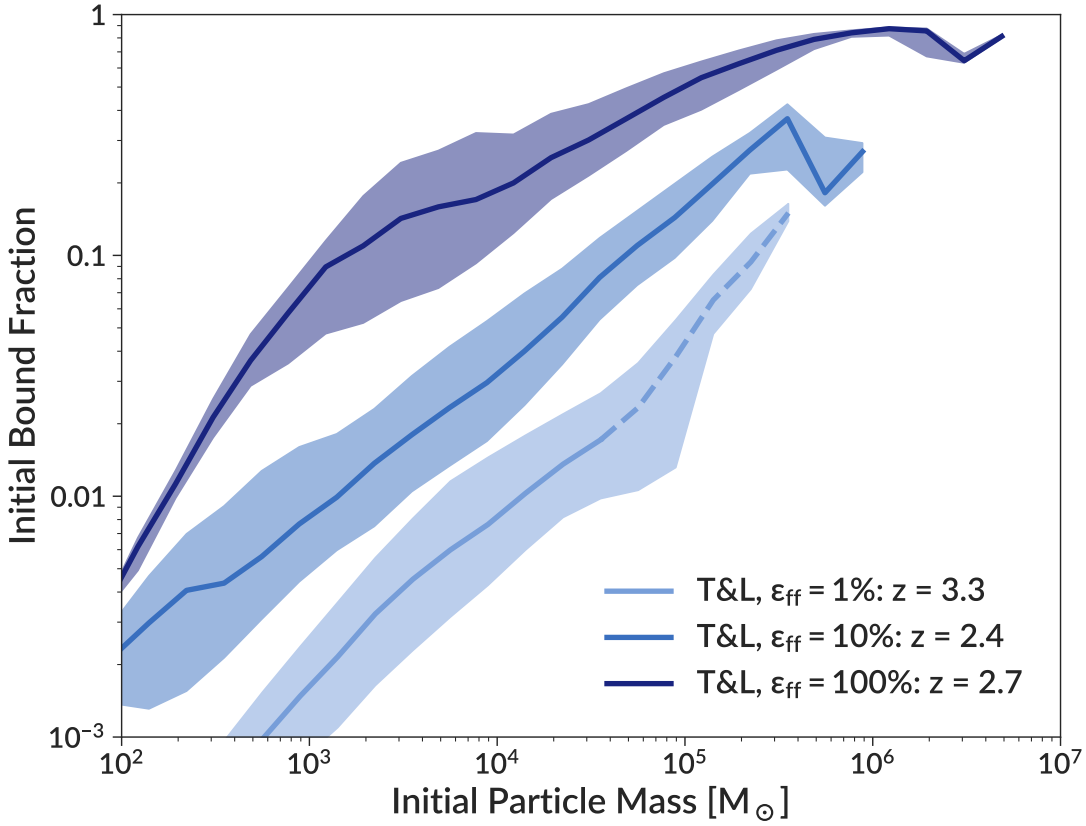

Figure 5.16 The initial bound fractions for runs using the `Thelma & Louise` IC and varying $\epsilon_{\rm ff}$. The solid line shows the median, with the shaded region showing the interquartile range of the distribution of the initial bound fraction at that mass. The mass plotted here is the particle mass at the end of cluster formation, not the bound cluster mass, so that the plotted variables are independent. For the $\epsilon_{\rm ff} = 1\%$ run with failed cluster formation, dashed lines indicate the range where cluster masses are likely underestimated. We only show the `Thelma & Louise` IC for clarity, but `Romeo & Juliet` and `Isolated MW` show the same behavior. All runs use $f_{\rm boost} = 5$, $f_{\rm HN,0} = 20\%$, and show all clusters formed before the redshift listed in the legend.

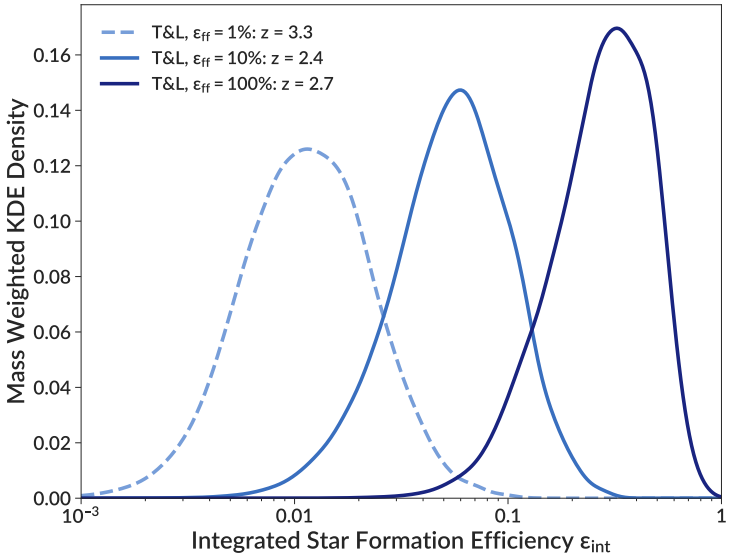

Figure 5.17 Kernel density estimation for the distribution of integrated star formation efficiency for clusters in the `Thelma & Louise` IC with variations in $\epsilon_{\rm ff}$. We use a Gaussian kernel with a width of 0.05 dex. Each curve is normalized to the same area for comparison purposes. All runs use $f_{\rm boost} = 5$, $f_{\rm HN,0} = 20\%$, and show all clusters formed before the redshift listed in the legend. We plot the $\epsilon_{\rm ff} = 1\%$ run with a dashed line as that run had many clusters that failed to finish forming.

low-mass clusters. Quantitatively, the mean value of $\epsilon_{\rm int}$ takes values of 1.2%, 7.2%, and 30% for $\epsilon_{\rm ff} = 1\%$, 10%, and 100%, respectively. As $\epsilon_{\rm ff}$ increases, the widths of these distributions decrease, with values of 0.30, 0.24, and 0.17 dex, respectively.

The trend of higher $\epsilon_{\rm int}$ with higher $\epsilon_{\rm ff}$ is a direct consequence of $\epsilon_{\rm ff}$ controlling the cluster formation rate (Equation 5.1). Higher $\epsilon_{\rm ff}$ leads to higher star formation rates, allowing the cluster to accrete more of the gas from its surroundings. This is reflected in the duration of cluster formation in runs with different $\epsilon_{\rm ff}$. A lower value $\epsilon_{\rm ff}$ causes clusters to form more slowly. With a slow star formation rate, feedback also starts before the cluster has accreted a significant fraction of the surrounding gas, leading to lower $\epsilon_{\rm int}$. The different timescales also likely lead to the change in width of the distributions. As low values of $\epsilon_{\rm ff}$ lead to longer timescales of cluster formation, there is more possibility for variation in the accretion history of the GMC. High values of $\epsilon_{\rm ff}$ form quickly, so they are forming mostly out of the gas that was present at cluster birth.

In addition to the initial bound fraction, we also need to account for stellar evolution and dynamical disruption, which both cause clusters to lose mass with time. These processes are calculated in simulation runtime. In general, the mass M_b bound to a cluster at time t can be written as

$$M_b(t) = M_i \, f_i \, f_{\rm se}(t) \, f_{\rm dyn}(t) \tag{5.17}$$

where M_i is the initial particle mass, f_i is the initial bound fraction, $f_{\rm se}(t)$ accounts for mass loss due to stellar evolution, and $f_{\rm dyn}$ accounts for mass lost due to tidal stripping (Li & Gnedin, 2019; Meng & Gnedin, 2022). Our feedback scheme self-consistently decreases the stellar mass of the cluster whenever mass is ejected into the ISM, and dynamical disruption is calculated as described in Section 5.2.3. In Figure 5.18, we show the impact of disruption on clusters of different mass, taking as an example the run using the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$. We choose a run using the `Isolated MW` IC as it reached a lower redshift, so clusters have more time to disrupt. Clusters with masses below $10^4\,{\rm M}_\odot$ are entirely disrupted within 500 Myr. Clusters of intermediate mass $10^4 - 10^5\,{\rm M}_\odot$ persist for a few Gyr, but do not survive until the present. However, clusters with masses above $10^5\,{\rm M}_\odot$ survive throughout the lifetime of the simulation. Tidal disruption only decreases the mass of these clusters by approximately 20% over the 4 Gyr length of this simulation.

Using these disruption calculations, we now present the mass function of bound clusters at various redshifts. In Figure 5.19, we show the mass function of the surviving clusters at $z = 4$ in the Local Group runs with varied $\epsilon_{\rm ff}$. This figure shows trends similar to those seen in Figure 5.12, with several trends more exaggerated now that bound cluster mass is included. First, we note similar shapes. Our mass functions have a sharp cutoff at high masses, a peak, and a shallower decrease to low masses. This shape is seen in all runs with $\epsilon_{\rm ff} \geq 10\%$. The position of the peak depends

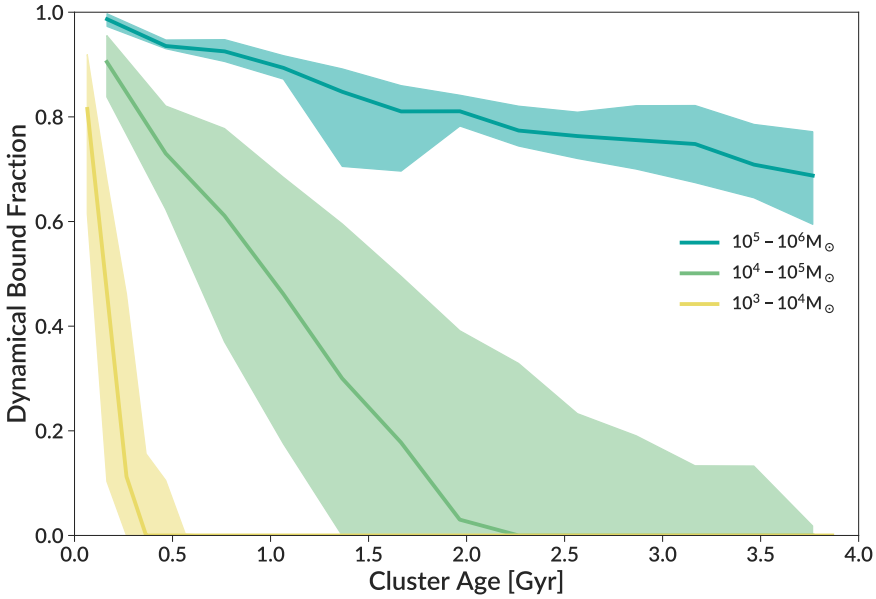

Figure 5.18 Evolution of the dynamical bound fraction $f_{\rm dyn}$ as a function of cluster age for clusters in different mass ranges. Lines show the median, with the shaded region showing the interquartile range. Clusters are grouped according to their initial bound mass at formation. The plot shows clusters in the central galaxy of the run using the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, $f_{\rm HN,0} = 0$, and shows all clusters formed before $z = 1.5$.

strongly on $\epsilon_{\rm ff}$. For $\epsilon_{\rm ff} = 100\%$ it is at approximately $10^5\,{\rm M}_\odot$, while it is closer to $10^4\,{\rm M}_\odot$ for $\epsilon_{\rm ff} = 10\%$. This is due to a combination of three effects. First, as seen in Figure 5.12, the initial particle masses are higher for higher values of $\epsilon_{\rm ff}$. Second, higher values of $\epsilon_{\rm ff}$ give higher bound fractions, as shown in Figure 5.16. The two effects magnify each other, such that higher values of $\epsilon_{\rm ff}$ result in cluster mass functions that reach to significantly higher masses. The disparity is further increased by the effects of disruption, which preferentially removes low-mass clusters (Figure 5.18). These three effects combine to produce dramatically different cluster mass functions when changing $\epsilon_{\rm ff}$. Of note, the $\epsilon_{\rm ff} = 1\%$ run has no existing clusters above $10^4\,{\rm M}_\odot$, while the $\epsilon_{\rm ff} = 10\%$ runs have no clusters above $3 \times 10^5\,{\rm M}_\odot$.

We also note that, as described in Section 5.3.2, the low-mass end of the mass function is sensitive to $f_{\rm boost}$, with higher values of $f_{\rm boost}$ decreasing the number of low-mass clusters. The runs shown in Figure 5.19 were all run with $f_{\rm boost} = 5$. Lower

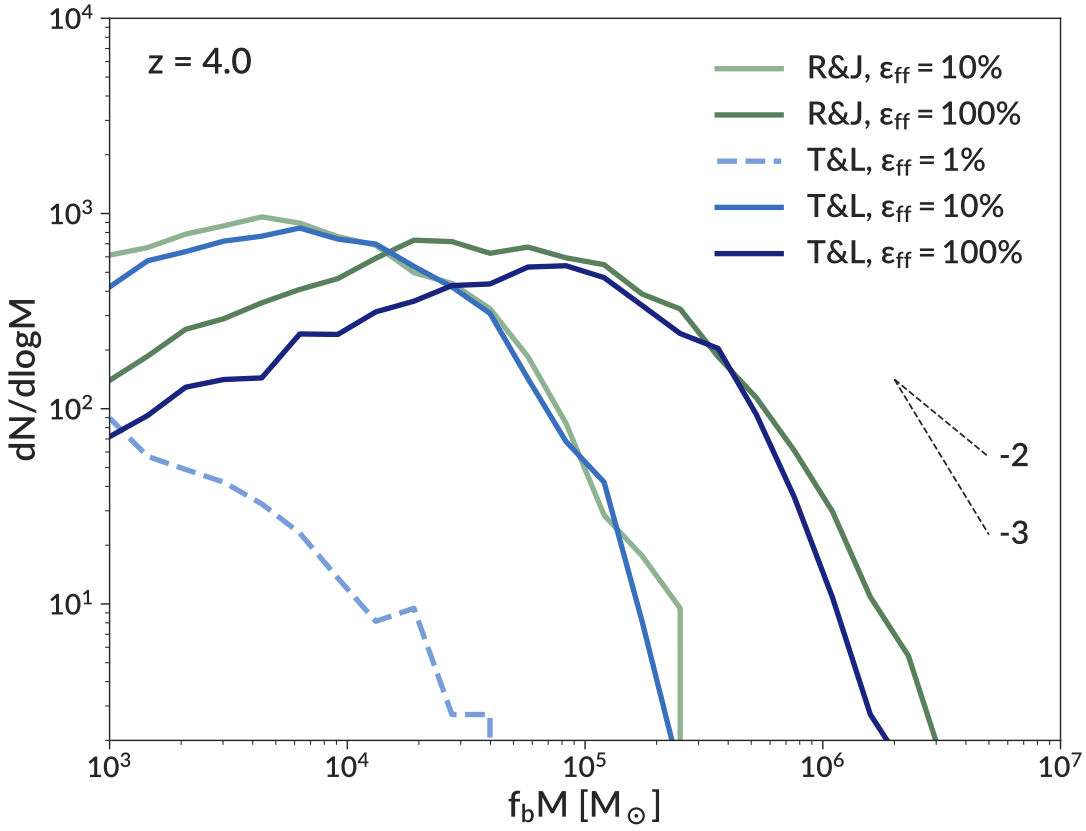

Figure 5.19 The bound mass function of all clusters present at $z = 4$ using the Local Group ICs for different values of $\epsilon_{\rm ff}$. We plot the $\epsilon_{\rm ff} = 1\%$ run with a dashed line as that run had many clusters that failed to finish forming. Black dashed lines indicate power-low slopes of -2 and -3. The lower limit of the plot corresponds to one cluster. All runs use $f_{\rm boost} = 5$ and $f_{\rm HN,0} = 20\%$.

values of $f_{\rm boost}$ would increase the number of low-mass clusters and give it a shape more similar to that seen in the local universe. Similarly, massive clusters tend to form in epochs of intense star formation, while low-mass clusters dominate in more quiescent epochs. As only the `Thelma` IC has any significant mergers after the redshift shown in this plot, we expect there to be more low-mass clusters as time progresses.

In Figure 5.20 we show the evolution of the bound cluster mass function from $z = 6$ to $z = 1.9$ for the run using the `Romeo & Juliet` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 5$, and $f_{\rm HN,0} = 20\%$. A significant fraction of clusters with masses above $2 \times 10^5\,{\rm M}_\odot$ are in place already at $z = 6$. More massive clusters form by $z = 4$, but we see little

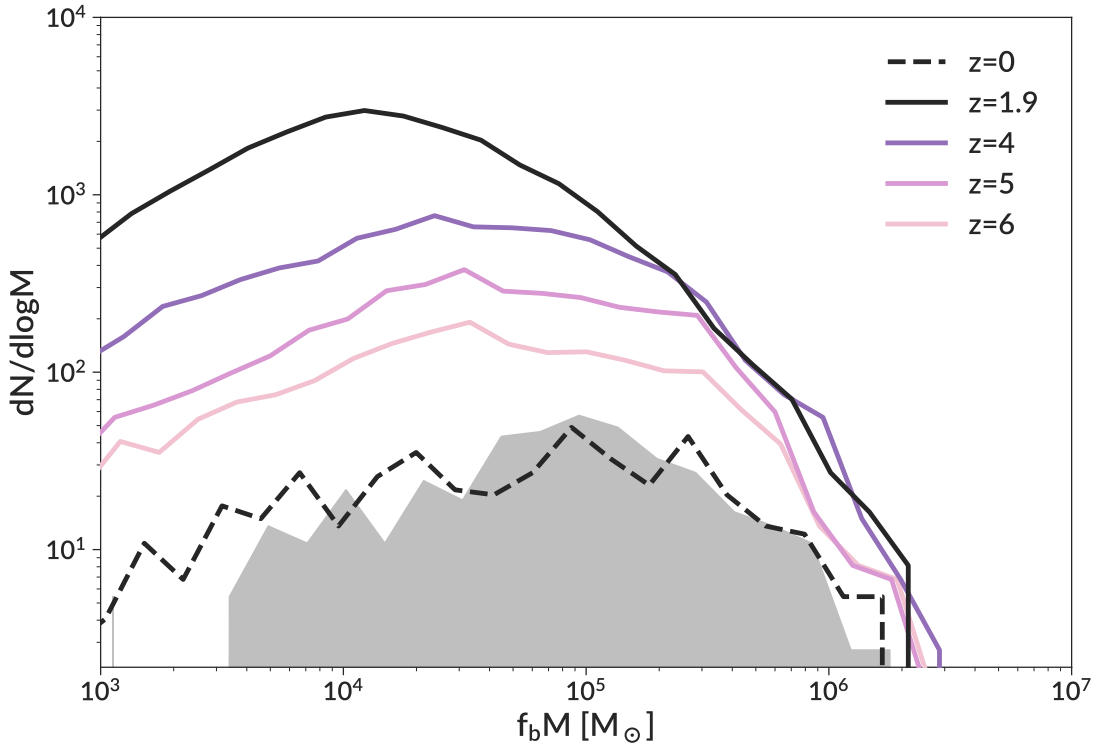

Figure 5.20 The bound mass function of all clusters present at a range of redshifts. All simulation lines are from the same run that uses the `Romeo & Juliet` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 5$, and $f_{\rm HN,0} = 20\%$. The dashed line indicates the cluster population analytically evolved from $z = 1.9$ to $z = 0$. The shaded region shows the observed mass function of clusters in the MW.

change in the massive end of the mass function beyond that redshift. At later epochs low-mass clusters dominate the mass function, particularly increasing the number of clusters around $10^4\,\mathrm{M}_\odot$. Clusters of low mass that appear in this plot are mainly from recent star formation. As Figure 5.18 shows, clusters with masses below $10^5\,\mathrm{M}_\odot$ disrupt within a few Gyr, and clusters below $10^4\,\mathrm{M}_\odot$ disrupt within several hundred Myr.

We also show an analytical evolution of star clusters from the last output of this run at $z = 1.9$ to the present, following Li & Gnedin (2019). The prescription for tidal disruption (Equations 5.6–5.8) depends on the galactic tidal field. In simulation runtime we calculate it self-consistently, but to extrapolate to $z = 0$ we simply assume

a constant value of $\Omega_{\rm tid} = 175$ Gyr^{-1}. This value was chosen to produce the same final number of clusters as are observed in the MW. It results in the disruption of most clusters with masses below $10^5\,{\rm M}_\odot$, and decreases the masses of all surviving clusters, shifting the distribution to lower masses and decreasing the normalization. This calculation also assumes that no new clusters form after $z = 1.9$.

Our chosen value of $\Omega_{\rm tid} = 175$ Gyr^{-1} is quite high. It is equivalent to a maximum eigenvalue of the tidal tensor $\lambda_m \approx 10^5$ Gyr^{-2}. Meng & Gnedin (2022) examined the tidal field for the L18 simulations, finding that clusters experience such strong tidal fields only shortly after their birth. As they migrate away from the high-density star-forming regions, the tidal field decreases significantly to typical values $\lambda_m \approx 3 \times 10^3$ Gyr^{-2}, or $\Omega_{\rm tid} \approx 30$ Gyr^{-1}. Choosing this low value of $\Omega_{\rm tid}$ would significantly increase the number of low-mass clusters surviving to $z = 0$ in our simulations. However, this analysis was done at $z > 1.5$. The value of the tidal field may increase over time as the galaxy grows. Our adopted value is also similar to that used by Choksi & Gnedin (2019a) in an analytic model for cluster formation and destruction. These authors find that $\Omega_{\rm tid} = 200$ Gyr^{-1} can reproduce several observational results, including the GC mass function at $z = 0$ and the relation between galaxy halo mass and mass of its globular cluster system.

We compare our results with the distribution of masses of the observed MW GCs. We use the V-band absolute magnitude from Harris (1996) along with the luminosity dependent mass-to-light ratio

$$\frac{M}{L_V} = 1.3 + \frac{4.5}{1 + \exp\left(2M_V + 21.4\right)} \tag{5.18}$$

from Harris et al. (2017) to obtain GC masses. We find good agreement between the two mass functions. While we match the normalization by construction through our choice of $\Omega_{\rm tid}$, the similarity of the MF shape to that in the MW system is a

test of the model. The maximum cluster mass matches the MW GCs well. We note that the Harris (1996) catalog includes both in-situ and ex-situ clusters in the MW. As the simulation $z = 0$ result comes from analytic evolution of all clusters in the central galaxies at $z = 1.9$, any later clusters that come in from later mergers would be missed. However, Figure 5.20 uses the `Romeo & Juliet` IC, which has quick early growth with no significant mergers after $z = 1.9$ (Figure 5.3). We therefore expect few clusters from later infalling satellites, making a comparison to the full MW population reasonable. We also note that the `Romeo & Juliet` IC has more massive clusters than the `Thelma & Louise` IC. This is likely becuase of its quick early growth (Figure 5.3), increasing the star formation density at early times and leading to the formation of more massive clusters. For the `Thelma & Louise` runs, a lower value of $\Omega_{\rm tid}$ is required to reproduce the high-mass end of the Galactic GC mass function, leading to too many simulated low-mass clusters.

In the runs with $\epsilon_{\rm ff} = 100\%$, our present-day mass functions have more clusters with masses above $3 \times 10^5\,{\rm M}_\odot$ than seen in L18. This is a consequence of our initial mass functions extending to higher masses than in L18. These changes are primarily driven by the addition of the virial criterion. As Figure 5.14 shows, the addition of this criterion significantly increases the number of massive clusters. The increase in the number of massive clusters allows us to increase the value of $\Omega_{\rm tid}$ from 50 ${\rm Gyr}^{-1}$ (used by L18) to 175 ${\rm Gyr}^{-1}$. In that work higher values of $\Omega_{\rm tid}$ would have disrupted too many clusters. In the runs presented in this work, stronger disruption is required to produce an agreement for the massive end of the mass function while reducing the number of low-mass clusters.

Similarly to the mass function at $z = 4$, the mass functions of surviving clusters at $z = 0$ depend strongly on $\epsilon_{\rm ff}$. For all runs with $\epsilon_{\rm ff} \leq 10\%$ (not shown), we find no clusters above $4 \times 10^5\,{\rm M}_\odot$, and the overall distributions shift to lower masses. That is clearly inconsistent with the observed mass function of MW GCs.

Finally, in Figure 5.21 we show the age-metallicity relation for clusters in the run using the `Thelma & Louise` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 5$, and $f_{\rm HN,0} = 0$. There is broad agreement with the observations of MW GCs presented in VandenBerg et al. (2013) and Leaman et al. (2013). Metal-rich clusters form systematically later than metal-poor clusters, as the galaxy enriches its interstellar medium with time. While the plot shows only one run using the `Thelma & Louise` initial condition, we see similar trends in all ICs. We note that the value of $\Omega_{\rm tid}$ used in the analytical disruption calculation slightly affects this result. A change in the disruption rate would affect which clusters that survive to the present. In particular, as most of the highest mass clusters form early, increased disruption tends to remove younger, higher metallicity clusters. While the shape of the age-metallicity relation changes little, the distribution of clusters within it does.

5.5 Discussion

5.5.1 Timing of Supernova Feedback

In Section 5.2.4.7 we discussed several prescriptions for the timing of SN, then in Section 5.3.1 we examined how these prescriptions affect the properties of star clusters. We find that later SN feedback leads to longer timescales for cluster formation and higher values of the integrated star formation efficiency. In this formalism, we assume that there is no difference in the formation time of low and high mass stars within a cluster. Individual stars of all masses have the same age. However, this assumption may be incorrect. For example, using a simulation of a star cluster forming out of a $2 \times 10^4\,{\rm M}_\odot$ GMC, Grudić et al. (2022) find that massive stars ($m > 10\,{\rm M}_\odot$) finish accreting 1 Myr later than the average star. Padoan et al. (2020) find a similar result using a simulation of $2 \times 10^6\,{\rm M}_\odot$ of gas in a $(250\text{ pc})^3$ box with several star-forming regions. The delay in massive star formation in turn delays the onset of feedback.

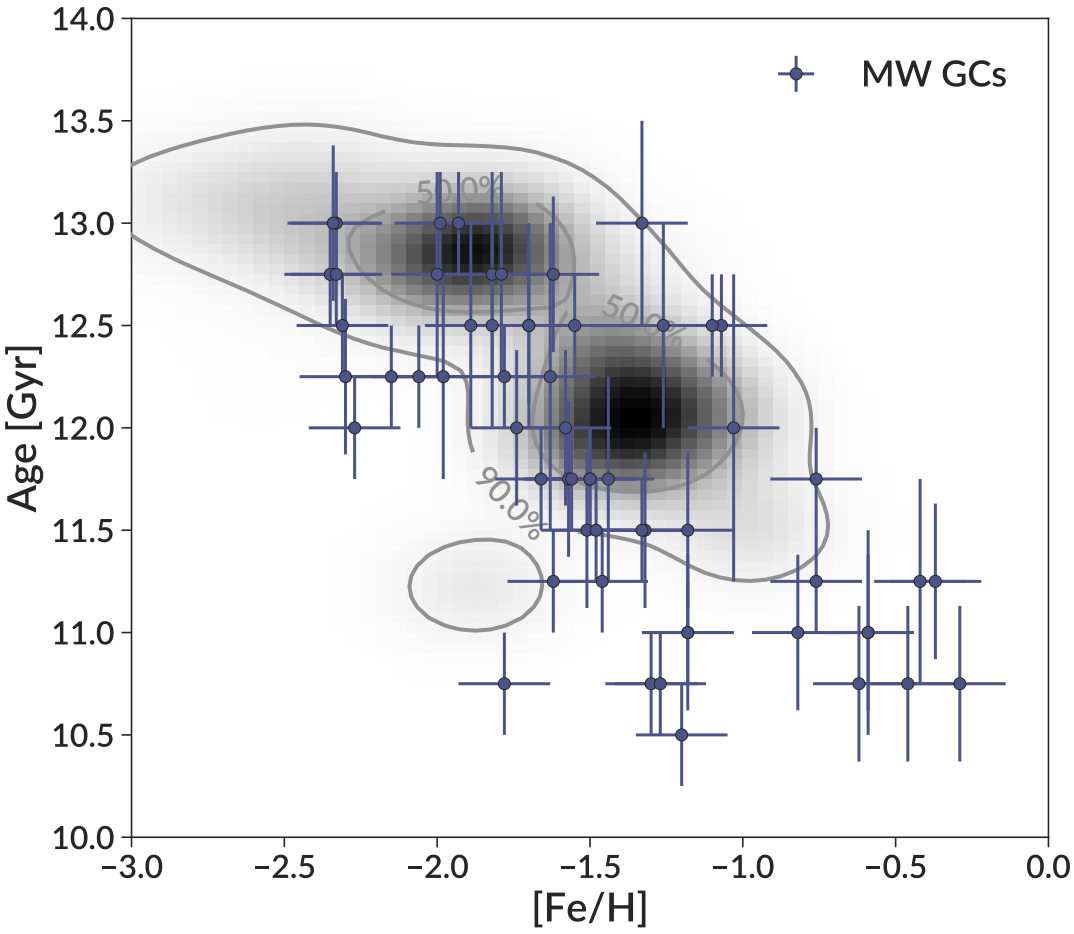

Figure 5.21 The age-metallicity relation for surviving clusters with masses above $3 \times 10^3\,\mathrm{M}_\odot$ at $z = 0$ in the run using the `Thelma & Louise` IC, $\epsilon_{\mathrm{ff}} = 100\%$, $f_{\mathrm{boost}} = 5$, and $f_{\mathrm{HN},0} = 0$. The final output of this run corresponds to an age of 10 Gyr, meaning that all regions on the plot are accessible to the simulation. Clusters from both central galaxies are included as there are no systematic differences between the two. Grey shaded regions and contours indicate the bulk of simulated clusters, with contours enclosing 50 and 90% of the sample. Data points show MW GCs from VandenBerg et al. (2013) and Leaman et al. (2013).

While our simulations account for the stellar age spread within the cluster when determining the timing of SN, they do not account for this systematic delay in the formation of individual stars.

As shown in Figure 5.2, our hybrid approach to the timing of SN feedback approximates well the delay in SN feedback due to the age spread of the stars, so it is our preferred model for future simulations. However, it may need to be further refined to account for the delay in massive star formation. In particular, one possible approach would be to calibrate a subgrid model for the timing of cluster feedback to the results of GMC-scale simulations such as in Grudić et al. (2022). Further delays in the onset of massive star feedback may increase the timescales of cluster formation and the integrated star formation efficiency, but these effects are likely to be small compared to the effects of other parameters, namely $\epsilon_{\rm ff}$.

5.5.2 Strength of Stellar Feedback

In Figures 5.6 and 5.7 we showed how the star formation rate of the central galaxies in our simulations changed when varying $f_{\rm boost}$. We found that $f_{\rm boost} = 5$ produces too little star formation in the current simulation setup. In the `Isolated MW` runs we find that $f_{\rm boost} = 1 - 2$ matches the UniverseMachine predictions fairly well, as do $f_{\rm boost} = 1 - 3$ in the `Thelma & Louise` runs. As we discuss more in Appendix B, updates to the hydrodynamics are primarily responsible for the change in preferred values of $f_{\rm boost}$. Such low values of $f_{\rm boost}$ are unexpected. L18 calibrated $f_{\rm boost}$, finding $f_{\rm boost} = 5$ to be their preferred value. Numerical tests in Semenov et al. (2017) have also shown that values of $f_{\rm boost} \approx 5$ best account for numerical losses of momentum as a SN shell moves across the simulation grid. Theoretical grounds for $f_{\rm boost} > 1$ also exist, with Gentry et al. (2017) finding that clustered SN can enhance momentum feedback by up to an order of magnitude relative to an isolated SN.

We also note that all of these runs, even with $f_{\rm boost} = 1$, show a large decrease in

the star formation rate at $z < 2$, in conflict with the abundance matching expectation. Both the hydrodynamics and feedback models have been updated to be more physically realistic than those used in L18, but produce worse agreement in the star formation histories. This may indicate that there is additional relevant physics that needs to be included in our simulation.

Our model assumes that all stars above $8\,\mathrm{M}_\odot$ explode as SN. However, this assumption may not hold. Simulations of SN find that some progenitors collapse directly to a black hole without a SN explosion (Heger et al., 2003; Horiuchi et al., 2014; Pejcha & Thompson, 2015). If we were to include such scenarios in the feedback scheme, it would decrease the total energy and momentum from SN. We would therefore require a higher value of f_{boost} to obtain reasonable star formation rates. Additionally, changing the minimum progenitor mass for SN makes a large difference in the energy injected by SN (Keller & Kruijssen, 2022). We assume $M_{\mathrm{min}} = 8\,\mathrm{M}_\odot$, but this value is uncertain. Increasing it would decrease the number of SN, again requiring a higher f_{boost} to compensate.

5.5.3 Constraints on Star Formation Efficiency

Figure 5.17 shows the distribution of ϵ_{int} for the runs varying ϵ_{ff}. While we find a clear trend that decreasing ϵ_{ff} decreases ϵ_{int}, we can also examine the ratio $\epsilon_{\mathrm{int}}/\epsilon_{\mathrm{ff}}$. Figure 5.22 shows this ratio for the `Thelma & Louise` runs, which can be directly compared with Figure 8 of L18. For all values of ϵ_{ff} we consider, we find higher values of $\epsilon_{\mathrm{int}}/\epsilon_{\mathrm{ff}}$ than did L18. Two of the changes discussed in Section 5.3 are responsible. First, the modified SN feedback prescription delays SN feedback compared to L18, which results in higher ϵ_{int} (Figure 5.5). Second, the introduction of the virial criterion leads to higher ϵ_{int} at a given ϵ_{ff} (Figure 5.15). Combined, these two effects shift our distributions of $\epsilon_{\mathrm{ff}}/\epsilon_{\mathrm{int}}$ to higher values.

We still see the same trend with ϵ_{ff} as did L18, where higher values of ϵ_{ff} lead to

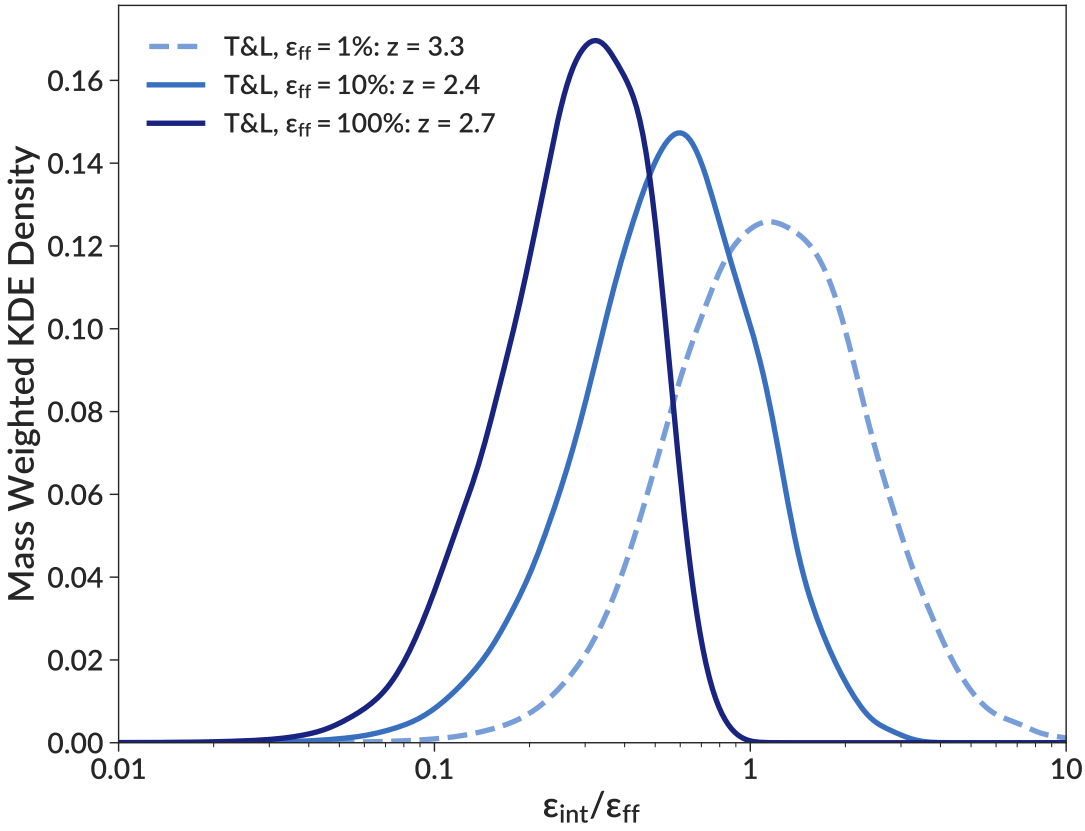

Figure 5.22 The distribution of $\epsilon_{\rm int}/\epsilon_{\rm ff}$ for different values of $\epsilon_{\rm ff}$ in the `Thelma & Louise` IC. All runs used $f_{\rm boost} = 5$, $f_{\rm HN,0} = 20\%$, and show all clusters formed before the redshift listed in the legend. We plot the $\epsilon_{\rm ff} = 1\%$ run with a dashed line as that run had many clusters that failed to finish forming.

smaller $\epsilon_{\rm int}/\epsilon_{\rm ff}$ ratios. Quantitatively, the mean value of this ratio drops from 1.15 to 0.57 to 0.30 for $\epsilon_{\rm ff} = 1\%$, 10%, and 100%, respectively. Conceptually, this ratio is proportional to the number of freefall times over which the cluster accreted material. As discussed in Section 5.3.4, lower values of $\epsilon_{\rm ff}$ lead to longer formation timescales, in accordance with this result.

As we have discussed throughout Section 5.3, the duration of a star formation episode is sensitive to cluster feedback and formation prescriptions. While difficult to constrain precisely, current observations indicate that the age spread within clusters is less than ≈ 6 Myr (see the compilation of age data in L18). These age spreads

can still be significantly larger than $t_{\rm ff}$. One example is the Orion Nebula Cluster (ONC), where star formation appears to have occurred over several freefall times (Da Rio et al., 2014; Caldwell & Chang, 2018; Kounkel et al., 2018). In contrast, simulations of individual molecular clouds generally show star formation ending after one $t_{\rm ff}$ (Grudić et al., 2022, e.g.). Our simulated age spreads, shown in Figure 5.10, are consistent with the observations for $\epsilon_{\rm ff} \geq 10\%$. We see a strong mass trend, but even for massive clusters the vast majority have age spreads smaller than 6 Myr. However, our results rule out $\epsilon_{\rm ff} = 1\%$, which has unphysically long age spreads for clusters of all masses.

The shape of the initial cluster mass function is another key observable. YMCs in the MW and nearby galaxies are found to follow the functional form of Schechter (1976), with a power law slope of -2 at the low-mass end (Portegies Zwart et al., 2010). Our mass functions have a positive power law slope at low mass, peak at a mass that depends on $\epsilon_{\rm ff}$ ($10^5\,{\rm M}_\odot$ for $\epsilon_{\rm ff} = 100\%$), then decline in a manner consistent with a power-law. In essence, our simulations are missing low-mass clusters. While our cluster formation algorithm only seeds clusters if they have an expected mass of $6 \times 10^3\,{\rm M}_\odot$, runs with $\epsilon_{\rm ff} = 100\%$ show the increasing mass function above this mass. This may indicate that $\epsilon_{\rm ff} = 100\%$ forms stars too efficiently, leading to too few low-mass clusters. However, $\epsilon_{\rm ff} \leq 10\%$ results in too few massive clusters, with no clusters projected to reach $z = 0$.

Lastly, $\epsilon_{\rm ff}$ has been measured in observations with several methods (Evans et al., 2014; Usero et al., 2015; Lee et al., 2016; Ochsendorf et al., 2017; Utomo et al., 2018). While the observations have somewhat different medians, uncertainties, and intrinsic scatter in $\epsilon_{\rm ff}$, a value of $\epsilon_{\rm ff} \approx 1\%$ is typical. However, we find that this value does not produce reasonable star cluster properties in our simulations. The timescales of cluster formation reach our algorithmically imposed limit of 15 Myr. Such timescales are in conflict with observations. Low values of $\epsilon_{\rm ff}$ also produce few massive clusters.

Even a value of $\epsilon_{\rm ff} = 10\%$ produces few clusters with high enough mass to reach $z = 0$ as GCs. Our simulations prefer higher values of $\epsilon_{\rm ff}$. Among the runs presented here, $\epsilon_{\rm ff} = 100\%$ produced the most realistic cluster properties, as it did in L18.

To compare with observations more directly, we postprocess the simulations to calculate an effective value of $\epsilon_{\rm ff}$ in a way analogous to how it is derived in observations. First we identify clusters that are actively forming in several simulation snapshots. Within a sphere of radius r centered on the cluster, we calculate the inferred value of $\epsilon_{\rm ff}$ as

$$\bar{\epsilon}_{\rm ff}(r) = \frac{\bar{t}_{\rm ff}(r)\,\dot{M}(<r)}{M_{\rm gas}(<r)} \tag{5.19}$$

where $\bar{t}_{\rm ff} \equiv \sqrt{3\pi/32G\bar{\rho}}$ is calculated using the mean density $\bar{\rho}$ within the sphere. In the rest of this section we will use $\bar{\epsilon}_{\rm ff}$ to refer to the inferred value from this procedure, while $\epsilon_{\rm ff}$ will refer to the value used in runtime of the simulation. To calculate $\dot{M}$, we use a procedure analogous to that used in studies that determine $\epsilon_{\rm ff}$ by counting young stellar objects (YSOs) to determine the star formation rate within a cloud (Evans et al., 2014; Heyer et al., 2016; Ochsendorf et al., 2017). These studies use YSOs to estimate the mass of recently formed stars, then divide it by the lifetime of the YSO phase typically set to a fixed time of 0.5 Myr. As we do not store the full accretion histories of simulated clusters, we cannot directly obtain the star formation rate over the last 0.5 Myr. Instead, we approximate it with the average star formation rate over the relevant timescale:

$$\dot{M} = \frac{M}{\max\left(\tau_{\rm spread}, 0.5\ {\rm Myr}\right)} \tag{5.20}$$

where M is the current mass of the actively forming cluster. For clusters with large age spreads this prescription gives the average star formation rate, while for clusters with short age spreads this matches the rate inferred observations assuming an 0.5 Myr timescale. We choose to use the cluster age spread rather than the full duration

as it more accurately reflects the timescale over which the bulk of cluster formation happens. The total $\dot{M}$ within a given sphere is the sum of $\dot{M}$ from all actively forming clusters in the sphere.

This calculation of $\bar{\epsilon}_{\rm ff}$ involves significant averaging both in time and space, compared to the application in simulation runtime. A typical local timestep at the highest refinement levels is 100–1000 yr, orders of magnitude shorter than even 0.5 Myr. Therefore, the finite difference calculation of the star formation rate $\dot{M}$ from Equation 5.1 is a much closer approximation to the true derivative than Equation 5.20. Considering spheres of radius $r > 5$ pc also introduces averaging of the stellar and gas mass on a larger scale than our adopted GMC radius. Both of these effects tend to shift $\bar{\epsilon}_{\rm ff}$ to smaller values than the input $\epsilon_{\rm ff}$.

In Figure 5.23 we show the distribution of values of $\bar{\epsilon}_{\rm ff}$ calculated for two choices of the averaging radius: 5 pc and 30 pc. The radius of 5 pc matches the GMC sphere actively participating in star formation. The inferred values peak at around 30% with large scatter but are significantly below the simulation input $\epsilon_{\rm ff} = 100\%$. The procedure to infer $\bar{\epsilon}_{\rm ff}$ uses the cluster formation timescale to average the star formation rate, which creates the wide spread and systematic shift. In addition, this procedure calculates $t_{\rm ff}$ and $M_{\rm gas}$ at one instant, which may not reflect typical conditions over the course of the cluster's growth.

Considering a larger sphere radius of 30 pc adds also spatial averaging. For an isolated cluster, increasing the size of the sphere would simply include more surrounding gas without increasing $\dot{M}$, leading to smaller inferred values of $\bar{\epsilon}_{\rm ff}$. However, we find that clusters often form in larger star-forming complexes with many clusters in close proximity of each other. Our choice of 30 pc corresponds to the typical radius of these star-forming complexes. Figure 5.23 shows that these complexes have a peak value of $\bar{\epsilon}_{\rm ff} \approx 10\%$, with less scatter than the values inferred on 5 pc scales. The lower mean value is due to the inclusion of more gas not participating in star formation,

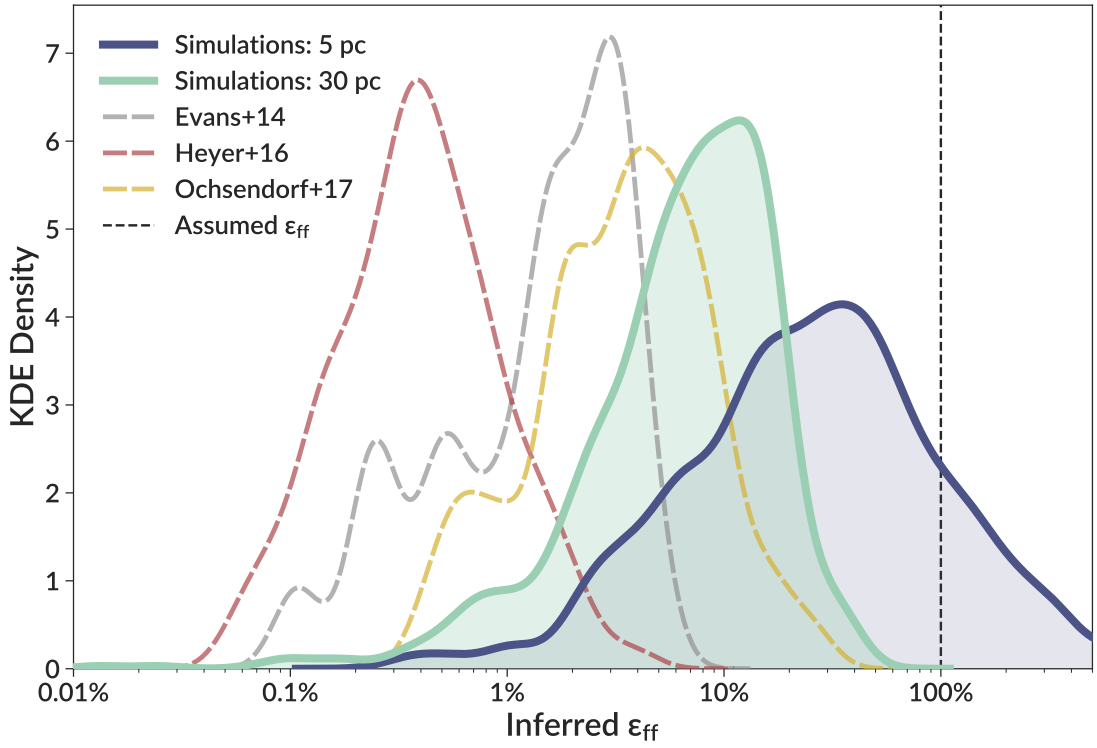

Figure 5.23 Kernel density estimation of the distribution of $\epsilon_{\rm ff}$ inferred from postprocessing simulated star clusters and from observations. We use a normalized Gaussian kernel with a width of 0.15 dex. The simulated clusters are from the run with the `Isolated MW` IC, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$. The 5 pc line shows $\bar{\epsilon}_{\rm ff}$ as inferred from the region actively participating in cluster formation, while the 30 pc line shows the value inferred for larger star-forming complexes.

while the decreased scatter comes from averaging together multiple clusters within each region.

This exercise shows that the inferred values of $\bar{\epsilon}_{\rm ff}$ are a factor of 10 lower than the simulation input. Still, for this run typical $\bar{\epsilon}_{\rm ff}$ are higher than those seen in observations. In Figure 5.23 we include observations from Evans et al. (2014), Heyer et al. (2016), and Ochsendorf et al. (2017), which all use the YSO method but do so on different scales. Evans et al. (2014) and Heyer et al. (2016) use clumps with typical radii of a few pc, while Ochsendorf et al. (2017) uses star-forming complexes with radii around 40 pc. Even with these differences of scale, all studies measure mean values of $\epsilon_{\rm ff}$ consistent with $\sim$1%. However, we note that we cannot make a direct comparison between these observations and our inferred values of $\bar{\epsilon}_{\rm ff}$. Each ingredient of the calculation of $\epsilon_{\rm ff}$ has systematic differences. The mass of recently formed stars is calculated differently, as we do not directly model the number of observable YSOs in each cluster. The timescales for calculating the star formation rate are also different, as many of our clusters have $\tau_{\rm spread}$ longer than the 0.5 Myr used in observations. Lastly, detailed modeling of CO and HCN abundances and ionization states is needed to calculate $M_{\rm gas}$ exactly as is done in observations. To resolve these differences would require a further analysis in the simulation runtime. Nevertheless, Figure 5.23 demonstrates that the discrepancy with observations is substantially smaller than appears from a straightforward comparison with the simulation input.

5.5.4 Failed Cluster Formation

In Section 5.3.4 we showed that in some runs with low $\epsilon_{\rm ff}$, clusters fail to finish formation before it is automatically ended at 15 Myr. In this section, we investigate the reasons for these failed clusters.

We find that no runs with the high value of $\epsilon_{\rm ff} = 100\%$ have failed cluster formation, all runs with the low value of $\epsilon_{\rm ff} = 1\%$ fail, and among the runs with the

intermediate value $\epsilon_{\rm ff} = 10\%$, only the run using the `Isolated MW` IC and $f_{\rm boost} = 1$ failed. All other runs with $\epsilon_{\rm ff} = 10\%$ used higher $f_{\rm boost}$ and did not fail.

These trends are due to an interplay between $\epsilon_{\rm ff}$ and $f_{\rm boost}$. When $\epsilon_{\rm ff}$ is low, cluster formation progresses slowly, leaving significant amounts of gas. We find low values of $\epsilon_{\rm int}$ for low $\epsilon_{\rm ff}$ (Figure 5.17), meaning that at the end of cluster formation, only a small fraction of gas has been turned into stars. This applies in the midst of cluster formation, too. We examine the gas densities of the host cells of clusters as they form and find that for lower values of $\epsilon_{\rm ff}$ there is more gas near the cluster at a given time after the beginning of cluster formation, meaning that GMCs are more massive with low values of $\epsilon_{\rm ff}$. In addition, the slower star formation with low $\epsilon_{\rm ff}$ leads to fewer stars to provide feedback. When SNe begin, they must first disperse the gas within the cluster. Higher values of $f_{\rm boost}$ make this process more efficient. Therefore, higher values of $f_{\rm boost}$ lead to shorter timescales for cluster formation when $\epsilon_{\rm ff}$ is low. In contrast, when $\epsilon_{\rm ff} = 100\%$, clusters consume a high fraction of the gas within their GMC. SN feedback of any $f_{\rm boost}$ is able to clear the smaller amounts of gas that remain.

In Figure 5.24 we illustrate this effect by presenting instantaneous distribution of the molecular gas density within galaxies with different combinations of $\epsilon_{\rm ff}$ and $f_{\rm boost}$. Runs with failed cluster formation have distributions that extend to higher densities than runs without failed cluster formation. As feedback cannot terminate star formation, gas continues to accrete onto the GMC, increasing the density. Of particular note is the gas at densities higher than the star formation threshold. The total gas number density must be greater than 10^3 cm^{-3} with a molecular fraction of 0.5, giving a minimum molecular number density of 500 cm^{-3}. Above this threshold, the failed runs have significantly more gas than runs that successfully terminate star formation.

The prescription for SN feedback also contributes to why these runs had failed

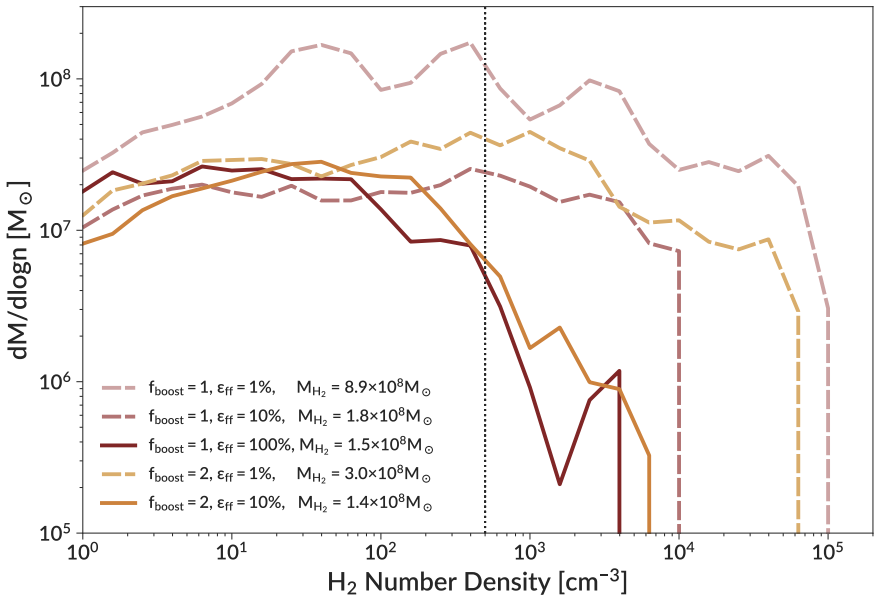

Figure 5.24 The distribution of cell mass-weighted molecular gas densities for different combinations of $\epsilon_{\rm ff}$ and $f_{\rm boost}$ in the `Isolated MW` IC. All runs used $f_{\rm HN,0} = 0$ and show the gas within 10 kpc of the galactic center. Dashed lines indicate runs with failed cluster formation. Runs with $f_{\rm boost} = 1$ are shown at $z = 1.56$, while those with $f_{\rm boost} = 2$ are shown at $z = 2.57$. The dotted line at 500 cm^{-3} indicates the minimum density for star formation.

cluster formation while the runs of L18 did not. In the 15 Myr timescale for cluster formation, the prescription of L18 injects significantly more energy than the updated model (see Figure 5.1). For low hypernova fractions typical of most clusters, the new prescription injects only 26% of the energy of L18 within the first 15 Myr, increasing to about 50% after 40 Myr. This is exacerbated by the lower $f_{\rm boost}$ used in the updated runs. The total momentum injected by SN feedback during the 15 Myr of cluster formation can be more than an order of magnitude lower than in L18. For low values of $\epsilon_{\rm ff}$, this results in SN feedback being unable to disperse the GMC.

The timing of SN also contributes to failed cluster formation. Compared to L18, SN start later in the new prescription due to the stellar lifetimes chosen (Figure 5.1). These runs also use the average approach for determining cluster feedback timing, as they were run before the hybrid approach was finalized. This average approach further delays the onset of SN (Figure 5.2). In addition, we find that runs with lower

$\epsilon_{\rm ff}$ have later average times of cluster formation, $t_{\rm ave}$, meaning that SN is delayed even further in these runs. These delays in the onset of SN gives the gas outside the GMC more time to accrete onto the GMC, leading to higher gas masses that SN feedback then needs to disperse. This combines with the effect described in the previous paragraphs to make GMCs more difficult to disperse for lower $\epsilon_{\rm ff}$, further explaining why we find that $f_{\rm boost}$ affects the timescales of cluster formation for low $\epsilon_{\rm ff}$ but not for $\epsilon_{\rm ff} = 100\%$.

To summarize our understanding of why massive star clusters fail to finish forming when $\epsilon_{\rm ff} \leq 10\%$, lower values of $\epsilon_{\rm ff}$ turn gas into stars at a slower pace. At a given time after the beginning of cluster formation this results in fewer stars, and therefore less stellar feedback, embedded in a more massive GMC. The onset of SN is delayed compared to L18 due to our choice of stellar lifetimes (see Figure 5.1), and then is delayed further after accounting for the age spread within the cluster. This allows more material to accrete onto the GMC, making it even more difficult for feedback to disperse. Once SN feedback starts, the updated feedback prescription injects less momentum than L18. The new prescription has fewer SN and is further exacerbated if low values of $f_{\rm boost}$ are chosen. Although lower values of $f_{\rm boost}$ produce more reasonable star formation rates for $z > 2$, these low values fail to provide enough feedback to disperse GMCs when $\epsilon_{\rm ff}$ is low. This may indicate that another source of feedback is needed at early times to help disperse GMCs.

5.6 Conclusions

We have described improvements to the implementation of star cluster formation and feedback in the ART code. We introduced a new criterion for the seeding of cluster particles, requiring the star-forming gas to be gravitationally bound. We also implemented a new prescription for the initial bound fraction of clusters based on simulations of individual GMCs. We added runtime tracking of C, N, O, Mg, S,

Ca, and Fe, with enrichment coming from SNIa, SNII, stellar winds, and AGB stars. We updated the SN feedback prescriptions significantly. We now implement SN as discrete events, with rates based explicitly on the stellar lifetimes and IMF. We also explored effects of hypernovae, which inject more energy and have different elemental yields. Lastly, we improved our prescription for the timing of SN to account for the age spread of stars within a cluster.

With these code updates, we ran 20 simulations using the initial condition from L18 and 9 simulations using two Local Group-like ICs from the ELVIS project. These runs have a range of parameters, including variations in $\epsilon_{\rm ff}$, $f_{\rm HN,0}$, $f_{\rm boost}$, and the timing of SN feedback. We explored how these parameters affect the properties of galaxies as well as the populations of star clusters within them. Our results are summarized as follows.

• Delaying the onset of SN (without changing the total energy injection) results in longer formation timescales for massive clusters and higher $\epsilon_{\rm int}$ (Figures 5.4, 5.5), but does not significantly change the galaxy star formation rate.

• Higher values of the momentum boosting factor for SN greatly decrease the galactic star formation rate (Figures 5.6, 5.7). While no value of $f_{\rm boost}$ can reproduce the abundance matching expectation for the full redshift range explored here ($z > 1.5$), we find that the range $1 < f_{\rm boost} < 3$ produces reasonable agreement for $z > 2$. Higher values of $f_{\rm boost}$ decrease the total stellar mass by decreasing the number of low-mass clusters that form, without changing the number of massive clusters (Figure 5.9).

• The hypernova fraction $f_{\rm HN,0}$ makes little difference to galaxy or cluster properties (Figures 5.6, 5.9). The strong decrease in $f_{\rm HN}$ with metallicity (Equation 5.11) results in limited change in the total energy injected by SN (Figures 5.1, 5.8).

• The local star formation efficiency per freefall time does not have a strong impact on the galactic star formation rate (Figure 5.13). However, it strongly changes cluster properties. Higher values of $\epsilon_{\rm ff}$ lead to more massive clusters (Figures 5.11, 5.12),

shorter timescales for cluster formation (Figure 5.10), higher initial bound fractions (Figure 5.16), and higher $\epsilon_{\rm int}$ (Figure 5.17).

• Adding the virial parameter criterion to require star-forming gas be gravitationally bound produces more high-mass clusters (Figure 5.14), longer timescales for cluster formation, and higher $\epsilon_{\rm int}$ (Figure 5.15).

• In runs with low values of $\epsilon_{\rm ff}$, we find a population of clusters that fail to finish forming after 15 Myr. Low values of $\epsilon_{\rm ff}$ form stars slowly, leaving massive GMCs that are difficult for feedback to disperse, especially with low values of $f_{\rm boost}$.

• We present the evolution of the observable mass function of clusters at various redshifts (Figure 5.20). Most massive clusters form at high redshifts when the star formation density is high, with low-mass clusters dominating in quiescent epochs.

• We analytically extrapolate the dynamical disruption of clusters from the last available output to $z = 0$ (Figure 5.20). We can match the observed mass function of MW GCs by assuming a high value for the cluster disruption rate. The surviving clusters also match the age-metallicity relation of MW GCs (Figure 5.21).

• Among the values of $\epsilon_{\rm ff}$ we explored, only $\epsilon_{\rm ff} = 100\%$ can match the MW GC mass function. Runs with $\epsilon_{\rm ff} = 1\%$ produces clusters with unphysically long age spreads (Figure 5.10), and runs with $\epsilon_{\rm ff} = 10\%$ produce too few high-mass clusters (Figures 5.11, 5.12, 5.19).

This exploration emphasizes the importance of well-calibrated subgrid models for modeling star clusters in simulations of galaxy formation. Some modeling choices, such as the optimal value of $\epsilon_{\rm ff}$, whether to enforce a virial criterion when seeding star clusters, and different prescriptions for the timing of SN feedback all affect the resulting cluster populations without significantly impacting global galaxy properties. A successful model of star formation and feedback in simulations must be able to reproduce not only galaxy-scale properties, but also the small-scale properties of individual star clusters.

CHAPTER VI

Future Directions and Conclusions

6.1 Summary

In this dissertation, I have investigated the formation and evolution of star clusters in both simulations and observations. In this chapter I will summarize my work then describe future directions.

In Chapter II, I present my measurements of the radii of over 6000 young star clusters in nearby galaxies. This is the largest catalog of young star cluster radii currently available. To effectively calculate these radii, I implemented a pipeline with several features to make it robust against contaminating objects, allowing it to produce reliable radii for as many clusters as possible. I find that most galaxies share a common radius distribution, with a peak at around 3 pc. However, we find two galaxies that have different distributions. While we do not currently understand the reason, this will be a fruitful avenue for future research. One of the most important results from this work is that we find a clear mass-radius relation of the form $R_{\rm eff} \propto M^{0.24}$, with a large intrinsic scatter of 0.25 dex. This is the clearest detection of such a relation thus far, enabled by the large catalog of radii. This relation provides constraints on models of clusters formation and gives the initial conditions for cluster disruption. We also calculate cluster densities, finding a roughly lognormal surface density distribution peaked at 100 $\rm M_{\odot} pc^{-2}$.

Turning to my work with cosmological simulations, in Chapter III I present my work analyzing the nuclear star clusters (NSCs) of an existing galaxy formation simulation. I developed a method to find NSCs in simulations, then analyzed their structural properties such as mass, size, and shape. I also examined the chemical properties of NSCs, finding that they are systematically more metal-rich than their host galaxies. Finally, I used these simulations to test the hypothesis that nuclear star clusters in dwarf galaxies may become the progenitors of the anomalous globular clusters seen in the MW. I find that NSCs contain spreads in iron abundance and age similar to those in anomalous GCs. This lends support to the idea that these GCs may be remnant NSCs of disrupted dwarf galaxies.

In Chapter IV I begin describing my work to develop and run a new suite of simulations. In this chapter I focus on the initial conditions (ICs) used in the cosmological simulations. While zoom-in ICs are very useful for studying the evolution of galaxies in their cosmological context, they can be difficult for simulation codes like ART to handle computationally. Specifically, the small fraction of the box occupied by the galaxies of interest makes it difficult for the load balancing algorithm to evenly divide the work. This makes simulation runtime impractically slow. As one of the long-term goals of the new simulation suite is to reach $z = 0$, improving the runtime of the simulation is important. In this chapter, I present a method to modify the grid structure of zoom-in ICs. I decrease the box size while increasing the resolution of the root grid cells. As the load balancing algorithm of ART operates on these root grid cells, increasing their number allows for a more even distribution of work among computing nodes, increasing computational efficiency. This method can produce a speedup of nearly a factor of two. Importantly, this is done without changing the growth histories of the galaxies in the zoom region. This enables my simulations to progress faster than would otherwise be possible, laying the foundations for the simulations I present in Chapter V.

I begin Chapter V with a description of the improvements I made to the model of cluster formation in cosmological simulations of Li et al. (2017) and Li et al. (2018). I add a criterion requiring that star forming gas is gravitationally bound before star clusters are seeded. I significantly update the stellar feedback prescription, making SN rates directly based on the IMF and stellar lifetimes. Some fraction of these SN can explode as hypernovae, injecting more energy into the surrounding ISM. Lastly, I updated the timing of SN feedback to better reflect the intrinsic age spreads within clusters.

I then ran a large suite of 29 simulations, 9 of which use the updated Local Group-like initial conditions as described in Chapter IV. While I did not reach $z = 0$ with these simulations yet, in the future these runs will be able to progress to $z = 0$ and make direct comparisons with the MW cluster population. I first perform a differential analysis with the goal of understanding how each of the updates affects star cluster populations. I systematically explore a range of parameters including the star formation efficiency per freefall time $\epsilon_{\rm ff}$, the momentum boost from supernovae $f_{\rm boost}$, and the initial hypernova fraction $f_{\rm HN,0}$. Some modeling choices, such as the value of $\epsilon_{\rm ff}$, whether or not to enforce a virial criterion when seeding star formation, and different prescriptions for the timing of SN feedback all change the resulting cluster populations without significantly impacting global galaxy properties. Observations of clusters are the only way to constrain these parameters. This exploration emphasizes the importance of well-calibrated subgrid models when modeling cluster formation in simulations of galaxy evolution.

I also examine the evolution of star cluster populations in these simulations over cosmic time. I find that most massive clusters form at high redshifts ($z > 4$) when the star formation rate is high, while low-mass clusters are dominant in more quiescent epochs. By analytically extrapolating cluster disruption to $z = 0$, I make a preliminary comparison of the simulated clusters to the MW. I find good agreement

with both the GC mass function and age-metallicity relation.

6.2 Future Work with Cluster Radii

In Chapter II, I presented my measurements of the radii of young star clusters. However, only the result that the cluster radius distribution is common among galaxies examined how cluster properties change with environment. This is an area with many potential avenues for investigation.

One simple investigation would be to correlate cluster properties with their position within galaxies. As tidal fields are stronger in denser regions near the center of galaxies, tidal truncation may result in a trend where clusters at larger galactocentric radii have larger radii. If this trend does not exist, it indicates that clusters do not fill their tidal radii. Results in the literature are mixed about the observed trend with galactocentric radius (Bastian et al., 2012; Ryon et al., 2015; Sun et al., 2016; Ryon et al., 2017). Using the large sample size provided by my cluster catalog would give the most definitive result yet.

Cluster properties can also be correlated with the properties of their host galaxy. For example, the analytical model of Choksi & Kruijssen (2021) predicts that the cluster radius depends on the surface density of starforming gas. This naturally arises out of the star formation process, affecting clusters at birth. By compiling gas surface density profiles for the galaxies in my sample, this can be directly tested across a wide range of galaxies. This investigation may also shed light on what is happening with the two galaxies with discrepant radius distributions. Similarly, the cluster disruption rate should depend on the host galaxy, as higher density causes increased disruption. Both cluster birth and evolution can be further tested by examining cluster radius as a function of host galaxy properties.

Finally, detailed analytical models of cluster evolution are needed to interpret the trends of cluster radius with age. My results show less evolution of the cluster

radius with time than is expected by many models. The toy models I created were unable to reproduce the key aspects of the observed distributions, and neither was the full model of Gieles & Renaud (2016). These simple models must be refined to match observations, both through improved constraints on cluster initial properties and more detailed modeling of tidal disruption within galactic contexts.

6.3 Future Work with Simulations

In this dissertation I have only scratched the surface of what can be done with my suite of simulations. I'll sketch a few potential avenues of exploration in this section. Most obviously, I was only able to run the Local Group simulations to $z \approx 2-3$. Reaching $z = 0$ would enable direct comparisons with the Milky Way. Instead of needing an analytical extrapolation of cluster disruption to obtain cluster masses at $z = 0$, the simulation will provide full self-consistent calculations of these masses. This would enable a direct comparison with the MW GC population. The GC mass function, metallicity distribution, and kinematics can be investigated. Taking the selection of surviving clusters and tracking them back to their formation sites would also reveal what conditions are needed at high redshift to form GCs. One particularly interesting application of $z = 0$ outputs would be an investigation of the kinematics of in-situ and ex-situ GCs. Many recent papers have attempted to use kinematics and/or chemistry to distinguish in-situ from ex-situ clusters in the MW (Massari et al., 2017; Helmi et al., 2018; Koppelman et al., 2019; Massari et al., 2019). In the simulation we know which clusters are in-situ and ex-situ, and can directly test proposed methods of reconstructing these populations.

These simulations will also be useful for tracking galactic chemical enrichment. I have implemented the runtime tracking of C, N, O, Mg, S, Ca, and Fe, but so far have not investigated the chemical properties of galaxies systematically. In particular, the high spatial and temporal resolution (3–6 pc, 10–50 Myr) and the state

of the art physics allows for physically accurate modeling of individual galaxies and their evolution over time, which is a powerful tool for disentangling the various processes that act on a galaxy's chemical content (Vincenzo & Kobayashi, 2018; Torrey et al., 2019). Do galaxies move parallel to the mean mass-metallicity relation (MZR) with some intrinsic scatter, or is the scatter due to galaxies taking more complex paths? Tracking quantities proposed as possible additional parameters in a fundamental mass-metallicity relation such as star formation rate or gas mass will show exactly how these parameters shape the MZR. This will enable tests of proposed analytic galaxy evolution models that make predictions regarding the mass-metallicity relation (Lilly et al., 2013; Forbes et al., 2014). In addition, my simulations are the first of their kind (to my knowledge) to include hypernovae. These did not have significant impacts on galactic structure or cluster properties, but their different chemical yields may change the metallicity evolution of galaxies.

Finally, integrating my radius measurements into the dynamical evolution calculations of the simulations would be a big step forward. Currently the disruption calculations do not depend on cluster radius, and it is not modeled in the simulation. There are several ways the radius could be implemented in the simulation, and I willl sketch a few ideas to close out this section. First, radius could be directly modeled in runtime. At cluster birth, a radius is assigned to the cluster by randomly drawing from the observed cluster radius distribution at the appropriate mass. This could be improved if future work on observed star cluster radii reveals a correlation with local properties such as the surface density of the gas. Once the radius is assigned, it would self-consistently evolve with time, accounting for both internal evolution and any tidal shocks the cluster experiences.

Alternatively, the radius information could be included when post-processing simulation outputs. The simulations output the tidal tensor at each global timestep (every 10-50 Myr), so the tidal evolution of each cluster can be reconstructed. Clus-

ters could be assigned radii in the same way as they would be assigned during runtime, but would then be evolved analytically using the tidal tensor of the cluster as directly resolved in the simulation. One advantage of this method is that multiple realizations of the cluster population could be created by redoing the initial random draws that assign cluster radii. Analytically evolving the tidal evolution of this new realization is much cheaper than rerunning the whole simulation. This would allow an exploration of how the cluster population statistics are sensitive to stochasticity, but also how individual clusters evolve differently with different assigned radii.

APPENDICES

APPENDIX A

Methods for Fitting Mass-Radius Relation

In the main text, we use the orthogonal fitting method described in sections 7 and 8 of Hogg et al. (2010) (hereafter H10), which we summarize here. In evaluating the Gaussian likelihood of each data point given a linear relation (with parameters of slope m and intercept b), we use the displacements and variances projected perpendicular to the line being evaluated. The projected displacement is given by Equation 30 of H10, and can also be written as:

$$\Delta_i = \frac{y_i - (mx_i + b)}{\sqrt{1 + m^2}} \tag{A.1}$$

where in our case x is the log of the mass, y is the log of the effective radius, $\beta = m$, and we can calculate b given β and our pivot point R_4.

The projected variance is given by Equation 31 of H10. In our case, the mass and radius errors are independent, so the off-diagonal terms of the covariance matrix are zero, allowing us to simplify that expression:

$$\sigma_i^2 = \frac{m^2\sigma_{x,i}^2 + \sigma_{y,i}^2}{1 + m^2} \tag{A.2}$$

Then we add an intrinsic scatter $\sigma_{\rm int}$ orthogonal to the line to the data variance,

giving the likelihood for a single data point of

$$\mathcal{L}_i = \frac{1}{\sqrt{2\pi\left(\sigma_i^2 + \sigma_{\rm int}^2\right)}} \exp\left(-\frac{\Delta_i^2}{2(\sigma_i^2 + \sigma_{\rm int}^2)}\right) \tag{A.3}$$

The total data likelihood is the product of this over all data points. In the main text, we maximize this likelihood to produce final parameter values.

This method does not incorporate any selection effects. As these do exist in the LEGUS sample, we implemented an additional method to attempt to incorporate those selection effects. We use a hierarchical Bayesian model, following Kelly (2007). Each cluster has its observed mass and radius ($M_{{\rm obs},i},\ R_{{\rm eff},i}$) with corresponding unobserved true quantities ($m_i,\ r_i$). We use the same relation as the main text (Equation 2.20), but define it using the unobserved true quantities rather than the observed values:

$$\hat{r}_i(m_i) = r_4 \left(\frac{m_i}{10^4\,{\rm M}_\odot}\right)^\beta \tag{A.4}$$

such that the normalizing factor r_4 is the underlying effective radius at $10^4\,{\rm M}_\odot$. In addition, we include an intrinsic lognormal scatter $\sigma_{\rm int}$.

For a given cluster, our data likelihood takes the general form:

$$\begin{aligned} P(R_{{\rm eff},i}, M_{{\rm obs},i}, r_i, m_i | r_4, \beta, \sigma_{\rm int}) = & P(R_{{\rm eff},i}|r_i) \\ & \times P(M_{{\rm obs},i}|m_i) \\ & \times P(r_i|m_i, r_4, \beta, \sigma_{\rm int}) \\ & \times P(m_i) \end{aligned} \tag{A.5}$$

where the first two terms are the likelihoods of the observed values given the unobserved true values (which we treat as independent), the third term is the mass-radius relation, and the final term is the prior on the true mass. We model the radius

distribution at a given mass as a lognormal distribution:

$$P(r_i|m_i, r_4, \beta, \sigma_{\rm int}) = \frac{1}{\sigma_{\rm int}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\log r_i - \log \hat{r_i}(m_i)}{\sigma_{\rm int}}\right)^2\right] \tag{A.6}$$

where $\log \hat{r_i}(m_i)$ is from Equation 2.20. The normalizing factor here is important, as it includes a variable of interest $\sigma_{\rm int}$.

We treat the mass and radius measurement errors as independent lognormal variables, with a width equal to the symmetrized observational uncertainties:

$$P(R_{{\rm eff},i}|r_i) \propto \exp\left[-\frac{1}{2}\left(\frac{\log r_i - \log R_{{\rm eff},i}}{\sigma_{R_{{\rm eff},i},{\rm err}}}\right)^2\right] \tag{A.7}$$

$$P(M_{{\rm obs},i}|m_i) \propto \exp\left[-\frac{1}{2}\left(\frac{\log m_i - \log M_{{\rm obs},i}}{\sigma_{M_{{\rm obs},i},{\rm err}}}\right)^2\right] \tag{A.8}$$

This allows us to analytically marginalize over the unobserved radius r_i.

$$\begin{aligned} P(R_{{\rm eff},i}|m_i, r_4, \beta, \sigma_{\rm int}) =& \int P(R_{{\rm eff},i}|r_i) P(r_i|m_i, r_4, \beta, \sigma_{\rm int}) dr_i \\ =& \frac{1}{\sqrt{2\pi\left(\sigma_{\rm int}^2 + \sigma_{R_{{\rm eff},i},{\rm err}}^2\right)}} \\ & \times \exp\left[-\frac{1}{2}\frac{\left(\log R_{{\rm eff},i} - \log \hat{r_i}(m_i)\right)^2}{\sigma_{\rm int}^2 + \sigma_{R_{{\rm eff},i},{\rm err}}^2}\right] \end{aligned} \tag{A.9}$$

We also need to include the selection effects. There are two key selection variables: radius and V band absolute magnitude.

To prevent contamination from unresolved sources, LEGUS selects clusters by examining the concentration index (CI), the magnitude difference between a 3 pixel and 1 pixel apertures. A hard boundary is drawn: anything above this is a cluster, anything below this is a star (Adamo et al., 2017). The V band absolute magnitude cut is simple: selected clusters have absolute V band magnitude brighter than -6. Unfortunately, this causes selection effects as a function of both mass and age, as

dying massive stars go away, making clusters fade with age. To properly account for this, we need to include both age and V band magnitude into our analysis. In what follows we will use T, τ for observed and true ages, respectively, and V, ν for observed and true V band absolute magnitudes.

To renormalize our likelihood function, we need to determine the likelihood of a given cluster of a given true mass and age being selected:

$$\Phi_i(m_i, \tau_i, r_4, \beta, \sigma_{\rm int}) = \int f(R_{\rm eff}) f(V) P(R_{\rm eff}|m_i, r_4, \beta, \sigma_{\rm int}) P(V|m_i, \tau_i) dR_{\rm eff} dV \quad \text{(A.10)}$$

where $f(R_{\rm eff}), f(V)$ are the selection functions. For V, this is a simple step function:

$$f(V) = \begin{cases} 1 & \text{for } V < -6 \\ 0 & \text{for } V \geq -6 \end{cases} \quad \text{(A.11)}$$

For radius, this selection function is more complicated. Adamo et al. (2017) made a first attempt at this in the NGC 628c field, finding that clusters with effective radii of 2 pc were entirely recovered, while those with 1 pc were recovered roughly 50% of the time. For this test, we simply represent the selection probability as

$$f(R_{\rm eff}) = \max\left(\frac{R_{\rm eff}}{0.05\,\text{arcsec}}, 1\right) \quad \text{(A.12)}$$

where 0.05 arcsec is roughly 2 pc at the distance of NGC 628. This gives a 100% selection at 2 pc and 50% selection at 1 pc. While this functional form is likely inaccurate, quantifying the selection effects in more detail is beyond the scope of this paper.

The final two terms in Equation A.10 are the likelihood of a given radius being selected at that mass (given by Equation A.9) and the likelihood of observing a

given V band magnitude given a cluster's mass and age. We obtain these using the Yggdrasil models matching those used in LEGUS (Zackrisson et al., 2011). We can then represent the V band term as a normal distribution:

$$P(V_i|m_i,\tau_i) \propto \exp\left[-\frac{1}{2}\left(\frac{V_i-\nu(m_i,\tau_i)}{\sigma_{V,\mathrm{err}}}\right)^2\right] \tag{A.13}$$

where $\nu(m_i,\tau_i)$ is the expected V band magnitude obtained from Yggdrasil, and $\sigma_{V,\mathrm{err}}$ is the observed uncertainty in V band magnitude. Note that since this term is for an arbitrary cluster with true mass m and true age τ, we use typical errors of 0.04 magnitudes.

We use this selection probability to renormalize the likelihood. Φ_i is calculated separately for each cluster and used to renormalize that cluster's likelihood. In addition, we need to incorporate τ into the likelihood as it enters the selection effects:

$$P(R_{\mathrm{eff},i},M_{\mathrm{obs},i},T_i,m_i,\tau_i|r_4,\beta,\sigma_{\mathrm{int}}) = \tag{A.14}$$

$$\Phi_i^{-1}(m_i,\tau_i,r_4,\beta,\sigma_{\mathrm{int}}) \tag{A.15}$$

$$\times P(R_{\mathrm{eff},i}|m_i,r_4,\beta,\sigma_{\mathrm{int}}) \tag{A.16}$$

$$\times P(M_{\mathrm{obs},i}|m_i)P(T_i|\tau_i)P(m_i)P(\tau_i) \tag{A.17}$$

We treat the age in the same way we treat the mass and radius errors, as a lognormal distribution with a width equal to the symmetrized observational error:

$$P(T_i|\tau_i) \propto \exp\left[-\frac{1}{2}\left(\frac{\log\tau_i-\log T_i}{\sigma_{T,\mathrm{err}}}\right)^2\right] \tag{A.18}$$

As this likelihood is for one cluster, the total likelihood is the product of the likelihood for all clusters. We also use Bayes' Theorem to turn this into posterior

likelihoods on our parameters:

$$P(r_4, \beta, \sigma_{\rm int} | R_{{\rm eff},i}, M_{{\rm obs},i}, m_i) \propto P(r_4)P(\beta)P(\sigma_{\rm int}) \times \prod_i P(R_{{\rm eff},i}, M_{{\rm obs},i}, m_i | r_4, \beta, \sigma_{\rm int}) \quad \text{(A.19)}$$

We use flat priors on all parameters. The slope β is uniform between -1 and 1, the normalization r_4 is uniform between 0.01 and 100 pc, and the intrinsic scatter σ is uniform between 0 and 1 dex.

We sample this posterior distribution using the *emcee* implementation of MCMC (Foreman-Mackey et al., 2013). This allows us to easily marginalize over m_i and τ_i in postprocessing of the MCMC chain.

As this implementation requires evaluating the integral in Equation A.10 for each cluster at each step of the MCMC chain, it is computationally expensive, and scales with the number of clusters. We therefore tested this method using a random sample of only 100 clusters. Figure A.1 shows the fit parameters for regular least squares, our fiducial orthogonal least squares method, and several variations of this method with different selection functions enabled.

This hierarchical Bayesian MCMC model produced larger error bars, but otherwise the results are consistent with those of the orthogonal fit. Additionally, removing the selection function terms from the fit does not change the result. Because of this, we decided to not use the hierarchical Bayesian method and instead use the simpler orthogonal fit described at the beginning of this appendix. Another reason to use the orthogonal fit is that we do not know the true selection effects in LEGUS, which would be needed to do a proper analysis of their impact. A full accounting of these is beyond the scope of this paper, making the functional form we assumed for the selection effects overly simplistic. We therefore choose to use the orthogonal fit as our method of choice in this paper.

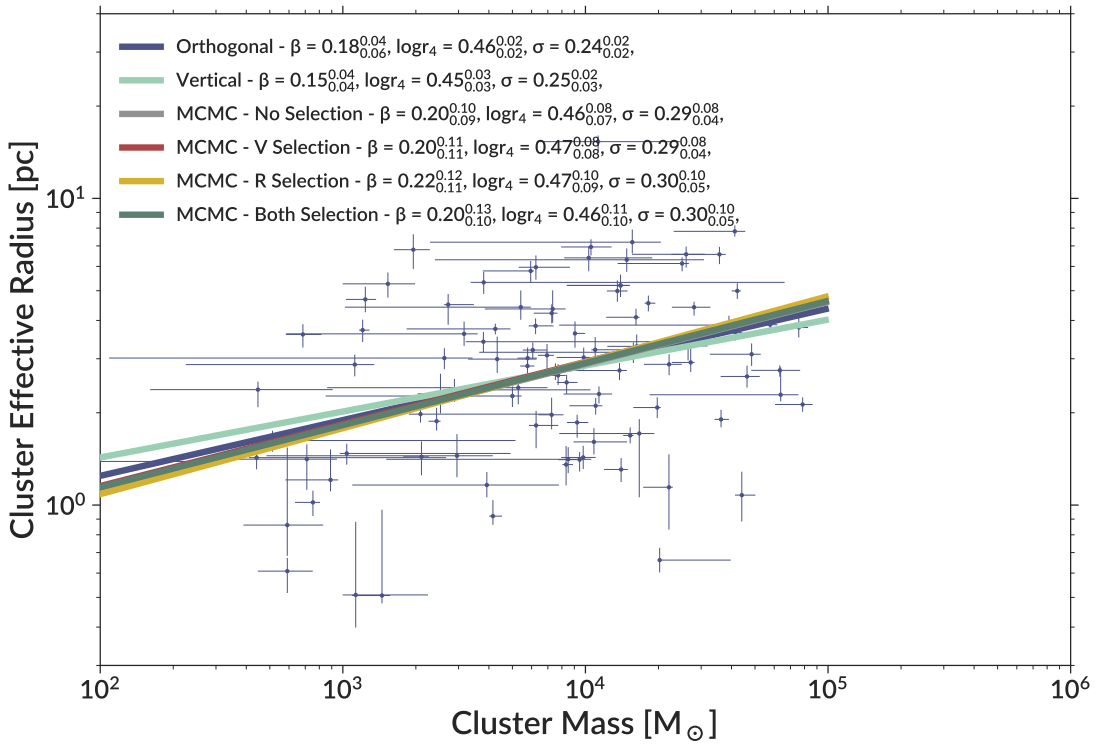

Figure A.1 The mass-radius relation fitted by several methods for a random sample of 100 clusters.

APPENDIX B

Hydrodynamics in the ART code

When updating from the version of the ART code used in L18, we changed the model of how internal energy is calculated in the presence of subgrid turbulence. The hydro solver independently tracks total energy, thermal energy, and energy of unresolved subgrid turbulence. The thermal energy and subgrid turbulence are assumed to evolve adiabatically (other than energy injection from sources such as stellar feedback). As these are calculated independently, there is no initial restriction for the sum of thermal, kinetic, and turbulent energies to equal the total. As the adiabatic assumption is not always correct for thermal energy (particularly in shocks), the new version calculates the thermal energy as $E_{\rm th} = E_{\rm tot} - E_{\rm kinetic} - E_{\rm turb}$. This energy synchronization allows for shocks to transfer energy from kinetic to thermal, as should happen. The adiabatic assumption is only used in cases where the gas is highly supersonic, such that $E_{\rm tot} \approx E_{\rm kinetic}$. In this case, the subtraction would be susceptible to numerical errors, so we revert to the adiabatic assumption. In the old version of the code, which *always* relied on the adiabatic assumption, shocks were not treated properly and energy that should have been transferred from kinetic to thermal was simply lost. This is visualized in the top row of Figure B.1, where we show the phase diagram of gas within the virial radius at $z = 13.3$ before stars have formed. The hydrodynamic scheme of L18 follows what is expected for pure adiabatic compression,

while the new scheme shows gas being heated by virial shocks.

While the newer version of the code is better physically motivated, it significantly changed the structure of modeled galaxies. We find large differences in temperature distributions of the gas. The bottom panel of Figure B.1 shows the phase diagram of gas within the virial radius at $z = 1.5$. Here the run with the updated hydrodynamics has significantly more hot, low-density gas in the halo. This hot gas prevented cold gas from accreting onto the disc, effectively reducing star formation. We show this star formation in Figure B.2 using test runs that vary both the hydro and feedback schemes. We test the stellar feedback model presented in this paper as well as the model used by L18. All runs use $f_{\rm boost} = 5$, yet runs with the new hydro scheme produce dramatically lower star formation rates.

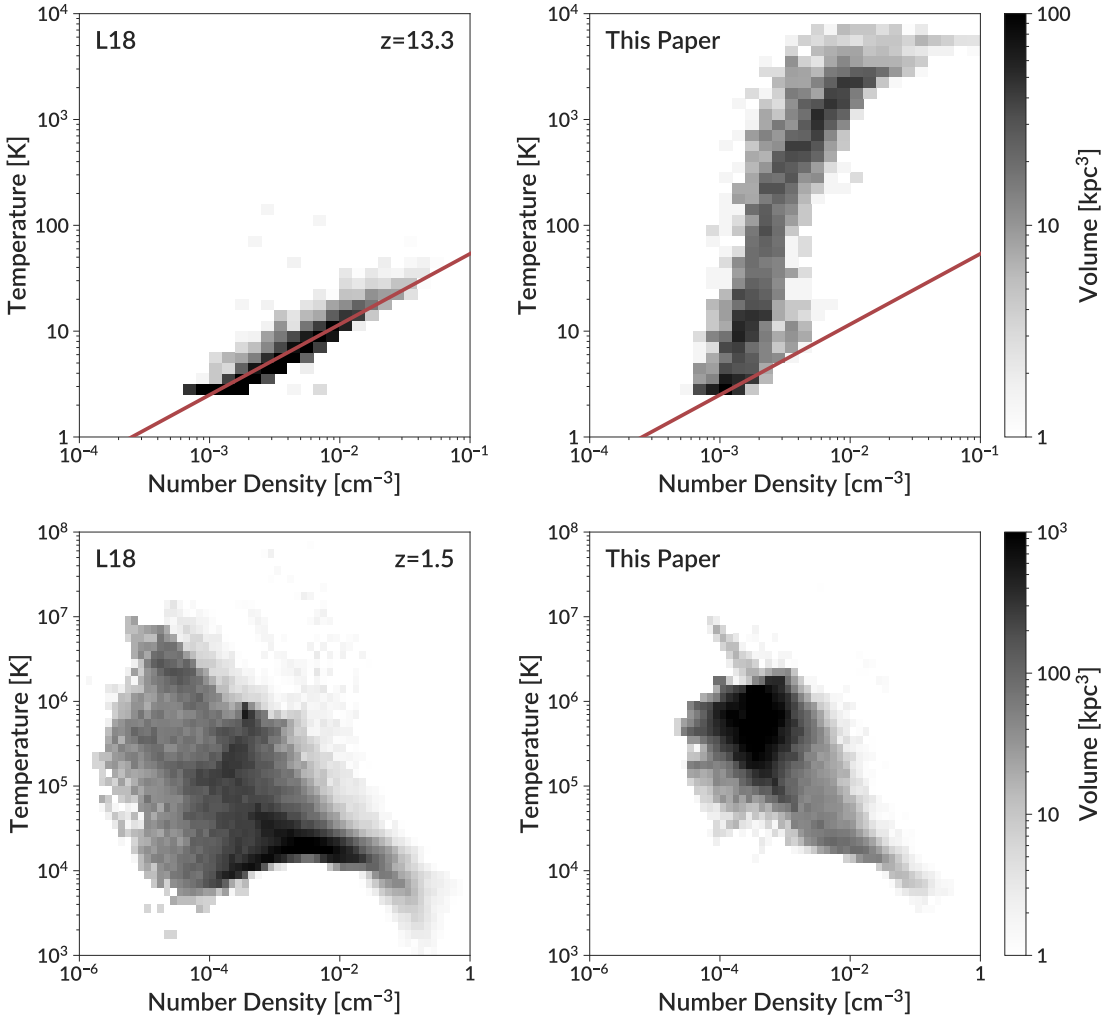

Figure B.1 Heatmap showing the temperature and density of gas within the virial radius of the largest halo. In each panel, the shading shows the volume of gas at the given temperature and density. The left column shows a run using the hydrodynamic scheme of L18, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 5$, and $f_{\rm HN,0} = 0$, while the right column shows the run with the energy-based hydrodynamics scheme that is the default in ART 2.0, $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 1$, and $f_{\rm HN,0} = 0$. The top row show these runs at $z = 13.3$ before any stars formed, while the bottom row shows the runs at $z = 1.5$. In the top panels, the red line shows the expected behavior for pure adiabatic compression. The code version of L18 exactly follows this line, while the updated version has extra heating from proper treatment of shocks.

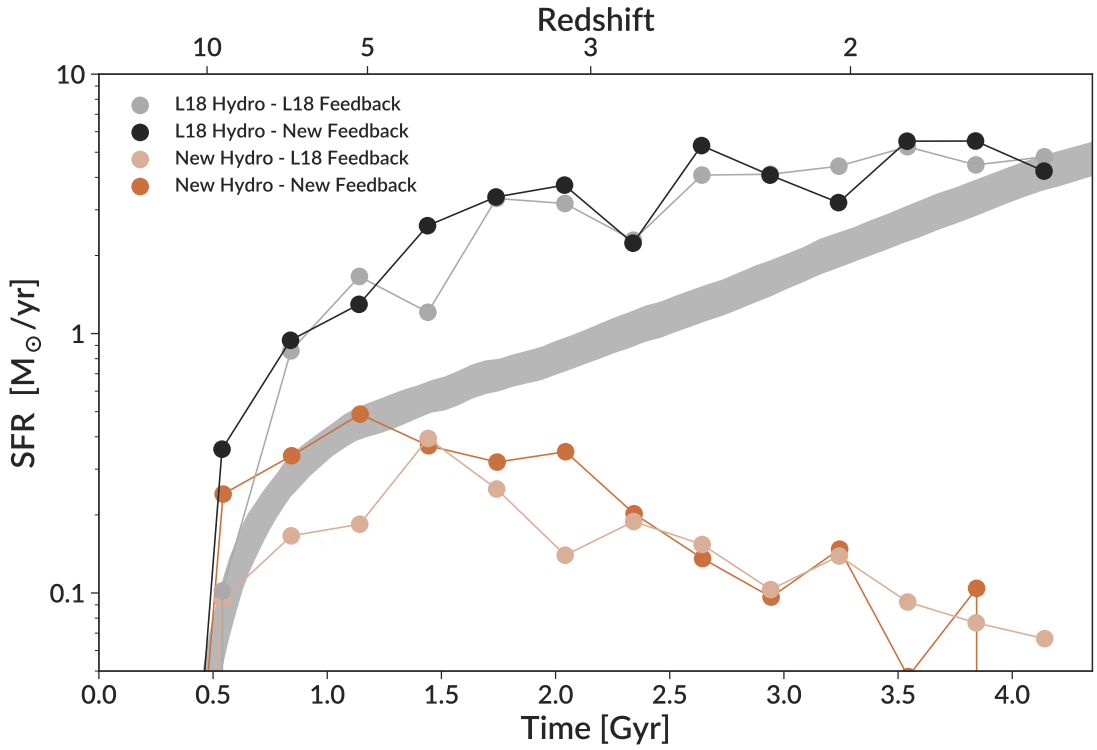

Figure B.2 The star formation history for galaxies in our test runs with varying hydrodynamics and feedback. We compare the feedback model of L18 to the feedback model presented in this paper. All runs use $\epsilon_{\rm ff} = 100\%$, $f_{\rm boost} = 5$, and $f_{\rm HN,0} = 0$. We compare to the UniverseMachine model (Behroozi et al., 2019). The change in hydrodynamics is solely responsible for the change in star formation rate, while our updates to feedback have little effect.

BIBLIOGRAPHY

BIBLIOGRAPHY

Adamo, A., Östlin, G., Zackrisson, E., et al. 2010, *MNRAS*, 407, 870

Adamo, A., Ryon, J. E., Messa, M., et al. 2017, *ApJ*, 841, 131

Adamo, A., Zeidler, P., Kruijssen, J. M. D., et al. 2020, *Space Sci. Rev.*, 216, 69

Agertz, O., Kravtsov, A. V., Leitner, S. N., & Gnedin, N. Y. 2013, *ApJ*, 770, 25

Agertz, O., Teyssier, R., & Moore, B. 2011, *MNRAS*, 410, 1391

Anderson, J., & King, I. R. 2000, *PASP*, 112, 1360

Andrews, B. H., Weinberg, D. H., Schönrich, R., & Johnson, J. A. 2017, *ApJ*, 835, 224

Antonini, F., Capuzzo-Dolcetta, R., Mastrobuono- Battisti, A., & Merritt, D. 2012, *ApJ*, 750, 111

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, 558, A33

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, *AJ*, 156, 123

Ballesteros-Paredes, J., André, P., Hennebelle, P., et al. 2020, *Space Sci. Rev.*, 216, 76

Bally, J. 2016, *ARA&A*, 54, 491

Bastian, N. 2008, *MNRAS*, 390, 759

Bastian, N., Cabrera-Ziri, I., Davies, B., & Larsen, S. S. 2013, *MNRAS*, 436, 2852

Bastian, N., Gieles, M., Lamers, H. J. G. L. M., Scheepmaker, R. A., & de Grijs, R. 2005, *A&A*, 431, 905

Bastian, N., & Lardo, C. 2018, *ARA&A*, 56, 83

Bastian, N., Adamo, A., Gieles, M., et al. 2012, *MNRAS*, 419, 2606

Baumgardt, H. 2001, *MNRAS*, 325, 1323

Beccari, G., Petr-Gotzens, M. G., Boffin, H. M. J., et al. 2017, *A&A*, 604, A22

Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, 488, 3143

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2013a, *ApJ*, 762, 109

Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., et al. 2013b, *ApJ*, 763, 18

Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212

Benincasa, S. M., Loebman, S. R., Wetzel, A., et al. 2020, *MNRAS*, 497, 3993

Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393

Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, 54, 529

Böker, T. 2008, *ApJ*, 672, L111

Bothwell, M. S., Kennicutt, R. C., & Lee, J. C. 2009, *MNRAS*, 400, 154

Bovy, J. 2017, *MNRAS*, 468, L63

Bradley, L., Sipőcz, B., Robitaille, T., et al. 2019, astropy/photutils: v0.7.2, doi:10.5281/zenodo.3568287

Brodie, J. P., & Strader, J. 2006, *ARA&A*, 44, 193

Brown, G., & Gnedin, O. Y. 2021a, *J. New Ast.*, 84, 101501

—. 2021b, *MNRAS*, 508, 5935

Brown, G., Gnedin, O. Y., & Li, H. 2018, *ApJ*, 864, 94

Cabrera-Ziri, I., Bastian, N., Longmore, S. N., et al. 2015, *MNRAS*, 448, 2224

Caldwell, S., & Chang, P. 2018, *MNRAS*, 474, 4818

Calzetti, D., Lee, J. C., Sabbi, E., et al. 2015, *AJ*, 149, 51

Cantat-Gaudin, T., Jordi, C., Vallenari, A., et al. 2018, *A&A*, 618, A93

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, *ApJ*, 345, 245

Carretta, E., Bragaglia, A., Gratton, R., & Lucatello, S. 2009a, *A&A*, 505, 139

Carretta, E., Bragaglia, A., Gratton, R. G., et al. 2009b, *A&A*, 505, 117

—. 2010a, *A&A*, 520, A95

—. 2010b, *ApJ*, 714, L7

Cen, R., & Ostriker, J. P. 1992, *ApJ*, 399, L113

Ceverino, D., Klypin, A., Klimek, E. S., et al. 2014, *MNRAS*, 442, 1545

Chandar, R., Whitmore, B. C., Dinino, D., et al. 2016, *ApJ*, 824, 71

Chandrasekhar, S. 1943, *ApJ*, 97, 255

Choksi, N., & Gnedin, O. Y. 2019a, *MNRAS*, 486, 331

—. 2019b, *MNRAS*, 488, 5409

Choksi, N., Gnedin, O. Y., & Li, H. 2018, *MNRAS*, 480, 2343

Choksi, N., & Kruijssen, J. M. D. 2021, *MNRAS*, 507, 5492

Claydon, I., Gieles, M., Varri, A. L., Heggie, D. C., & Zocchi, A. 2019, *MNRAS*, 487, 147

Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. 1994, *MNRAS*, 271, 781

Colombo, D., Rosolowsky, E., Ginsburg, A., Duarte-Cabral, A., & Hughes, A. 2015, *MNRAS*, 454, 2067

Conroy, C., Loeb, A., & Spergel, D. N. 2011, *ApJ*, 741, 72

Cook, D. O., Lee, J. C., Adamo, A., et al. 2019, *MNRAS*, 484, 4897

Corbelli, E. 2003, *MNRAS*, 342, 199

Côté, P., Piatek, S., Ferrarese, L., et al. 2006, *ApJS*, 165, 57

Cottrell, P. L., & Da Costa, G. S. 1981, *ApJ*, 245, L79

Cuevas-Otahola, B., Mayya, Y. D., Puerari, I., & Rosa-González, D. 2020, *MNRAS*, 492, 993

—. 2021, *MNRAS*, 500, 4422

Da Rio, N., Gouliermis, D. A., & Gennaro, M. 2010, *ApJ*, 723, 166

Da Rio, N., Tan, J. C., & Jaehnig, K. 2014, *ApJ*, 795, 55

de Mink, S. E., Pols, O. R., Langer, N., & Izzard, R. G. 2009, *A&A*, 507, L1

Decressin, T., Meynet, G., Charbonnel, C., Prantzos, N., & Ekström, S. 2007, *A&A*, 464, 1029

Dejonghe, H. 1987, *MNRAS*, 224, 13

Denissenkov, P. A., & Hartwick, F. D. A. 2014, *MNRAS*, 437, L21

Dessauges-Zavadsky, M., Richard, J., Combes, F., et al. 2019, Nature Astronomy, 3, 1115

Do, T., Kerzendorf, W., Winsor, N., et al. 2015, *ApJ*, 809, 143

D'Souza, R., & Bell, E. F. 2018, Nature Astronomy, 2, 737

Elmegreen, B. G. 2010, *ApJ*, 712, L184

Elson, R. A. W., Fall, S. M., & Freeman, K. C. 1987, *ApJ*, 323, 54

Emerick, A., Bryan, G. L., & Mac Low, M.-M. 2019, *MNRAS*, 482, 1304

Escala, A., & Larson, R. B. 2008, *ApJ*, 685, L31

Evans, Neal J., I., Heiderman, A., & Vutisalchavakul, N. 2014, *ApJ*, 782, 114

Evans, T. A., Fattahi, A., Deason, A. J., & Frenk, C. S. 2020, *MNRAS*, 497, 4311

Fall, S. M., & Efstathiou, G. 1980, *MNRAS*, 193, 189

Fall, S. M., & Zhang, Q. 2001, *ApJ*, 561, 751

Feldmeier-Krause, A., Kerzendorf, W., Neumayer, N., et al. 2017a, *MNRAS*, 464, 194

Feldmeier-Krause, A., Zhu, L., Neumayer, N., et al. 2017b, *MNRAS*, 466, 4040

Fernández-Trincado, J. G., Zamora, O., García-Hernández, D. A., et al. 2017, *ApJ*, 846, L2

Ferrarese, L., Côté, P., Dalla Bontà, E., et al. 2006, *ApJ*, 644, L21

Ferraro, F. R., Massari, D., Dalessandro, E., et al. 2016, *ApJ*, 828, 75

Forbes, J. C., Krumholz, M. R., Burkert, A., & Dekel, A. 2014, *MNRAS*, 443, 168

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306

Fouesneau, M., Johnson, L. C., Weisz, D. R., et al. 2014, *ApJ*, 786, 117

Freeman, K. C. 1993, in IAU Symp. 153: Galactic Bulges, 263

Fujii, M. S., Saitoh, T. R., Hirai, Y., & Wang, L. 2021, *PASJ*, 73, 1074

Fukushige, T., & Heggie, D. C. 2000, *MNRAS*, 318, 753

Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, 616, A1

Gallart, C., Monelli, M., Mayer, L., et al. 2015, *ApJ*, 811, L18

Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, 362, 41

Garrison-Kimmel, S., Boylan-Kolchin, M., Bullock, J. S., & Lee, K. 2014, *MNRAS*, 438, 2578

Gentry, E. S., Krumholz, M. R., Dekel, A., & Madau, P. 2017, *MNRAS*, 465, 2471

Georgiev, I. Y., Böker, T., Leigh, N., Lützgendorf, N., & Neumayer, N. 2016, *MNRAS*, 457, 2122

Gieles, M., Baumgardt, H., Heggie, D. C., & Lamers, H. J. G. L. M. 2010, *MNRAS*, 408, L16

Gieles, M., & Portegies Zwart, S. F. 2011, *MNRAS*, 410, L6

Gieles, M., Portegies Zwart, S. F., Baumgardt, H., et al. 2006, *MNRAS*, 371, 793

Gieles, M., & Renaud, F. 2016, *MNRAS*, 463, L103

Gingold, R. A., & Monaghan, J. J. 1977, *MNRAS*, 181, 375

Gnedin, N. Y. 2014, *ApJ*, 793, 29

Gnedin, N. Y., & Abel, T. 2001, *J. New Ast.*, 6, 437

Gnedin, N. Y., & Draine, B. T. 2014, *ApJ*, 795, 37

Gnedin, N. Y., & Kravtsov, A. V. 2011, *ApJ*, 728, 88

Gnedin, O. Y., & Ostriker, J. P. 1997, *ApJ*, 474, 223

Gnedin, O. Y., Ostriker, J. P., & Tremaine, S. 2014, *ApJ*, 785, 71

Goldman, S. R., van Loon, J. T., Zijlstra, A. A., et al. 2017, *MNRAS*, 465, 403

Goodwin, S. P., & Bastian, N. 2006, *MNRAS*, 373, 752

Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, *MNRAS*, 467, 179

Grasha, K. 2018, PhD thesis, University of Massachusetts

Grasha, K., Calzetti, D., Adamo, A., et al. 2019, *MNRAS*, 483, 4707

Gratton, R., Sneden, C., & Carretta, E. 2004, *ARA&A*, 42, 385

Gratton, R. G., Carretta, E., & Bragaglia, A. 2012, *A&A Rev.*, 20, 50

Grevesse, N., & Sauval, A. J. 1998, *Space Sci. Rev.*, 85, 161

Grimmett, J. J., Karakas, A. I., Heger, A., Müller, B., & West, C. 2020, *MNRAS*, 496, 4987

Grudić, M. Y., Guszejnov, D., Hopkins, P. F., Offner, S. S. R., & Faucher-Giguère, C.-A. 2021, *MNRAS*, 506, 2199

Grudić, M. Y., Guszejnov, D., Offner, S. S. R., et al. 2022, *MNRAS*, 512, 216

Guth, A. H. 1981, *Phys. Rev. D*, 23, 347

Haardt, F., & Madau, P. 2001, in Clusters of Galaxies and the High Redshift Universe Observed in X-rays, ed. D. M. Neumann & J. T. V. Tran, 64

Hahn, O., & Abel, T. 2011, *MNRAS*, 415, 2101

Hammer, F., Puech, M., Chemin, L., Flores, H., & Lehnert, M. D. 2007, *ApJ*, 662, 322

Harris, W. E. 1991, *ARA&A*, 29, 543

—. 1996, *AJ*, 112, 1487

Harris, W. E., Blakeslee, J. P., & Harris, G. L. H. 2017, *ApJ*, 836, 67

Hartmann, M., Debattista, V. P., Seth, A., Cappellari, M., & Quinn, T. R. 2011, *MNRAS*, 418, 2697

Hayden, M. R., Bovy, J., Holtzman, J. A., et al. 2015, *ApJ*, 808, 132

Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, *ApJ*, 591, 288

Helmi, A., Babusiaux, C., Koppelman, H. H., et al. 2018, *Nature*, 563, 85

Heyer, M., Gutermuth, R., Urquhart, J. S., et al. 2016, *A&A*, 588, A29

Hills, J. G. 1980, *ApJ*, 235, 986

Hislop, J. M., Naab, T., Steinwandel, U. P., et al. 2022, *MNRAS*, 509, 5938

Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv e-prints, arXiv:1008.4686

Holtzman, J. A., Faber, S. M., Shaya, E. J., et al. 1992, *AJ*, 103, 691

Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, *MNRAS*, 445, 581

Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, *MNRAS*, 480, 800

Horiuchi, S., Nakamura, K., Takiwaki, T., Kotake, K., & Tanaka, M. 2014, *MNRAS*, 445, L99

Ibata, R., Nipoti, C., Sollima, A., et al. 2013, *MNRAS*, 428, 3648

Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, *ApJS*, 125, 439

Iwamoto, K., Mazzali, P. A., Nomoto, K., et al. 1998, *Nature*, 395, 672

Jacobs, B. A., Rizzi, L., Tully, R. B., et al. 2009, *AJ*, 138, 332

Jeans, J. H. 1902, Philosophical Transactions of the Royal Society of London Series A, 199, 1

Johnson, C. I., Pilachowski, C. A., Michael Rich, R., & Fulbright, J. P. 2009, *ApJ*, 698, 2048

Johnson, L. C., Seth, A. C., Dalcanton, J. J., et al. 2012, *ApJ*, 752, 95

Joo, S.-J., & Lee, Y.-W. 2013, *ApJ*, 762, 36

Jordán, A., Côté, P., Blakeslee, J. P., et al. 2005, *ApJ*, 634, 1002

Jordán, A., McLaughlin, D. E., Côté, P., et al. 2007, *ApJS*, 171, 101

Jordán, A., Peng, E. W., Blakeslee, J. P., et al. 2009, *ApJS*, 180, 54

Katz, N. 1992, *ApJ*, 391, 502

Katz, N., Quinn, T., Bertschinger, E., & Gelb, J. M. 1994, *MNRAS*, 270, L71

Keller, B. W., & Kruijssen, J. M. D. 2022, *MNRAS*, 512, 199

Keller, B. W., Kruijssen, J. M. D., Pfeffer, J., et al. 2020, *MNRAS*, 495, 4248

Keller, B. W., Wadsley, J., Benincasa, S. M., & Couchman, H. M. P. 2014, *MNRAS*, 442, 3013

Kelly, B. C. 2007, *ApJ*, 665, 1489

Kennicutt, Robert C., J. 1998, *ARA&A*, 36, 189

Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, 142, 35

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R. D. 2013, *A&A*, 558, A53

Kim, J.-h., Agertz, O., Teyssier, R., et al. 2016, *ApJ*, 833, 202

Kim, J.-h., Ma, X., Grudić, M. Y., et al. 2018, *MNRAS*, 474, 4232

Kirby, E. N., Cohen, J. G., Guhathakurta, P., et al. 2013, *ApJ*, 779, 102

Kissler-Patig, M., Jordán, A., & Bastian, N. 2006, *A&A*, 448, 1031

Klessen, R. S., & Burkert, A. 2000, *ApJS*, 128, 287

Kobayashi, C., Umeda, H., Nomoto, K., Tominaga, N., & Ohkubo, T. 2006, *ApJ*, 653, 1145

Koppelman, H. H., Helmi, A., Massari, D., Price-Whelan, A. M., & Starkenburg, T. K. 2019, *A&A*, 631, L9

Kounkel, M., Covey, K., Suárez, G., et al. 2018, *AJ*, 156, 84

Kourkchi, E., & Tully, R. B. 2017, *ApJ*, 843, 16

Krabbe, A., Genzel, R., Drapatz, S., & Rotaciuc, V. 1991, *ApJ*, 382, L19

Krause, M., Charbonnel, C., Decressin, T., Meynet, G., & Prantzos, N. 2013, *A&A*, 552, A121

Kravtsov, A. V. 1999, PhD thesis, New Mexico State University

—. 2003, *ApJ*, 590, L1

Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, *ApJS*, 111, 73

Kroupa, P. 2001, *MNRAS*, 322, 231

Kruijssen, J. M. D. 2009, *A&A*, 507, 1409

—. 2014, Classical and Quantum Gravity, 31, 244006

—. 2015, *MNRAS*, 454, 1658

Kruijssen, J. M. D., & Lamers, H. J. G. L. M. 2008, *A&A*, 490, 151

Kruijssen, J. M. D., Pelupessy, F. I., Lamers, H. J. G. L. M., et al. 2012, *MNRAS*, 421, 1927

Kruijssen, J. M. D., Pelupessy, F. I., Lamers, H. J. G. L. M., Portegies Zwart, S. F., & Icke, V. 2011, *MNRAS*, 414, 1339

Kruijssen, J. M. D., Pfeffer, J. L., Reina-Campos, M., Crain, R. A., & Bastian, N. 2019a, *MNRAS*, 486, 3180

Kruijssen, J. M. D., Schruba, A., Chevance, M., et al. 2019b, *Nature*, 569, 519

Krumholz, M. R. 2014, *Phys. Rep.*, 539, 49

Krumholz, M. R., McKee, C. F., & Bland -Hawthorn, J. 2019, *ARA&A*, 57, 227

Krumholz, M. R., Adamo, A., Fumagalli, M., et al. 2015, *ApJ*, 812, 147

Küpper, A. H. W., Lane, R. R., & Heggie, D. C. 2012, *MNRAS*, 420, 2700

Küpper, A. H. W., MacLeod, A., & Heggie, D. C. 2008, *MNRAS*, 387, 1248

Lada, C. J., & Lada, E. A. 2003, *ARA&A*, 41, 57

Lahén, N., Naab, T., Johansson, P. H., et al. 2019, *ApJ*, 879, L18

Lane, H. B., Grudić, M. Y., Guszejnov, D., et al. 2022, *MNRAS*, 510, 4767

Larsen, S. S. 1999, *A&AS*, 139, 393

—. 2002, *AJ*, 124, 1393

Larsen, S. S., Strader, J., & Brodie, J. P. 2012, *A&A*, 544, L14

Larson, R. B. 1974, *MNRAS*, 169, 229

—. 1981, *MNRAS*, 194, 809

Leaman, R., VandenBerg, D. A., & Mendel, J. T. 2013, *MNRAS*, 436, 122

Lee, E. J., Miville-Deschênes, M.-A., & Murray, N. W. 2016, *ApJ*, 833, 229

Lee, J. C., Gil de Paz, A., Tremonti, C., et al. 2009, *ApJ*, 706, 599

Leigh, N., Böker, T., & Knigge, C. 2012, *MNRAS*, 424, 2130

Leitherer, C., Robert, C., & Drissen, L. 1992, *ApJ*, 401, 596

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, *ApJS*, 123, 3

Li, H., & Gnedin, O. Y. 2014, *ApJ*, 796, 10

—. 2019, *MNRAS*, 486, 4030

Li, H., Gnedin, O. Y., & Gnedin, N. Y. 2018, *ApJ*, 861, 107

Li, H., Gnedin, O. Y., Gnedin, N. Y., et al. 2017, *ApJ*, 834, 69

Li, H., Vogelsberger, M., Marinacci, F., & Gnedin, O. Y. 2019, *MNRAS*, 487, 364

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, *ApJ*, 772, 119

Lotz, J. M., Telford, R., Ferguson, H. C., et al. 2001, *ApJ*, 552, 572

Lu, J. R., Do, T., Ghez, A. M., et al. 2013, *ApJ*, 764, 155

Ma, X., Quataert, E., Wetzel, A., et al. 2020, *MNRAS*, 498, 2001

Mackey, A. D., & Gilmore, G. F. 2003a, *MNRAS*, 338, 120

—. 2003b, *MNRAS*, 338, 85

Maiolino, R., & Mannucci, F. 2019, *A&A Rev.*, 27, 3

Maíz Apellániz, J. 2009, *ApJ*, 699, 1938

Majewski, S. R., Schiavon, R. P., Frinchaboy, P. M., et al. 2017, *AJ*, 154, 94

Mannucci, F., Cresci, G., Maiolino, R., et al. 2009, *MNRAS*, 398, 1915

Maoz, D., & Graur, O. 2017, *ApJ*, 848, 25

Marino, A. F., Milone, A. P., Piotto, G., et al. 2011, *ApJ*, 731, 64

Marino, A. F., Milone, A. P., Karakas, A. I., et al. 2015, *MNRAS*, 450, 815

Martizzi, D., Faucher-Giguère, C.-A., & Quataert, E. 2015, *MNRAS*, 450, 504

Massari, D., Koppelman, H. H., & Helmi, A. 2019, *A&A*, 630, L4

Massari, D., Posti, L., Helmi, A., Fiorentino, G., & Tolstoy, E. 2017, *A&A*, 598, L9

Massari, D., Mucciarelli, A., Ferraro, F. R., et al. 2014, *ApJ*, 795, 22

Mathew, S. S., & Federrath, C. 2021, *MNRAS*, 507, 2448

McKee, C. F., & Ostriker, E. C. 2007, *ARA&A*, 45, 565

McLaughlin, D. E., & van der Marel, R. P. 2005, *ApJS*, 161, 304

Meng, X., & Gnedin, O. Y. 2020, *MNRAS*, 494, 1263

—. 2021, *MNRAS*, 502, 1433

—. 2022, arXiv e-prints, arXiv:2201.09826

Meng, X., Gnedin, O. Y., & Li, H. 2019, *MNRAS*, 486, 1574

Meurer, G. R., Heckman, T. M., Leitherer, C., et al. 1995, *AJ*, 110, 2665

Milone, A. P., Piotto, G., Renzini, A., et al. 2017, *MNRAS*, 464, 3636

Miville-Deschênes, M.-A., Murray, N., & Lee, E. J. 2017, *ApJ*, 834, 57

Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution

Mo, H. J., Mao, S., & White, S. D. M. 1998, *MNRAS*, 295, 319

Mok, A., Chandar, R., & Fall, S. M. 2021, *ApJ*, 911, 8

Moustakas, J., Kennicutt, Jr., R. C., Tremonti, C. A., et al. 2010, *ApJS*, 190, 233

Muratov, A. L., & Gnedin, O. Y. 2010, *ApJ*, 718, 1266

Murray, N. 2009, *ApJ*, 691, 946

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493

Navarro, J. F., & White, S. D. M. 1994, *MNRAS*, 267, 401

Neumayer, N., Seth, A., & Böker, T. 2020, *A&A Rev.*, 28, 4

Neyrinck, M. C., Hamilton, A. J. S., & Gnedin, N. Y. 2004, *MNRAS*, 348, 1

Nomoto, K., & Leung, S.-C. 2018, *Space Sci. Rev.*, 214, 67

Nomoto, K., Tominaga, N., Umeda, H., Kobayashi, C., & Maeda, K. 2006, Nuclear Physics A, 777, 424

Ochsendorf, B. B., Meixner, M., Roman-Duval, J., Rahman, M., & Evans, Neal J., I. 2017, *ApJ*, 841, 109

Offner, S. S. R., Klein, R. I., McKee, C. F., & Krumholz, M. R. 2009, *ApJ*, 703, 131

Olivares E., F., Hamuy, M., Pignata, G., et al. 2010, *ApJ*, 715, 833

Ostriker, J. P., Spitzer, L. J., & Chevalier, R. A. 1972, *ApJ*, 176, L51

Padoan, P., Haugbølle, T., & Nordlund, Å. 2012, *ApJ*, 759, L27

Padoan, P., Pan, L., Juvela, M., Haugbølle, T., & Nordlund, Å. 2020, *ApJ*, 900, 82

Pejcha, O., & Thompson, T. A. 2015, *ApJ*, 801, 90

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, 124, 266

—. 2010, *AJ*, 139, 2097

Pfeffer, J., Kruijssen, J. M. D., Crain, R. A., & Bastian, N. 2018, *MNRAS*, 475, 4309

Pignatari, M., Herwig, F., Hirschi, R., et al. 2016, *ApJS*, 225, 24

Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, 473, 4077

Piskunov, A. E., Schilbach, E., Kharchenko, N. V., Röser, S., & Scholz, R. D. 2007, *A&A*, 468, 151

Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, *ARA&A*, 48, 431

Press, W. H., & Schechter, P. 1974, *ApJ*, 187, 425

Prieto, J. L., & Gnedin, O. Y. 2008, *ApJ*, 689, 919

Puzia, T. H., Kissler-Patig, M., Thomas, D., et al. 2005, *A&A*, 439, 997

Raiteri, C. M., Villata, M., & Navarro, J. F. 1996, *A&A*, 315, 105

Renaud, F., Agertz, O., & Gieles, M. 2017, *MNRAS*, 465, 3622

Renzini, A., D'Antona, F., Cassisi, S., et al. 2015, *MNRAS*, 454, 4197

Ritter, C., Herwig, F., Jones, S., et al. 2018, *MNRAS*, 480, 538

Roman-Duval, J., Jackson, J. M., Heyer, M., Rathborne, J., & Simon, R. 2010, *ApJ*, 723, 492

Romano, D., Karakas, A. I., Tosi, M., & Matteucci, F. 2010, *A&A*, 522, A32

Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, *ApJ*, 672, 19

Ryon, J. E., Bastian, N., Adamo, A., et al. 2015, *MNRAS*, 452, 525

Ryon, J. E., Gallagher, J. S., Smith, L. J., et al. 2017, *ApJ*, 841, 92

Sabbi, E., Calzetti, D., Ubeda, L., et al. 2018, *ApJS*, 235, 23

Santistevan, I. B., Wetzel, A., El-Badry, K., et al. 2020, *MNRAS*, 497, 747

Sawala, T., Frenk, C. S., Fattahi, A., et al. 2016, *MNRAS*, 457, 1931

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, 446, 521

Schechter, P. 1976, *ApJ*, 203, 297

Scheepmaker, R. A., Haas, M. R., Gieles, M., et al. 2007, *A&A*, 469, 925

Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103

Schmidt, M. 1959, *ApJ*, 129, 243

Sellwood, J. A., & Binney, J. J. 2002, *MNRAS*, 336, 785

Semenov, V. A., Kravtsov, A. V., & Diemer, B. 2021, arXiv e-prints, arXiv:2107.14240

Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2016, *ApJ*, 826, 200

—. 2017, *ApJ*, 845, 133

Seth, A. C., Cappellari, M., Neumayer, N., et al. 2010, *ApJ*, 714, 713

Siegel, M. H., Dotter, A., Majewski, S. R., et al. 2007, *ApJ*, 667, L57

Silva-Villa, E., Adamo, A., Bastian, N., Fouesneau, M., & Zackrisson, E. 2014, *MNRAS*, 440, L116

Silva-Villa, E., & Larsen, S. S. 2011, *A&A*, 529, A25

Spitzer, Lyman, J. 1958, *ApJ*, 127, 17

Spitzer, L. 1987, Dynamical Evolution of Globular Clusters (Princeton: Princeton University Press)

Springel, V., Frenk, C. S., & White, S. D. M. 2006, *Nature*, 440, 1137

Springel, V., & Hernquist, L. 2003, *MNRAS*, 339, 289

Springel, V., Wang, J., Vogelsberger, M., et al. 2008, *MNRAS*, 391, 1685

Stinson, G., Seth, A., Katz, N., et al. 2006, *MNRAS*, 373, 1074

Strader, J., Brodie, J. P., Cenarro, A. J., Beasley, M. A., & Forbes, D. A. 2005, *AJ*, 130, 1315

Sun, J., Leroy, A. K., Schruba, A., et al. 2018, *ApJ*, 860, 172

Sun, W., de Grijs, R., Fan, Z., & Cameron, E. 2016, *ApJ*, 816, 9

Tailo, M., Di Criscienzo, M., D'Antona, F., Caloi, V., & Ventura, P. 2016, *MNRAS*, 457, 4525

Timmes, F. X., Woosley, S. E., & Weaver, T. A. 1995, *ApJS*, 98, 617

Toomre, A. 1964, *ApJ*, 139, 1217

Torrey, P., Vogelsberger, M., Marinacci, F., et al. 2019, *MNRAS*, 484, 5587

Tremaine, S. D., Ostriker, J. P., & Spitzer, L., J. 1975, *ApJ*, 196, 407

Trujillo-Gomez, S., Kruijssen, J. M. D., Reina-Campos, M., et al. 2021, *MNRAS*, 503, 31

Tsatsi, A., Mastrobuono-Battisti, A., van de Ven, G., et al. 2017, *MNRAS*, 464, 3720

Tully, R. B., Courtois, H. M., & Sorce, J. G. 2016, *AJ*, 152, 50

Urquhart, J. S., König, C., Giannetti, A., et al. 2018, *MNRAS*, 473, 1059

Usero, A., Leroy, A. K., Walter, F., et al. 2015, *AJ*, 150, 115

Utomo, D., Sun, J., Leroy, A. K., et al. 2018, *ApJ*, 861, L18

van der Marel, R. P., & Kallivayalil, N. 2014, *ApJ*, 781, 121

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28

VandenBerg, D. A., Brogaard, K., Leaman, R., & Casagrande, L. 2013, *ApJ*, 775, 134

Vázquez, G. A., & Leitherer, C. 2005, *ApJ*, 621, 695

Vesperini, E., Zepf, S. E., Kundu, A., & Ashman, K. M. 2003, *ApJ*, 593, 760

Villanova, S., Geisler, D., Gratton, R. G., & Cassisi, S. 2014, *ApJ*, 791, 107

Vincenzo, F., & Kobayashi, C. 2018, *MNRAS*, 478, 155

Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, Nature Reviews Physics, 2, 42

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, 444, 1518

Wagner-Kaiser, R., Sarajedini, A., Dalcanton, J. J., Williams, B. F., & Dolphin, A. 2015, *MNRAS*, 451, 724

Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, *MNRAS*, 454, 83

Webb, J. J., Harris, W. E., Sills, A., & Hurley, J. R. 2013, *ApJ*, 764, 124

Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, *ApJ*, 827, L23

Whitmore, B. C., Chandar, R., Schweizer, F., et al. 2010, *AJ*, 140, 75

Wiersma, R. P. C., Schaye, J., Theuns, T., Dalla Vecchia, C., & Tornatore, L. 2009, *MNRAS*, 399, 574

Williams, J. P., & McKee, C. F. 1997, *ApJ*, 476, 166

Willman, B., & Strader, J. 2012, *AJ*, 144, 76

Woosley, S. E., & Weaver, T. A. 1995, *ApJS*, 101, 181

Zackrisson, E., Rydberg, C.-E., Schaerer, D., Östlin, G., & Tuli, M. 2011, *ApJ*, 740, 13

Zeldovich, Y. B. 1970, *A&A*, 500, 13

Zemp, M., Gnedin, O. Y., Gnedin, N. Y., & Kravtsov, A. V. 2011, *ApJS*, 197, 30

Zepf, S. E., Ashman, K. M., English, J., Freeman, K. C., & Sharples, R. M. 1999, *AJ*, 118, 752