Cosmological Small-scale Structure: The Formation of The First Stars, Galaxies, and Globular Clusters

by

Alexander L. Muratov

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Astronomy and Astrophysics)
in The University of Michigan
2013

Doctoral Committee:

Associate Professor Oleg Gnedin, Chair
Professor August Evrard
Professor Mario Mateo
Associate Professor Eric Bell
Associate Professor Mateusz Ruszkowski

To all of my family, friends, and teachers. Let the good times roll.

ACKNOWLEDGMENTS

First and foremost, I must thank my advisor Oleg Gnedin for his six years of guidance that made all of this work possible. With his scientific acumen and resourcefulness, he helped shaped every idea expressed in this thesis. With his attention and time, he ensured that I was mostly productive, and that my analysis met a high standard of quality previously unknown to me. With his keen eye for visual design and knack for efficient communication, he helped me find the most effective ways to present my research, and also did much in the way of promoting my work directly. He also left me plenty of room to grow independently as a scientist and as a person. I can only hope the lessons I've learned from him stay with me for the rest of my life.

I thank my collaborator Marcel Zemp for his patience and attention while I was learning the ropes of analyzing cosmological simulations. Despite all of my training in programming and cosmology, I was completely clueless the first time I had to look at simulation input and output files. Thanks to Marcel, I quickly became sufficiently fluent in an abstract, jargon-laden language. He has also generously provided me with his analysis tools, which I made extensive use of for the production of my results.

I am deeply grateful to Nick Gnedin, who basically acted as a second advisor for the development and analysis of my simulations. I frequently relied on his expertise in computational astrophysics and theoretical cosmology to make sense of my work. He went above and beyond the call of duty of a collaborator by always responding to emails in a timely and thorough manner, even when they were the most tedious and mundane requests. I should also mention that he generated all of the initial conditions for my simulations, provided the original configuration files, and generously hosted me when I visited Fermilab, where we developed my branch of the ART code. I'll add here that I am also thankful to Andrey Kravtsov and Doug Rudd for providing support with the code.

The Astronomy Department at The University of Michigan has provided a friendly and intellectually stimulating atmosphere for my research. I thank the members of my dissertation committee: Eric Bell, Gus Evrard, Mario Mateo, and Mateusz Ruszkowski for productive suggestions in both the content of this thesis, and the general advice for being a good scientist. I thank Mary Putman for allowing me to work on a small project which resulted in my first co-authorship on a refereed paper. I also particularly thank Nuria Calvet, who has been a personal guardian for me during my entire time here, and who brought me here in the first place.

I want to mention that without the support of the IT staff of the University of Michigan, my projects could have been dead in the water several times. Roy Bonser and Jeremy Hallum both worked very hard to fix all of my computing issues though they were responsible for so many other things, and were always underappreciated. I have also found all of the support staff for the Flux cluster to be incredibly competent, and their intervention led to several performance improvements of my code. I will mention Brock Palen, Marlin Whitaker, Mark Montague, and Tony Markel, by name. I apologize for using up so much disk space on the cluster.

I have made many great friends over the years here, and they will hopefully remain my friends for the rest of my life. My girlfriend, Sonu Mishra, has provided constant companionship, encouragement, and warmth during the writing of this thesis. My office mates, Shannon Schmoll, Nathan Crockett, Joel Lamb, and Ashley King have had to tolerate my shenanigans on an every-day basis, and somehow they still talk to me. Joel Lamb and Tina Hsu were particularly supportive in that they literally helped to feed and shelter me. Mike Anderson and Nathan Crockett have become my brothers-in-arms, and without them I would have probably given up many times over. Mike Anderson and Laura Ingleby were instrumental in helping me find a job, and I would probably be going hungry next year if not for them. I will also name-check some other friends in Ann Arbor who are associated with the university: Colin Slater, Kiwi Davis, Ilse Cleeves, Sarah Mange, Mark Haynes, Robin Stottlemeyer, Tom Eyster, and Lou Chang. I'm sorry if I left anyone else out, but if I did, I will gladly buy you a beer as acknowledgment. My friends from college also made a big difference in my life during this thesis work, and though none of them ended up in

Ann Arbor, I have received on the order of 3,000 emails from them since we graduated from college. Many have come to visit me and have kept me sane over long nights of gmail chat. Sorry guys, it'd take too long to write all your names and I'd forget someone, but I'll particularly mention Taejin Min for a decade's worth of commiserating with me over academia.

My family has been nothing but loving and supportive over the years. Without my parents' immigration to the United States twenty-one years ago, I would most likely not have had the opportunity to pursue something as interesting as theoretical astrophysics. My mom, Victoria Muratov, still finds time and energy to listen to all of my problems, and carry them as her own. My dad, Leonid Muratov, has mentored me in all matters of life, and has taught me the foundation of mathematics that allowed me to pursue this work. My grandmother, Vanda Muratov, has generously shared the wisdom of her long and fruitful life with me. My grandfather, Semyon Vekstein, fought for our freedom in World War 2. My sister, Katie Muratov, has had to grow up without me, but still appears to look up to me for some reason.

I thank David Bowie, Nick Cave, Björk, Beth Gibbons, Bert Jansch, Leonard Cohen, Vladimir Vysotsky, Viktor Tsoi, Elliott Smith, Radiohead, and Dinosaur Jr. for providing the soundtrack for much of this work.

Chapter 2 of this thesis is a reproduction of a paper titled "Revisiting The First Galaxies: The effects of Population III stars on their host galaxies". It has been accepted for publication in the Astrophysical Journal as of May 2013 as Muratov et al. (2013a). Similarly, Chapter 3 is taken from a paper titled "Revisiting The First Galaxies: The epoch of Population III stars", which has also been accepted to the Astrophysical Journal as of June 2013 as Muratov et al. (2013b). Both of these works were supported in part by NSF grant AST-0708087, NSF grant AST-0708154, NASA grant NNX12AG44G, the DOE at Fermilab, and a 985 grant from Peking University. Chapter 4 is a reproduction of "Modeling the Metallicity Distribution of Globular Clusters", published in the Astrophysical Journal as Muratov & Gnedin (2010). I would like to thank the Rackham Graduate School for supporting me in the final year of my thesis via the Rackham Pre-doctoral Fellowship, and for providing me with many travel grants which allowed me to attend national and international

conferences at which this work was presented.

CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Though the majority of stars now live in large, massive galaxies, understanding the origins of all galaxies ab initio requires fully comprehensive modeling of cosmological small-scale structure. In this thesis, I present a theoretical study of galaxy formation that focuses on low-mass halos. These halos are the sites for the formation of the first stars and galaxies at high redshift, and they also they play a role in forming massive globular clusters in the outskirts of the Milky Way.

I develop a physical model for Population III star formation and feedback, and implemented it into the Eulerian hydrodynamic Adaptive Refinement Tree (ART) code. With this code, I designed, performed, and analyzed a suite of cosmological simulations that resolve the formation of the first stars and galaxies. I quantify the extent of the dynamical signatures Population III stars can impart on their host galaxies, and derive a characteristic mass threshold $3 \times 10^6 M_\odot$, above which Population III stellar feedback is no longer dynamically significant over significant cosmic timescales.

I measure the duration of time for which Population III stars are the dominant drivers of feedback in the universe. Due to the inhomogeneous and patchy enrichment of the intergalactic medium, I find Population III stars can continue forming in some environments well after the end of the cosmic dark ages. However, in individual galaxies that are sufficiently massive, Population II star formation takes over soon after the efficient enrichment by a single pair-instability supernova. Globally, Population II is dominant at cosmic epochs later than redshift ($z \approx 15$).

Finally, I construct a semi-analytical model for globular cluster formation in hierarchical cosmology, and use it to demonstrate a plausible scenario for the formation of the Milky Way's globular cluster system. My model is successful in matching both the metallicity and mass distributions of galactic globular clusters. In particular, the bimodal nature

of the metallicity distribution is for the first time explained by the single mechanism of the merging of protogalaxies.

CHAPTER 1

Introduction

1.1 Introduction

Though firmly rooted in the physical sciences, one might describe cosmology as a branch of history that deals with the universe as a whole. We are currently living in what some refer to as the golden age of cosmology. For the first time since the dawn of mankind, astronomical observations and theoretical physics have come together to form a coherent understanding of what things were like in the past, even prior to the dawn of planets, stars, and galaxies. The key to this understanding lies in the theoretical prediction (Gamow, 1948), and subsequent discovery of the cosmic microwave background radiation (Penzias & Wilson, 1965) as an afterglow of the "Big Bang". When this observation was combined with Edwin Hubble's study (Hubble, 1926) on the universal recession of galaxies, and expressed in the framework of Einstein's General Relativity (Einstein, 1916), functional equations that connected matter, space, and time were revealed (Friedmann 1922, reprinted as Friedmann 1999). The extragalactic Astronomers of the past few decades have focused on filling in the missing pieces of the puzzle of cosmic evolution between the dawn of light and the cosmos of the present day. Although the conclusions that have been drawn are not necessarily definitive, and are by no means fully comprehensive, we can now certainly present a convincing, scientifically-motivated narrative for the beginning moments of time and space and what happened thereafter, leading up to the present day.

We begin our story with the first stage of cosmic history, with the process known as "inflation", which is indeed the bang in the Big Bang. During this time, an initial tiny volume was rapidly stretched into an immense horizon, while the quantum perturbations of the initial density field were blown up to larger scales (Guth & Pi, 1982). Immediately

following inflation, the universe is well described as a nearly homogeneous soup of matter and radiation, with only slight ripples in the density distribution of matter corresponding to those initial perturbations. These small ripples, however, are enough for gravity to begin molding structure into the universe by preferentially pulling in matter where the ripples are already the most pronounced.

The properties of these overdense ripples and the environment they inhabit informs the initial conditions for the modern theory of the genesis of the first stars and galaxies. The initial cosmic fluctuations continue to grow, eventually reaching a critical threshold where they undergo rapid collapse, as gravity's pull triumphs locally over universal expansion. The density in the central region of this local spherical "halo" becomes so great that gas is eventually compressed sufficiently to form stars (Press & Schechter, 1974; White & Rees, 1978). Halos can continue to accrete matter from the ambient field, and merge with one another, building up into larger and larger structures.

How and when this process took place depends on the nature of unseen "dark matter", which is estimated to make up ~80 % of all mass in the universe. The first hints that the universe had such a component came from the presence of an inexplicable apparent excess of gravitatioanl force observed in galaxy clusters (Zwicky, 1933) and later, the rotation of spiral galaxies (Rubin et al., 1980). Specifically, Rubin's work and all subsequent analysis of spiral galaxies revealed that the rotational velocities of stars in the outer parts of galaxies were much faster than the expected gravitational force from all galactic visible matter would seem to suggest. Analysis of smaller galaxies showed a similar trend (Mateo, 1998). Dark matter is now understood to be non-baryonic - in other words, it is not composed of protons, electrons, and neutrons (Bond & Efstathiou, 1984). As the dark matter particle has yet to be discovered at the time of writing, we may only attempt to derive its properties *a posteriori*, that is by inferring the kinds of particles necessary to mold cosmic structure into its present form. Currently, evidence points to a "cold" dark matter particle which travels at sub-relativistic velocities, and can clump up into very small halos at early times, thus instigating "bottom-up" structure formation. This scenario is confirmed by the high abundance of galaxies in the Hubble deep field (Madau et al., 1996) at early cosmic times, the high density of absorption systems in the intergalactic medium, and the fact that larger structures

like galaxy clusters formed only recently. "Hot" dark matter, such as neutrinos, would conversely suggest a "top-down" history (Eggen et al., 1962), where enormous galaxy clusters would form first, with smaller galaxies collapsing at later times. The top-down scenario is not consistent with observations (Blumenthal et al., 1984). It has been shown that all dynamically stable cold dark matter halos are structurally similar, and the density profile is well fit by a single equation with two parameters (Navarro et al., 1997).

Let us continue to recount the history of the universe, assuming that dark matter is indeed cold, which is known as the so-called Cold Dark Matter (CDM) paradigm. Dark matter halos that are immune to the effects of cosmic streaming begin runaway collapse only 20 Million years (Myr) after the big bang (or, at cosmic redshift z~100). When the gravitational potential becomes sufficient, gas can begin to flow into these halos. This begins to occur about 100 Myr into cosmic history (z~40-20) when the halos reach a mass of $10^4 - 10^5\,M_\odot$. Once they reach the total mass of $10^6\,M_\odot$ gas is finally able to cool and condense into the core of the galaxies and reach sufficient density and temperature to initiate star formation (Tegmark et al., 1997; Haiman et al., 1996), thus bringing about the end of the cosmic "dark ages", so termed because there is nothing at all in the universe that our eyes could detect prior to these stars (for a review, see Miralda-Escudé 2003). Finally, once halos grow to approximately $10^8 M_\odot$, which happens 400 Myr after the big bang (z~12) hydrogen line cooling is able to rapidly transport gas into the inner parts of the halo and efficient, sustained star formation can begin via disk instability, allowing the halos to for the first time resemble the galaxies of modernity.

Though the dark ages formally end with the creation of the first stars, the universe does not immediately become accessible to our telescopes. Hydrogen in the then-neutral intergalactic medium absorbs much of the light emitted by these stars and galaxies, hiding them from our view (Gunn & Peterson, 1965). The situation is made worse by the implications of bottom-up structure formation, as it dictates that these galaxies are far too small and faint to be detectable at their current distance. Things get somewhat more observable when the next historical milestone is reached at a cosmic time of about one billion years (z~6): the epoch of reionization. During this phase, the output of ionizing radiation from the young galaxies, and from the first massive black holes becomes sufficient to heat and ionize the

majority of the intergalactic medium. Though black holes were initially suspected to be the primary agents of reionization, it is now believed that young stars in normal galaxies did most of the work (Barkana & Loeb, 2001; Loeb & Barkana, 2001). Supernovae from these stars are also the sources of the first elements heavier than helium in the universe and begin the process of polluting the intergalactic medium, thereby influencing the enrichment history of all future galaxies (Carr et al., 1984).

This broad narrative is now almost unanimously accepted, but many aspects of the end of the dark ages remain poorly understood. Indeed, there are very limited empirical constraints on the details of the processes mentioned above. Ultimately, we still have no idea what most galaxies looked like prior to the epoch of reionization. We do not know what kind of stars they contained, or how those stars were able to reionize and enrich the universe. At the time of writing, this field is extremely data-starved, and observational facilities are simply unable to probe deep enough into the epoch to resolve what is actually happening, forcing modelers to fend for themselves.

After the epoch of reionization, the history of the universe is traditionally told from the perspective of "large-scale structure". Star formation is increasingly dominated by large, massive galaxies, which are built from the merging of many smaller progenitors and continued accretion from filaments. Subsequently, galaxies agglomerate into clusters and walls, connected by filaments, eventually tracing out an enormous cosmic web (for an excellent review of galaxy formation in general, see Benson 2010). Most stars in the universe form after the epoch of reionization, in these large galaxies and clusters over the next 12 billion years. Large scale structure is much easier to observe, and massive galaxies like our own are naturally studied extensively. Many of the low-mass halos that hosted stars prior to reionization were at some point dynamically disrupted, and the stars and dark matter were incorporated into larger galaxies.

However, the small, low-mass halos that were so important during that first billion years do not completely fade from history. Today, any small halo that also hosts stars is known as a "dwarf galaxy", or simply a "dwarf". Many of these dwarfs are now referred to as the "substructure" of large galaxies, which influence them gravitationally. Over the past few decades, our knowledge of this substructure has expanded rapidly, as we are able

to find dwarfs that are increasingly faint surrounding our own Milky Way (Mateo, 1998; Belokurov et al., 2007). Some of these dwarfs are thought to be the fossil remnants of the first galaxies (Ricotti & Gnedin, 2005), as star formation after the epoch of reionization can be suppressed by a variety of mechanisms (Gnedin & Kravtsov, 2006; Slater & Bell, 2013). The gas-rich environments required for the formation of massive globular star clusters known to surround all large galaxies may occur during the merger of some of the dense, clumpy substructure (Kravtsov & Gnedin, 2005; Brodie & Strader, 2006). We can even watch the process of dwarfs being destroyed by tidal forces, leaving their stars behind in streams that are incorporated into the Milky Way's outskirts (Belokurov et al., 2006). In the CDM paradigm, one cannot tell the complete story of a galaxy without understanding the story of each of the smaller protogalaxies that went into it, and the dwarfs that surround it. The present-day universe stands on the shoulders of dwarfs.

In this thesis, I will shed further light onto the "small-scale structure" of the universe. In the remainder of this introduction, I outline the relevant concepts and review the existing literature on the formation of Population III (Pop III) stars, high-redshift galaxies, and globular clusters. I also briefly review the numerical techniques crucial to attacking these problems theoretically. In Chapter 2, I present a series of hydrodynamic cosmological simulations that study the conditions of Pop III star formation, as well as the immediate feedback effects they have on their surroundings. In Chapter 3, I explore the subsequent evolution of the halos that hosted the first stars, and quantify the cosmic transition to "normal" (Pop II) star formation. In Chapter 4, I present a semi-analytical model that follows the assembly of a Milky Way-like galaxy from the epoch of reionization to the present epoch, that explains the formation of the Milky Way's massive globular star clusters. Finally, I conclude with several afterthoughts that bridge this work, and explore potentially relevant astrophysical modeling that will enhance our understanding of this field.

1.2 The First Stars

By gazing deeper and deeper into the sky, Astronomers find light that has been traveling for longer and longer intervals - such is the consequence of a finite speed of light in the paradigm of Einstein's relativity. This principle implies that we are granted the remarkable

ability to look directly into the past. Unfortunately, for a variety of reasons discussed in this thesis, it is likely that we will still never be able to observe the first individual stars that formed in the universe, which is a shame because we have many reasons to believe that the unique environment of the early universe imparted some truly bizarre properties onto these stars. This section largely draws content from several excellent existing reviews (Bromm & Larson, 2004; Wise, 2012; Bromm, 2013).

Let us consider the nature of primordial gas prior to the onset of star formation. The process during the early stages of the Big Bang that produced the first elementary particles, which then combined into atoms of the elements composing the majority of baryonic matter today is known is Big Bang Nucleosynthsis (Alpher et al., 1948). Despite this theory's relative age, the methods employed have proven fairly robust. However, the landmark study of Burbidge et al. (1957) demonstrated that only Hydrogen, Helium, and a very small amount of Lithium is formed through this process. The remaining elements found in the universe today must have been produced in the interior of stars.

The dearth of heavier elements, or "metals", leads to potentially significant consequences on the way that the first primordial gas clouds collapse into stars. To see this, we recall that metals play a critical role in present-day star forming gas to aid the gas in cooling and fragmentation during collapse, giving form to the statistical distribution of stellar masses at the time of their birth, or Initial Mass Function (IMF) (for a review of present-day star formation, see McKee & Ostriker (2007)). Without this cooling mechanism, the temperature of star-forming gas remains excessively high, which in turn affects the scale at which the gas can continue its rapid collapse to a protostar, or the Jeans mass of the gas (Jeans, 1902), which is approximately equal to the Bonner-Ebert mass (Clarke & Bromm, 2003), and can be expressed as:

$$M_J \approx 700\,\mathrm{M}_\odot \left(\frac{T}{200\ \mathrm{K}}\right)^{3/2} \left(\frac{n}{10^4\ \mathrm{cm}^{-3}}\right)^{-1/2}. \tag{1.1}$$

Where T is the temperature and n is number density of the gas cloud. Since a higher temperature implies a higher Jeans Mass, the clumps that formed the first stars were inherently excessively massive, which in term correlates to rapid accretion rate onto the pro-

tostellar core. This does not implicitly guarantee that the first stars themselves were very massive, but that is just what the first three-dimensional simulations capable of resolving these clouds and modeling the proper chemistry at the turn of the millennium revealed (Bromm et al., 1999; Abel et al., 2000). The ubiquity of these massive stars appearing suggested metal-free gas a very different IMF from all stars that we know of today. Because of this difference, it is necessary to distinguish these primordial star (Population III, or Pop III for short) from subsequent generations.

Though the chemistry of gas in the primordial universe is substantially simpler than present due to the small number of atomic species in existence, properly modeling the properties of hydrogen gas alone has proven to be far more difficult and uncertain than naive expectation may suggest. Indeed, given that Hydrogen is the primary constituent of the universe even today, much effort has been invested in its astrochemical modeling. Between the emission of the CMB and the epoch of reionization, the primary form of hydrogen in the atomic neutral phase. However, the key to understanding the formation of the first stars lies in the small fraction of molecular hydrogen, or H_2 in the densest regions of the pre-galactic halos. Most H_2 in the present day is produced on dust grains, small amounts can be formed in the primordial universe via a reaction involving H^-. Through ro-vibrational transitions, H_2 can cool gas to low temperatures, allowing this dense gas to undergo gravitational collapse necessary for protostar formation. The details of this process must be illuminated in great detail if we are to constrain the nature of Pop III stars ab initio. See Bromm & Larson (2004) for a detailed review.

Here, I will briefly mention that developments in understanding the protostellar chemistry and physics of primordial gas and how they have advanced the field, particularly focusing on phenomenological modeling simulations, which are at present the only ways to study these primordial objects. It was first realized by Saslaw & Zipoy (1967) that H_2 was the most important coolant in primordial gas. Once the CDM paradigm of galaxy formation became accepted, it became clear that low-mass structures relied on this cooling mechanism to form the first stars (Couchman & Rees, 1986; Haiman et al., 1996). The landmark study of (Tegmark et al., 1997) constrained the functional form for the minimum halo mass that can undergo H_2 cooling. The 3D simulations of (Abel et al., 1998; Bromm

et al., 1999; Abel et al., 2000) were the first to offer fully self-consistent ab-initio physical representations of the massive regions that formed the first stars. The next generation of simulations by the same groups probed deeper into the core, and concluded that primordial stars were inevitably massive (Abel et al., 2002; Bromm et al., 2002; Yoshida et al., 2003). These conclusions rested on the absence of fragmentation prior to the formation of the core, and the very high mass accretion rate onto the core owing to the large Jeans mass of the cloud.

For the next several years, studies of Pop III star formation focused on simulating environments that slightly deviated from the fully primordial regime. H_2-dissociating radiation in the Lyman-Werner band was identified as an important mechanism that could suppress early star formation (Machacek et al., 2001; Johnson et al., 2007; Wise & Abel, 2007; O'Shea & Norman, 2008). Others (Yoshida et al., 2007; Johnson & Bromm, 2006) explored how Pop III star formation may evolve once hydrogen becomes significantly ionized, finding that the characteristic mass of this "second" generation of stars to be smaller. Other works moved explicitly beyond the primordial regime, and studied gas that was somewhat chemically enriched, but still in the domain of having lower metallicity than any observed star-forming regions in the local universe (Bromm et al., 2001a; Omukai et al., 2005; Smith et al., 2009). Their work focused on searching for a "critical" metallicity at which fragmentation of star-forming clouds begins to resemble present-day molecular clouds, where the IMF does not appear to explicitly depend on metallicity. In the meantime, studies that were both more sophisticated and systematic continued supporting evidence for a top-heavy initial mass function among the first stars (Yoshida et al., 2006; O'Shea & Norman, 2007), though they also suggested that the IMF may depend on environment and accretion history of the protogalaxy.

Most recently, a new generation of simulations employing the sink particle technique to follow the evolution of high-density primordial gas for a longer period of time than previously possible, and discovered that fragmentation could occur in this gas despite its primordial chemical composition (Stacy et al., 2010; Greif et al., 2011; Clark et al., 2011). Instead of collapsing onto a single, massive protostar, the primordial clump instead formed a Keplerian disk, distributing the central gas mass over a larger area and eventually allowing

for the formation of multiple distinct cores. Some other simulations suggest that previously unaccounted radiative processes during primordial accretion could blow away some gas, shutting down the growth of protostars prior to their achieving massive status (Yoshida et al., 2008; Stacy et al., 2012; Hosokawa et al., 2011). Simulations that still do not employ sink particles find a degree of fragmentation in primordial gas that was absent from the earlier works, but do not currently suggest that fragmentation during protostellar collapse is ubiquitous, or that it completely suppresses the creation of massive stars (Turk et al., 2009; Greif et al., 2012, 2013). Sink particles are a useful numerical tool, but ultimately may oversimplify the relevant physics. Once the particle is created, the mass of gas sunk into the particle can no longer interact with its environment. In simulations without sink particles, this gas may later merge into larger clumps, enhancing growth.

For now, let us assume that Pop III stars are on average very massive compared to present day stars, as the first generation of simulations suggested. This ansatz leads to two very significant implications regarding the strength of their ability to impact their environment via stellar feedback. First, let us recall that Planck's law suggests that massive stars are very hot at their surface, when combined with Wien's law, this makes them factories for enormous quantities of ionizing photons. Second, stellar explosions, or supernovae are common for massive stars. In particular, stars that are between $140\,M_\odot$ and $260\,M_\odot$ (Heger & Woosley, 2002) are subject to enormously powerful stellar explosions, called pair instability supernovae, each of which is capable of unbinding large quantities of gas. It is also useful here to assume that each protogalactic "minihalo" hosts only a single massive star. This assumption is justified by early simulations (e.g. Abel et al. 2002) and is easily understood once it is seen just how suppressive this stellar feedback really is to further star formation.

Ionizing photons and the first HII regions

Regardless of whether or not the IMF of primordial stars was different, metal-free populations are generally expected to produce 50% more ionizing radiation than modern-day, metal-enriched stars (Tumlinson & Shull, 2000). However, if the IMF in primordial gas is indeed discrepantly top-heavy, this emphasizes the difference all the more. The ionizing

photon production rate grows highly non-linearly with respect to stellar mass, and the production of photons per unit mass of a stellar population is expected to be as high as 10-20 times larger than a Pop II population with a normal IMF (Bromm et al., 2001b; Schaerer, 2002). The rapid ionization of gas does more than simply change the electrical state of the atom. Absorption of ionizing photons will immediately change the temperature of hydrogen to ~$10^4 K$, and as the ionizing photons travel at the speed of light, the gas temperature globally changes at a rate much faster than the thermal sound speed. This means that the ionizing wake creates a hydrodynamic shock, rapidly lowering the density in the central regions of the galaxy.

The relics of these vast ionizing wakes are analogs to the environments, known as HII regions, that surround massive stars today. However, these first HII regions extend over the entire virial radius of the galaxy and beyond, with a single star ionizing several kiloparsecs worth of hydrogen (Alvarez et al., 2006; Whalen et al., 2004; Kitayama et al., 2004; Abel et al., 2007). The conditions for star formation are not favorable within HII regions, as the density of the gas is too low and the temperature too high. Even when the gas becomes able to recombine again, The IMF of stars in the slightly ionized gas is expected to follow the "Pop III.2" model proposed in Yoshida et al. (2007) rather than the top-heavy IMF of the first Pop III stars. Typically, these HII regions recombine within 50 Myr (Wise & Abel, 2008).

Pair-instability supernovae

It is believed that the pair-production instability can operate in the interior of particularly hot stars at the end of their lives, and cause catastrophic explosions known as pair-instability supernovae (PISN) (Barkat et al., 1967; Bond et al., 1984; Woosley & Weaver, 1986; Heger & Woosley, 2002). When the temperature of the stellar core climbs above a certain level (~$5 \times 10^9 K$), electron-positron pairs may form when energetic gamma rays interact with each other and with atomic nuclei. The creation of electron-positron pairs modifies the equation of state in the core, causing the star to contract and heat up, leading to an explosive ignition of oxygen burning. What follows is an explosion unlike anything we've ever had the chance to observe. Unlike the typical supernovae known to occur for

massive stars (Type II SNe), this explosion leaves no remnant, and the entire mass of the core is expelled into the intergalactic medium. The total energy of the explosion depends on the initial mass, but is typically an order of magnitude larger than the energy of Type II SNe (Fryer et al., 2001; Heger et al., 2003).

Following a supernova, there is no longer any starlight to drive the ionizing front that carved out the HII region, and the shock is terminated. However, the supernova explosion results in a blastwave that can easily plow through and re-shock the newly diffuse and warm inter-stellar medium. The energy that is given off is indeed typically comparable to the binding energy of the entire galaxy, making it unsurprising that the supernova is sometimes single-handedly capable of expelling all gas from the protogalaxy (Bromm et al., 2003). The flip side of this powerful feedback mechanism is that the entire core of the exploding star is ejected, and it is loaded with atomic species heavier than helium. Just one of these explosions is enough to change the fate of an entire galaxy before it ever forms. If the cosmic landscape and the end of the dark ages was peppered with them, their impacts reverberate all throughout cosmic history. While there have been significant efforts to study galaxies that form in the aftermath of a fully resolved PISN (Wise & Abel, 2008; Greif et al., 2010), a more systematic study is needed to quantify the extent of their damage.

It is worth noting that stars that are too massive (with mass in excess of 260 $M_{\odot}$) or not massive enough (below 140 $M_{\odot}$) avoid the fate of pair instability supernova. In these cases, the fate of the star is less certain, but likely involves the formation of a black hole remnant, and potentially an explosion with an energy comparable to a type II SNe (Heger et al., 2003). Such explosions don't have quite the same dramatic impact of the PISNe described above, but are still significant sources of feedback when considering the low mass of protogalaxies. As for the black hole remnants, they are hypothesized to potentially be the seeds of modern-day supermassive black holes, which are observed to be extremely massive as far back as z~6. This thesis does not deal with black hole growth and feedback, but see Volonteri & Bellovary (2012) for a review.

1.3 The First Galaxies

Following the first stars, the next stage of cosmic evolution saw the rise of "conventional" galaxies that could sustain entire stellar populations, a multi-phase interstellar medium, and structure in the distribution of stars and gas that begin to resemble modern day galaxies. These galaxies are also at least in part responsible for permanently changing the ionization state, temperature, and chemical composition of the intergalactic medium. Current observational facilities are edging closer and closer to gathering data that will sufficiently constrain this epoch, allowing for comprehensive theoretical groundwork that bridges the gap between the emission of the cosmic microwave background and modern galactic structure.

For now, cosmological simulations remain the most sensible way to study this epoch. It would of course be ideal if we were able to continue our simulations of the universe by resolving the core of every individual star to form in each galaxy, like the studies mentioned in the previous section were able to. However, it is beyond the capacity of today's supercomputers to progress substantially far when simultaneously resolving protostellar and galactic processes. We must therefore employ subgrid physics imbued with the constraints on Pop III stars provided in the previous section, and work towards simulations that eventually produce realistic galaxies. Progress in this branch of research has been fruitful over the past decade. For other excellent reviews of this content, see Bromm & Yoshida (2011); Wise (2012).

The two groups who so successfully laid the groundwork of Pop III star formation (Bromm et al. 2002; Abel et al. 2002) would also go on to study their prolonged effects on galaxy formation (Wise & Abel, 2008; Greif et al., 2010). They were once again in agreement in showing just how powerful the feedback of single PISNe could shape the fate of an entire galaxy. But while these Pop III stars are truly awe-inspiring, we must now ask the following question: were Pop III stars really the dominant mode of star formation for a significant segment of cosmic history? It was quickly realized that the strong chemical feedback of these stars makes them "suicidal" (Yoshida et al., 2004). Gas that has been mixed with the plethora of metals released from a single PISN can hardly be called pristine

anymore, and indeed, it may already closely resemble gas in nearby star-forming regions. Will this gas still fragment inefficiently, thereby leading to a top-heavy IMF (Pop III), or will it already be able to fragment in the standard way observed in the present-day universe (Pop II)?

To answer these questions, the collapse of low-metallicity gas has been studied systematically by various groups. At first, the answer appeared simple, as (Bromm et al., 2001a) found that a single critical value of the metallicity, $log(Z/Z_{\odot}) = -3.5$ separated the two regimes, where Z is the total fraction of metals in the gas, and $Z_{\odot}$ is the solar value, which is taken at 0.02. This value was later reconfirmed by subsequent studies (Smith et al., 2009). However, the process is significantly complicated when dust chemistry is included in the calculation. Omukai et al. (2005) stress that gas-phase and dust-phase metals have different efficacy in aiding fragmentation, and that the critical metallicity can be as low as $log(Z/Z_{\odot}) = -5$ if enough dust forms in the interstellar medium, or is ejected from supernovae. Dust formation in the local universe is not well understood, and the prospects of constraining its properties in these galaxies will be challenging in the foreseeable future.

In any case, it is clear that only a small amount of metals are required for Pop II stars to begin appearing in the universe. Unlike Pop III stars, Pop II is thought to be fairly well constrained and can be modeled with the same formation criteria, feedback properties, and IMF that describe modern-day star formation. Though the works that focus exclusively on Pop III star formation are interesting, simulations that deal with both Pop III and Pop II star formation are essential to paint a realistic picture of the epoch.

Ostriker & Gnedin (1996) and Gnedin & Ostriker (1997) were among the first of such modern efforts, where the roles of Pop III and Pop II were distinguished in heating, enriching, and reionizing the universe in full hydrodynamic simulations with radiative transfer. Already, these simulations identified that galaxies alone could be the drivers of the epoch of reionization, and provided the metals to enrich the intergalactic medium. Independent of the chemical feedback, the era of protogalactic Pop III "minihalos" that were driven by H_2 cooling is also brought to an end once the "virial temperature" of the host dark matter halo reaches $10^4 K$, allowing atomic hydrogen to be collisionally ionized and cool via recombination. The virial temperature is given as follows (from Bromm & Yoshida 2011):

$$T_{vir} = \frac{\mu m_H V_c^2}{2k_B} \approx 10^4 \left(\frac{\mu}{0.6}\right) \left(\frac{M}{10^8\,\mathrm{M}_\odot}\right)^{2/3} \left(\frac{1+z}{10}\right) \mathrm{K}. \quad (1.2)$$

Where μ is the mean molecular weight, and M is the total mass of the halo. Oh & Haiman (2002) showed that these galaxies cooled much more efficiently, with substantially more cool gas being transported to the dense regions of the galaxy where star formation can proceed on a sustained level. More sophisticated simulations (Ricotti et al., 2002a,b; Tornatore et al., 2007) that featured new methods for radiative transfer and separate spectral energy distributions for Pop III and Pop II stars further highlighted the failure of the mini-halos that relied exclusively on H_2 to efficiently form stars, and showed that the ionizing feedback of Pop III stars could suppress star formation in nearby galaxies. Semi-analytical modeling by other groups (Scannapieco et al., 2003; Yoshida et al., 2004; Schneider et al., 2006; Trenti & Stiavelli, 2009) serve as an alternative and useful approach to complement simulations, and have focused on the global transition from Pop III to Pop II star formation using physically-motivated parameters.

Despite the success of these studies, we are still uncertain of whether Pop III star formation was shut off by internal or external sources, and whether Pop III star formation can continue once Pop II star formation is in full swing. A pair instability supernova of a $170\,\mathrm{M}_\odot$ stars releases $80\,\mathrm{M}_\odot$ of metals (Heger & Woosley, 2002), which is enough to enrich all of the gas above the critical metallicity even in a galaxy of total mass as high as $M_h = 10^8\,\mathrm{M}_\odot$, assuming that the ejecta stay in the galaxy. If the ejecta leave the galaxy, whether they be from a PISN or from the metal feedback of an early Pop II cluster, then they can enrich external galaxies and cause them to skip the primordial phase altogether. The works referenced above, as well as a few concurrent efforts have begun to constrain this process (Pawlik et al., 2011; Maio et al., 2011; Wise et al., 2012a; Johnson et al., 2013; Pawlik et al., 2013), but more systematic effort must be taken to fully understand and quantify the results.

In a similar vein, if ionizing radiation escapes from the galaxies efficiently, star formation can be delayed, either by the destruction of H_2 by Lyman Werner radiation, or in some cases, by the expulsion of all gas via intergalactic ionizing shocks. Such a process

ultimately sets a lower limit to the mass galaxies can have during and after the epoch of reionization (Gnedin, 2000). However, it has thus far been notoriously difficult to accurately observationally measure the fraction of ionizing radiation in a given galaxy that actually escapes into the intergalactic medium. In the meanwhile, theoretical research seems to diverge on the answer. Escape fraction of early galaxies (Alvarez et al., 2006) estimate that it should be 70-90% for the HII regions enveloping the massive first stars using semi-analytical models. Radiative hydrodynamic simulations of Wise & Abel (2008) estimated that while 25% of the IGM surrounding a Pop III star relic HII region can be fully ionized, the gas typically recombines within 50 Myr. This complicates the matter by showing that even though the escape of the Pop III star's ionizing radiation was very efficient, it ultimately doesn't make much of a difference. Wise & Cen (2009) explored how the escape fraction can be quite high in a galaxy after a Pop III star has spent its life carving out a hole in the interstellar medium. If there are other stars in the galaxy at this time, their feedback can most certainly escape at high rates. Other researchers (Gnedin et al., 2008) maintain that the escape fraction is quite low in low-mass galaxies, particularly early in the history of Pop II SF, when the efficiency of star formation is still quite low.

The latest estimates from the groups leading the most ambitious observing campaign with the Hubble Space Telescope show that a significant fraction of escaping ionization must come from sources too faint to have yet been seen (Robertson et al., 2013). This highlights the need for more extensive theoretical work to understand these faint galaxies, as teasing out any physical information from future detections will take considerable effort. Some semi-analytical modeling that assumes very high escape fractions from low-mass galaxies claim to match all observational constraints of reionization (Robertson et al., 2013; Ahn et al., 2012; Alvarez et al., 2012; Trenti et al., 2009), but it is still unclear whether this is the proper way to treat these galaxies. Instead of relying on Pop III stars, we may need to search for other physical mechanisms that steer the history of reionization.

1.4 Globular Cluster Formation

We turn now to modern day massive galaxies, which were assembled gradually from the coalescence of many low-mass halos described in the previous section, as well as additional

growth from cosmic filaments. Some of the oldest stars in the Milky Way which may hold the clues to the early stage of its formation rest in massive, compact star clusters known as globular clusters (GCs). Unlike the majority of the Milky Way's stars, which reside in the disk and the central bulge, globular clusters are distributed in a sphere that extends from the galactic center all the way to the outer part of the halo (Shapley, 1918). This suggests that their formation mechanism is considerably different from the stars that are forming in the disk. About 150 exist in the Milky Way, and they appear to also exist abundantly around massive nearby spiral galaxies, as well as giant ellipticals. Smaller dwarf galaxies like the Large Magellanic Cloud also have a few, but for now, let us focus on the globular cluster populations of massive galaxies. For an excellent review of what is known about extragalactic globular clusters and how they are related to galaxy formation, see (Brodie & Strader, 2006).

Globular clusters have long been of interest to astronomers for a wide variety of reasons. Because they appear to be a dense ball of stars that is essentially evolving in isolation, they are as close as one can get to an idealized N-body problem, where gravity governs the interactions of $N \approx 100,000$ stars. In addition, the stars in a globular cluster all appear to have the same age, and metallicity, implying that the cluster formed in a single burst. This is not so unusual for stellar clusters, but is particularly striking when considering the large number of stars within each one. The stellar mass of a single globular cluster exceeds the stellar mass of many individual dwarf galaxies. But perhaps the most interesting thing about them is their potentially very old age - in fact, they are so old that they were once estimated to be older than the age of the universe as derived from cosmological parameters.

While theories that attempt to explain GC formation have been around as long as theories of galaxy formation (Peebles & Dicke, 1968), less definitive progress has been made in understanding them. The trouble is, the high stellar density and large mass of globular clusters suggests they formed in very large, dense molecular clouds that are simply absent from the Milky Way, or at least the part that we can observe. The resolution needed to study such molecular complexes in other nearby galaxies is hard to achieve. Furthermore, it may just be that galaxies like the Milky Way are no longer capable of forming such massive star clusters - after all, the age of GC's suggest that the Milky Way was a very different place

when they were forming.

A major milestone occurred with the Hubble Space Telescope's discovery that massive young star clusters were forming during the merger of the Antennae galaxies (Whitmore et al., 1999). This mode of star formation were the first observations for a plausible progenitor to the globular clusters observed around massive galaxies: the clusters were massive, dense, and likely destined to be ejected from the disks of their host galaxies due to the dynamical merging event. Star formation can then be triggered by strong shocks or high pressure in the interstellar medium (van den Bergh, 2001; Ashman & Zepf, 2001). But galaxy mergers are fairly rare events in today's universe. The Milky Way probably hasn't had one for at least 5 billion years, and it will be another 5 billion years before it merges with Andromeda. Fortunately, the CDM paradigm offers a degree of reconciliation. The Milky Way is built up of many smaller progenitors that all merged with one another at very high redshifts. Those progenitors carried plenty of low-metallicity gas, and each merger could result in the fireworks of the Antennae: young, massive, dense star clusters that were destined to become GCs after billions of years of dynamical evolution. For a review of young, massive star clusters see Portegies Zwart et al. (2010).

One way to explore this formation mechanism is via cosmological simulations. Kravtsov & Gnedin (2005) was a landmark study that for the first time actually resolved molecular complexes in the buildup of the Milky Way that could plausibly generate clusters as dense and massive as the ones in the Antennae. In their simulations, these mergers largely happened after the epoch of reionization, when the milky way progenitor galaxies were massive enough to host disks of their own. However, these simulations were not run to the present day, as maintaining this resolution quickly becomes computationally expensive. Other simulations focused on a potential for them to form in low-mass halos prior to the epoch of reionization (Bromm & Clarke, 2002; Boley et al., 2009). However, once again, due to computational limitations, these simulations have thus far been only carried out at high redshifts and do not progress to the present.

The properties of the globular clusters today can be used as excellent constraints on whether the theories are really doing a good job in describing their formation. Ideally, a theory should be able to explain the origin of the observed distributions of mass, metallicity,

age, and positions. In the Milky Way, we have good knowledge of all four properties, though all but position are typically derived indirectly through observations of brightness (mass), color (metallicity, age), and the absence or presence of stars above a certain main sequence mass threshold (age).

We can start with the mass function, which has been successfully reproduced in a number of studies. The current mass distribution of globular clusters is significantly different from the initial mass function of clusters as measured in starbursting mergers like The Antennae (Fall & Zhang, 2001). This, however, can be explained by the gradual destruction of globular clusters over cosmic time (Gnedin & Ostriker, 1997). Prieto & Gnedin (2008) showed that the globular clusters simulated in Kravtsov & Gnedin (2005) would undergo dynamical mass loss and eventually have a log-normal mass distribution centered around $10^5\,M_\odot$, matching observations. However, there is still ongoing debate within the community about whether external or internal processes are the most relevant for the mass evolution of star clusters (Gieles, 2010).

Though we know that globular clusters are old, it is difficult to judge exactly how old. Previous estimates suggested ~20 Gyr, which is older than the age of the universe as derived from cosmology. Thanks to improved distance measurements and understanding of stellar evolution, the estimates have since come down, and span the range of 6 - 15 Gyr with a typical error of 1 Gyr (Marín-Franch et al., 2009). This still means that there is an uncomfortable tension between the oldest GCs and the age of the universe as estimated by cosmology (Gnedin et al., 2001). The uncertainty for age estimation limits our ability to constrain whether the oldest GCs form prior to or after the epoch of reionization. This distinction makes a significant difference when attempting to determine the mass of protogalaxies that hosted globular cluster formation, as only massive halos can host sufficiently dense gas after reionization. Some authors have explored the idea of globular clusters even being the sources of reionization (Ricotti, 2002; Griffen et al., 2010, 2013; Katz & Ricotti, 2013).

Globular clusters are distributed spatially in a spherical profile centered on the inner regions of the galaxy. This property was once used to estimate the distance between the sun and the galactic center. Typically, galaxy clusters are thought to be associated either

with the bulge or the stellar halo. Their relative positions may offer clues to their formation mechanisms, though the old age of GCs implies that the orbits have had much time to evolve. Fully self-consistent calculations of orbital evolution are difficult in the context of CDM, but they were attempted by Prieto & Gnedin (2008). In general, observed globular clusters seem to have a more compact spatial distribution than what is suggested by the models of Prieto & Gnedin (2008). However, their calculations did not take into account some hydrodynamic effects that may have later brought the globular clusters closer in to the galactic center (Naab et al., 2009).

The final well-studied, but not well-understood constraint is the apparent bimodality of the metallicity distribution of globular clusters. This bimodality appears not only in the Milky Way's system, but also in the globular cluster population of giant ellipticals (Zepf & Ashman, 1993) and nearby spirals (Zinn, 1985). Clusters are generally divided into two groups: metal-rich and metal-poor, or alternatively "red" and "blue", respectively. The discrepancy in color comes about because of the discrepancy in metallicity. Often, the two populations are considered separately, though their internal properties are fairly similar. One hint stems from the fact that the red clusters are typically associated with the "bulge" population that are typically closer to the galactic center, while blue globular clusters are associated with the halo (Brodie & Strader, 2006). Studies that attempt to semi-analytically model the formation of the two population generally use different mechanisms for their formation (Beasley et al., 2002; Griffen et al., 2010). The general tendency is for blue globular clusters to be associated with the old population that may have formed prior to reionization, with some forming in dwarf galaxies that were accreted to the galactic halo. Red clusters are assumed to be more closely associated with the [early] disk of the main galaxy. This may imply that globular clusters have an inherent age-metallicity relation, but it is not clear whether this is observed (Forbes & Bridges, 2010).

Some additional challenges to the modeling community have recently been exposed. For one thing, if globular clusters were ever associated with dark matter halos, their appears not to be much of evidence of this in their internal velocity dispersions, and the dark-to-stellar mass ratio for several globular clusters has been constrained to be $M_{DM}/M_* < 1$ (Conroy et al., 2011). Another recent issue has been the revelation that some individual

globular clusters in fact have two distinct stellar populations within them, implying that their formation could have taken place over several different bursts (Conroy & Spergel, 2011). As the secrets of globular clusters continue to reveal themselves to us, it becomes increasingly clear that understanding them may be the missing link towards a complete model of galaxy formation.

1.5 Numerical Simulations

The disadvantage of working in the regime of the early history of the universe are plentiful and obvious. Though it is possible to observe the universe at early times, and new observing campaigns yield a rapidly growing sample of galaxies that are at increasingly high redshifts, we are still always limited by the inverse square law of light and the relative faintness of the first galaxies due to the small number of stars that initially form within them.

On the other hand, there is one very significant advantage to working in this regime: as alluded to in the opening paragraphs of this thesis, the initial conditions of the universe at the start of galaxy formation are well understood, and therefore define a fairly well-posed problem that can be solved ab initio. Data from the Wilkinson Microwave Anisotropy Probe (WMAP) provides us with the best constraints on the expansion rate of the universe, H_0, as well as the density of matter and dark energy (Komatsu et al., 2011). These are the the numerical parameters necessary to constrain the Friedmann equation, which in a flat universe simplifies to:

$$\frac{H(t)}{H_0} = \left(\frac{\Omega_m}{a^3} + \Omega_\Lambda + \frac{\Omega_r}{a^4}\right)^{1/2} \tag{1.3}$$

Where Ω_m, Ω_Λ, Ω_r are the ratios of the density of matter, dark energy, and radiation, respectively, to the critical density of the universe. $H(t)$ is the Hubble Constant which measures the expansion rate of the universe as a function of time, and H_0 is the value of the Hubble Constant at the present epoch. The latest measurements are: $\Omega_m = 0.28$, $\Omega_\Lambda = 0.72$, $H_0 = 70 km/s/Mpc^3$, and $\Omega_r \approx 0$. a is the universal expansion factor, which is defined in terms of the cosmological redshift, z, as: $a = \frac{1}{1+z}$. Equation 1.3 then describes the expansion rate of the universe. We can combine it with the equations of hydrodynamics, which simply guarantee continuity of mass, momentum, and energy, as well as the polytropic equation

of state for the gas:

$$\frac{\partial \rho}{\partial t} + \nabla \rho \mathbf{u} = 0, \tag{1.4}$$

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\nabla \Phi - \frac{\nabla P}{\rho}, \tag{1.5}$$

$$\frac{\partial E}{\partial t} + \nabla \cdot [(E + P)\mathbf{u}] = -\rho \mathbf{u} \cdot \nabla \Phi + (\Gamma - L), \tag{1.6}$$

$$\varepsilon = \frac{1}{\gamma - 1}\frac{P}{\rho}. \tag{1.7}$$

Where ρ is density, P is pressure, **u** is velocity, ε is internal energy, E is total energy, Γ is the heating rate, L is the cooling rate, Φ is the gravitational potential, and γ is the polytropic index.

Cold dark matter can be represented by numerical Lagrangian particles that only interact via gravity. The only additional equations needed to govern dark matter particles are:

$$\frac{\partial \mathbf{r}}{\partial t} = \mathbf{u}; \frac{\partial \mathbf{u}}{\partial t} = \nabla \Phi, \tag{1.8}$$

$$\nabla^2 \Phi = 4\pi G\rho - \Lambda. \tag{1.9}$$

Where the latter is Poisson's equation in a cosmological context, where ρ is the total matter density, G is the gravitational constant, and Λ is the cosmological constant for dark energy. Both gas and dark matter are affected by the gravitation induced by Poisson's equation. These equations can be embedded into simulation codes. Since the early work of Evrard (1988), cosmological hydrodynamical simulations have been a fundamental method in theoretical astrophysics. For all studies presented in this thesis, I use the Adaptive Refinement Tree (ART) code, (Kravtsov, 1999). It employs the adaptive mesh refinement technique to discretize gas into a grid of "cells". Each cell is divided into more "child"

cells if the density becomes sufficiently high, making the grid especially fine where high resolution is needed to capture the relevant physics. Each numerical dark matter particle represents an equal mass of material, and together they broadly trace the dark matter distribution. The motions of these particles are computed using N-body dynamics techniques. Periodic boundary conditions are employed in each "box" to ensure the conservation of mass. We simulate a fixed comoving volume, which means that the proper physical volume is different at each time.

The code has undergone decades of development and has very sophisticated treatment of a variety of gas processes. The most recent incarnation prior to my modifications is described in Gnedin & Kravtsov (2011). Some of the processes taken into account include: primordial chemistry, cooling by metals, various heating processes, star formation, stellar feedback by SNe, and coupled radiative transfer. The radiative transfer technique employed is called OTVET, and is described in Gnedin & Abel (2001).

Each simulation begins at $z \approx 150$, with initial conditions generated randomly from the known form of the primordial power spectrum for matter as given by WMAP. From there, the simulation marches forward in time incrementally in "timesteps", advancing the equations for each cell and particle. At some epochs, which can be specified at the beginning of each run, the simulation writes a "snapshot" file which contains the positions and velocities of all stellar and dark matter particles, as well the hydrodynamical properties of each gas cell. These snapshots are run through a "halo finder" to generate a catalog of the locations and masses of every region of the universe which meets a minimum requirement for mass and overdensity.

In the case of the simulations described in Chapters 2 and 3, galactic profiles were constructed using the analysis routines of Zemp et al. (2012). These profiles were used in conjunction with all of the raw data about the particles and cells to perform the subsequent analysis necessary to obtain our results. In the case of the simulations used in Chapter 4, I used merger trees from Kravtsov et al. (2004) to trace the fate of every halo from the time of formation until $z = 0$.

CHAPTER 2

The effects of Population III stars on their host galaxies

2.1 Introduction

A natural consequence of Big Bang Nucleonsynthesis is that the first stars in the universe formed from gas that was completely devoid of elements heavier than lithium. Since such conditions do not exist in any known star-forming regions today, star formation in the primordial regime has thus far only been explored by theory and simulations. From first principles, it can be deduced that gas which is free of metals would not be able to cool efficiently, and would therefore inherently have a higher Jeans mass than stars forming in metal-enriched gas. Bromm et al. (1999) showed that metal-free gas settles into disks, then fragments into clumps with $M_J \approx 10^3\,\mathrm{M}_\odot$, which undergo runaway collapse to densities of $n_H > 10^8\,\mathrm{cm}^{-3}$. In a follow-up study, it was shown that this process was robust to initial conditions in Smoothed Particle Hydrodynamics (SPH) simulations (Bromm et al., 2002). Using independent Adaptive Mesh Refinement (AMR) grid techniques, Abel et al. (2000) confirmed that such dense clumps and cold pockets can indeed form in primordial gas. Abel et al. (2002) presented even more realistic simulations, and concluded that single massive stars would form via H_2 cooling at the center of these clumps, and the radiative feedback would halt accretion onto the star and prevent further star formation in the parent halo. Subsequent work with higher resolution reaffirmed these conclusions (Yoshida et al., 2006; O'Shea & Norman, 2007). However, recent studies with both SPH (Stacy et al., 2010, 2012; Clark et al., 2011) and grid techniques (Turk et al., 2009; Greif et al., 2011) have suggested that angular momentum imparted on the gas during collapse could still lead to fragmentation, causing the cores and the resulting stars to be substantially smaller. Ultimately, the next generation of super-zoomed simulations will need to

follow the proto-stellar systems for longer periods of time, with more detailed treatments of radiative transfer and magnetohydrodynamics, to conclude decisively on the true nature of Pop III stars (Greif et al., 2012).

Though we do not yet have any direct observational evidence for Pop III stars, we know that they must have existed in some form, as gas in the universe inevitably transitioned from having primordial composition to being enriched with enough metals to allow present-day star formation to commence. The nature of this transition is of key importance for understanding the dawn of galaxy formation: if Pop III stars did indeed form with a top-heavy initial mass function (IMF), a significant fraction of them may have been prone to end their lives as pair-instability supernovae (PISNe). Such supernovae generate up to ten times as much thermal energy as Type II SNe (Heger & Woosley, 2002), quickly heating the gas in their host halos. In addition, metal-free stars are able to produce enormous quantities of ionizing photons: a metal-free star will always have a higher surface temperature than an enriched counterpart of equal mass (Schaerer, 2002; Tumlinson & Shull, 2000). If the IMF is indeed top-heavy, the effect is even more drastic as many stars would have their emission spectra peak in the UV regime. Both the supernovae and the ionizing photons serve to heat and disperse neutral gas in the vicinity the star, creating a large HII region (Whalen et al., 2004; Alvarez et al., 2006; Johnson et al., 2007; Abel et al., 2007), and delaying the onset of steady, continuous star formation. On the other hand, PISNe release a large amount of metals, rapidly enriching previously primordial gas (Wise & Abel, 2008; Greif et al., 2010; Maio et al., 2011; Wise et al., 2012a), and potentially leading to quick termination of the Pop III epoch (Yoshida et al., 2004). The balance of these effects works to determine star formation rates in galaxies and their subsequent evolution.

While Pop III stars are no longer thought to be a major driver of the global reionization of the intergalactic medium (Ciardi et al., 2000; Ricotti et al., 2002b; Ricotti & Ostriker, 2004; Mesinger et al., 2009), constraints on the cosmological electron scattering optical depth from WMAP suggest that halos less massive than $10^8\,M_\odot$ may have contributed to the photon budget at the beginning of reionization. In order for this scenario to work, low mass halos must permit the escape of ionizing photons effectively (Alvarez et al., 2012; Ahn et al., 2012). Pop III stellar feedback has been explored as a mechanism to create

windows of time during which such high escape fractions are made possible (Wise & Cen 2009, but see Gnedin et al. 2008; Ricotti et al. 2008).

We revisit these important conclusions with novel high-resolution cosmological simulations that feature separate star formation criteria and feedback prescriptions for Pop III and Pop II stars. While we cannot resolve all stages of their formation, we do resolve the clumps of gas on a ~1 pc scale which inevitably collapse into Pop III stars. In addition to the commonly accepted parameters for Pop III formation and feedback, we also explore models in which the radiative and SNe feedback are taken at extreme values, as well as models in which the IMF of Pop III stars is taken to be identical to Pop II counterparts.

Our analysis in this paper focuses on the dynamical effects that Pop III stars can impart onto their host galaxies. We quantify the ability of Pop III stars to suppress star formation, expel gas, and enrich the medium both within and outside of the galaxies in which they form. Using a wide suite of simulations, we show that mass resolution and mesh refinement criteria affect the derived importance of Pop III stars. In a forthcoming second paper (Muratov et al. 2013b, see Chapter 3), we will explore the nature of the transition between Pop III and Pop II star formation, and assess the relative importance of feedback effects from the two stellar populations.

2.2 Simulations

We perform cosmological simulations with the Eulerian, gasdynamics+N-body adaptive refinement tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999, 2003; Rudd et al., 2008). The latest version of the code incorporates a new phenomenological prescription for molecular hydrogen formation on dust grains and self-shielding, as well as shielding by dust introduced in Gnedin et al. (2009) and developed further in Gnedin & Kravtsov (2011). Having such a detailed account of molecular gas, as well as excellent spatial resolution at high redshift makes it practical to consider a star formation recipe that is also based on molecular gas. This formulation has previously been shown to much better reproduce the Kennicutt-Schmidt relation for high-redshift, low-mass, and low metallicity galaxies, and it now enables more realistic simulations of the early universe (Tassis et al. 2008; Gnedin & Kravtsov 2010, 2011). Radiative transfer, including Hydrogen and Helium ionization, as

well as Lyman-Werner H_2 dissociating feedback, is computed using the OTVET approximation (Gnedin & Abel, 2001), employing the same Eddington tensor considered in that work. Stellar particles are treated as point sources of radiation. Diffuse radiation from the CMB, Compton heating, recombination, and bremsstrahlung is also computed.

For our baseline runs, we use a 1 h^{-1} Mpc comoving box and the WMAP-7 cosmology (Ω_m = 0.28, Ω_Λ = 0.72, h = 0.7, σ_8 = 0.817, Ω_b = 0.046, Ω_{DM} = 0.234). Additional runs were performed with 0.5 h^{-1} and 0.25 h^{-1} Mpc comoving boxes to explore the effects of mass resolution (hereby referred to as H Mpc and Q Mpc runs, respectively). The details for each run performed in our simulation suite are presented in Section 2.2.3. Numbers quoted in this paragraph, as well as Sections 2.2.1 and 2.2.2 refer to our baseline 1 h^{-1} Mpc runs. These runs start with a 256^3 root grid, which sets the DM particle mass to m_{DM} = $5.53 \times 10^3 \mathrm{M}_\odot$ and the base comoving resolution of 5.56 kpc. We employ Lagrangian refinement criteria, refining cells when the DM mass approximately doubles compared to the initial mass in the cell, specifically, when it exceeds $2 \times m_{DM} \times \frac{\Omega_m}{\Omega_{DM}} \times 0.8$, or the gas density surpasses an approximately equivalent value modulated by the cosmic baryon fraction $0.3 \times m_{DM} \times \frac{\Omega_m}{\Omega_{DM}} \times 0.8$. In both refinement conditions, the extra factor of 0.8 is the split tolerance. We use a maximum of 8 additional levels of refinement, giving us a final resolution of $10^6 h^{-1}/256/2^8 \approx 22$ comoving pc. Since we are studying high-redshift galaxies, it is important to note that this translates to about 2 physical pc at the endpoint of our simulations, z = 9, and 1 physical pc at z = 20 when the first stars begin to form. This spatial resolution is sufficient to capture the detailed multi-phased structure of the interstellar medium (ISM) (e.g. Ceverino & Klypin 2009).

In order to simulate several representative regions of the universe, we employ the "DC mode" formulation presented in Sirko (2005) and Gnedin et al. (2011). Running simulations with different DC modes allows us to sample cosmologically over- and under-dense regions without actually changing the total mass within each box. A single parameter Δ_{DC}, which is constant at all times for a given simulation box, represents the fundamental scale of density fluctuations present in the box. At sufficiently early times, when perturbations on the fundamental scale of the box are in the regime of linear growth, Δ_{DC} is related to the overdensity, $\delta_{DC}(a) \equiv D_+(a)\,\Delta_{DC}$, where D_+ is the linear growth factor. The expansion rate

of the individual box relates to the expansion rate of the universe by the following relation, which is Equation 3 of Gnedin et al. (2011):

$$a_{box} = \frac{a_{uni}}{[1+\Delta_{DC} D_{+}(a_{uni})]^{1/3}}, \tag{2.1}$$

where a_{box} is the local scale factor of the simulation box, while a_{uni} is the global expansion factor of the universe. For this study, we have used three different setups with $\Delta_{DC} = -2.57, -3.35$, and 4.04, labeled 'Box UnderDense–', 'Box UnderDense+' and 'Box OverDense', respectively. At the endpoint of our simulations, $z = 9$, these values translate to overdensities of $\delta_{DC} = -0.257, -0.335$, and 0.404, respectively. Boxes UnderDense– and UnderDense+ have negative DC modes, implying they are underdense regions of the universe. However, while Box UnderDense– is representative of a void, and hosts only low-mass galaxies which collapse relatively late, Box UnderDense+ hosts the first star-forming galaxy among all three simulation boxes. This galaxy is also more massive than any of those in Box OverDense until $z \approx 13$.

Box OverDense hosts several massive star-forming galaxies which statistically dominate the sample of simulated galaxies. The H Mpc and Q Mpc boxes used $\Delta_{DC} = 5.04$ and 6.11, respectively. These runs are primarily designed to explore the earliest possible epoch of Pop III star formation, tracing only the most overdense regions with even higher mass resolution. None of our simulations continue past $z = 9$, as the boxes are too small to capture the nonlinear growth of large-scale modes at later epochs.

We construct catalogs of halo properties from simulation outputs using the profiling routine described in Zemp et al. (2012). We take the virial radius, R_{vir}, as the distance from the center of a halo which encloses a region that has an overdensity of 180 with respect to the critical density of the universe.

2.2.1 Population II star formation

Following Gnedin et al. (2009) and Gnedin & Kravtsov (2011), we set the threshold for Pop II star formation in a gas cell when the fraction of molecular hydrogen exceeds the threshold $f_{H_2} \equiv 2n_{H_2}/n_H = 0.1$. Tests described in the above studies showed that the ex-

act value of this threshold was not important for overall star formation rates, but mainly regulated the number and mass of stellar particles produced. The simulations performed in that study employed a non-zero floor (minimum value) of the dust-to-gas ratio in cells, which was meant to account for unresolved pre-enrichment. Since our current simulations spatially resolve regions where the first stars are expected to form, it is unnecessary and inappropriate to use such a floor. This means that in our runs, prior to the metal feedback from the first generation of stars, only primordial chemistry is used in H_2 formation reactions. We find that such primordial reactions with our resolution do not yield molecular fractions $f_{H_2} > 0.01$ on relevant timescales. Therefore, in order to form the first (Pop III) generation of stars, a separate prescription is required and is outlined in the next section.

In cells where the molecular fraction exceeds the f_{H_2} threshold, Pop II stellar particles are formed with a statistical star formation delay, $dt_{SF} = 10^7$ yr, implemented by drawing a random number, P, between 0 and 1, and forming stars only if $P > \exp\left(-\frac{dt}{dt_{SF}}\right)$, where dt is the length of the timestep at the cell's refinement level. Each particle represents a stellar population with a Miller & Scalo (1979) IMF from $0.1\,\mathrm{M}_\odot$ to $100\ \mathrm{M}_\odot$. The mass of a stellar particle is determined by the following the relation:

$$\dot{\rho_*} = \frac{\epsilon_{ff}}{\tau_{SF}} \rho_{H_2}. \tag{2.2}$$

The star formation efficiency per free-fall time is set to $\epsilon_{ff} = 0.01$, based on recent results of Krumholz et al. (2012). We use a constant star formation timescale $\tau_{SF} = 8.4 \times 10^6$ yr corresponding to the free-fall time at hydrogen number density $n_H = 50\,\mathrm{cm}^{-3}$. This approach differs from Gnedin & Kravtsov (2011), where the timescale was computed by using the physical density of molecular clouds. However, we find that in our simulations the density-dependent timescale instills a strong resolution dependence on the star formation rate. Pop II stellar particles are treated as statistical ensembles of stars for which the appropriate metal yield and fraction of stars to go supernovae is computed by integration of the IMF. The number of SN II explosions is 75 per $10^4\,\mathrm{M}_\odot$ formed, while the amount of SN II metals generated by a stellar particle is 1.1% of its initial mass. Each supernova releases 2×10^{51} erg thermal energy which is deposited over the course of 10^7 yr. Following the notation of

Hummels & Bryan (2012), this implies fraction of the rest mass energy of stars which is available for thermal SNe feedback is $E_{SN}/Mc^2 = 8.4 \times 10^{-6}$. This value is relatively high, but consistent with values chosen by other researchers (Hummels & Bryan, 2012).

2.2.2 Population III star formation

Formation criteria

We model the formation of Pop III stars based on criteria derived from simulations of Abel et al. (2002). These authors showed that once the core density of a proto-cloud reached 1000 cm^{-3}, further collapse to a massive stellar object was imminent. Analyzing their results, we found that for gas at any given density n_H past this threshold, the time of collapse to a stellar core is approximately six times the free-fall time for that density, $6t_{ff}(n_H)$. This collapse time is 9 Myr for $n_H = 1000\,\mathrm{cm}^{-3}$ and scales as $n_H^{-1/2}$. For our fiducial runs, we use $n_{H,min} = 10000\,\mathrm{cm}^{-3}$ as the threshold and $dt_{SF} = 2.8$ Myr as the statistical star formation delay, simulating the collapse time. This value is lower than the one used for Pop II stars. This density threshold value was chosen to ensure Pop III stars would form primarily when cells have been maximally refined, but is low enough such that the collapsing gas clouds are still fully resolved in our simulations. Further discussion is presented in Section 2.2.3.

We also set a threshold for the minimum fraction of molecular hydrogen at 10^{-3} to reflect that primordial gas clouds must cool primarily via ro-vibrational transitions of H_2 to form the first stars (Couchman & Rees, 1986; Tegmark et al., 1997). The precise value of this threshold is rather arbitrary, as we do not attempt to model the actual chemistry of stellar core formation. We have chosen this value because it is lower than, but close to the typical value for the molecular hydrogen fraction in cold, dense primordial gas around $z = 20$, which we have found empirically to be 2×10^{-3} (see Section 2.2.3 and Figure 2.3). A minimum threshold for molecular fraction ensures that the H_2-dissociating Lyman Werner radiation from recently-formed Pop III stars will realistically suppress further Pop III star formation in the region.

Pop III stars form in gas that has metallicity $\log_{10} Z/Z_\odot < -3.5$. This threshold is chosen to match the critical metallicity discovered by Bromm et al. (2001a), and has held up in later studies (Smith et al., 2009). Though the exact value of this critical metallicity

is still uncertain and can be affected by the presence of dust (Omukai et al., 2005), we find that it is not very important as the majority of Pop III stars form in truly primordial, or nearly primordial gas. Compiling all of our simulations, we found that only ~10% of Pop III stars form with $\log_{10} Z/Z_{\odot} > -5$.

We summarize the formation criteria for Pop III stars with the following set of equations,

$$\begin{aligned} n_H &> n_{H,min} = 10^4\ \mathrm{cm}^{-3} \\ f_{H_2} &> f_{H_2,min} = 10^{-3} \\ \log_{10} Z/Z_{\odot} &< -3.5. \end{aligned} \tag{2.3}$$

Values given for each variable represent the fiducial choices.

IMF and supernova feedback

The IMF of Pop III stars is currently a hotly debated and active area of research. It is still unclear whether the high Jeans mass of primordial gas results in a top-heavy IMF as predicted by early studies (Abel et al., 2002; Bromm et al., 1999; Yoshida et al., 2003), or if the angular momentum and radiative effects during infall can fragment the cloud and generate relatively low-mass cores (Greif et al., 2011; Stacy et al., 2012; Hosokawa et al., 2011; Clark et al., 2011). It is even likely that the Pop III IMF can be considerably variable depending on environment (O'Shea & Norman, 2007) and ionization state of the collapsing gas (Yoshida et al., 2007). We choose not to explore various analytic forms for the IMF, as constraining it is beyond the scope of this paper. Instead, we consider that the main way by which the Pop III IMF can influence galaxy formation, in contrast to the known Pop II IMF, is by enhancing the output of ionizing radiation and the number and intensity of supernovae. In particular, PISNe, which are hypothetically plausible from stars in the mass range $140 - 260\,\mathrm{M}_{\odot}$ (or for lower masses if rotation is considered, see e.g. Stacy et al. 2013), would potentially be dramatic singular events in the evolution of any galaxy (Bromm et al., 2003; Whalen et al., 2008). To account for the occurrence of PISNe we use two different particle masses for Pop III stars. Every newly formed Pop III stellar

particle is randomly assigned to be either a $170\,M_\odot$ star, which is to explode in a PISN, or $100\,M_\odot$ star, which only generates a mild explosion before collapsing into a black hole (Heger & Woosley, 2002, 2010). The proportion of these two types of particle mass and fate is governed by a single parameter, P_{PISN}, which is the fraction of PISNe progenitors ($170\,M_\odot$ stars) that form when the Pop III star formation criteria are met. In our fiducial runs, we set $P_{PISN} = 0.5$. This value was chosen to test the maximum possible impact of PISNe on galaxy evolution, and probably represents the most top-heavy the primordial IMF can possibly be. Since the atmospheres of Pop III stars are free of metals, they are unable to drive stellar winds and therefore do not enrich the ISM in any way other than through supernovae. Pop III stars which have masses too low to produce SNe are ignored in our model.

For our fiducial value of $n_{H,min}$ (as well as all other parameters considered in Section 2.2.3), we found that the gas mass in a maximally refined cell at $z \approx 20$ is sometimes not sufficient to form a $170\,M_\odot$ star. Therefore, we prevent further refinement in metal-free cells that have $n_H > 0.5 n_{H,min}$ and whose splitting would leave insufficient mass to form the star. Through tests, we have checked that this refinement restriction never artificially slows down Pop III star formation. It becomes especially relevant in the super-Lagrangian runs discussed in Section 2.2.3 and in the H and Q Mpc boxes which inherently have very high resolution.

The PISN from a $170\,M_\odot$ star releases 27×10^{51} erg of thermal energy, as well as $80\,M_\odot$ of metals into the ISM (Heger & Woosley, 2002). As suggested by Wise & Abel (2008), we use a delay of 2.3 Myr from the formation of a $170\,M_\odot$ particle to its PISN event, representing the main sequence lifetime of this type of star (Schaerer, 2002). After the supernova goes off, the cell which hosts it often winds up with super-solar metallicity. The cooling functions employed in our code are not accurate for these high-temperature high-metallicity conditions associated with the early phases of supernova remnants. We found that while the blastwave expanded regardless of whether or not cooling was turned on, the inner regions of the supernova remnant overcooled. We therefore turned off all metal cooling for gas at temperatures higher than 10^4K. According to the models of Heger & Woosley (2002), a $170\,M_\odot$ star is completely disrupted by its PISN event, and all gas

mass from the stellar interior would be ejected into the ISM leaving no remnant. The ejecta then consists of $80\,M_\odot$ of metals from the core, as well as the primordial envelope which is $22\,M_\odot$ of He and $68\,M_\odot$ of H.

A $100\,M_\odot$ star does not explode as a PISN, but its actual fate is still uncertain and depends on the details of the stellar rotation and magnetic structure (Heger & Woosley, 2010). Before undergoing core collapse, such a star would experience thermonuclear pair-instability pulses that would eject the outer layers of H and He, with possible traces of the elements C, N, and O. The energy released in such pulses is of the order or smaller than the energy of a normal SN type II. If the remaining core has enough rotation to trigger a gamma-ray burst in the collapsar model (Woosley & Heger, 2012), it may then lead to a powerful explosion with over 10^{52} erg of energy and the ejection of significant mass of metals. Without rotation, the core may drive a weak collapsar explosion or no explosion at all, when all the remaining mass recollapses. Given these uncertainties, and to contrast with the case of a full PISN explosion for the $170\,M_\odot$ stars, for the $100\,M_\odot$ stars we assume that no substantial metals are deposited into the ISM and that the released energy corresponds to a standard SN type II. About $50\,M_\odot$ of gas is released into the ISM, while also leaving behind a remnant black hole of ~$50\,M_\odot$.

To prevent artificial radiative losses, PISN energy and mass ejecta are distributed within a sphere of constant density, with a radius 1.5 cell lengths, centered at the middle of the PISN host cell. Each of the 27 cells within such a sphere, consisting of the star's host cell and its immediate neighbors, receive a dose of energy and metal-rich gas proportional to the actual volume of the cell contained within the sphere. This prescription is physically motivated as we found that our typical timestep (about 450 yr) is too coarse compared to the typical timescale of the early free expansion phase of the SN remnant. For example, it takes the ejecta ~500 yr to traverse half of the typical Pop III star host cell length (4.5 pc) at z~20 if it travels at the free-expansion velocity. This velocity is computed here by assuming that all of the PISN energy of 27×10^{51} erg goes into kinetic energy of the ejecta. In practice, we found that this model did not significantly affect the geometry of the blastwave relative to simulations where we injected the metals and energy into a single cell.

Radiative feedback

In addition to the supernova feedback, all Pop III stars have enhanced radiative feedback relative to Pop II counterparts, due to the lack of metals in their atmospheres (Schaerer, 2002). We use the same spectral shape for the ionizing feedback of all stellar particles (the Pop II SED from Figure 4 of Ricotti et al. 2002a), which has a characteristic energy of 21.5 eV for ionizing photons, however we enhance the radiative output of Pop III stars by a factor of 10 relative to Pop II, following Wise & Cen (2009). After a Pop III stellar particle undergoes supernova, radiative feedback from the star is completely shut off. On the other hand, Pop II stellar particles shine according to a light curve fit from Starburst 99 model results (Leitherer et al., 1999). This light curve consists of a flat component for the first 3×10^6 yr, followed by a steep power-law falloff. Radiative feedback from Pop II stellar particles becomes insignificant after 3×10^7 yr. Since Pop II stars shine longer than both types of Pop III stars, the factor of 10 radiative enhancement does not translate into a proportionally higher number of ionizing photons per lifetime. Pop II stars emit 6,600 ionizing photons per stellar baryon per lifetime. In our fiducial runs, Pop III stars emit 38,800 and 34,500 photons per baryon per lifetime for the $100\,\mathrm{M}_\odot$ and $170\,\mathrm{M}_\odot$ stars, respectively.

2.2.3 Convergence Study & Setting Fiducial Parameters

In this section, we describe the test runs that justify the numerical setup and the choice of parameters for our main runs. Since the Pop III star formation recipe described above is one of the critical components of our study, we focus on testing the key elements of this model. In Table 1 we list the details of the simulations performed in our suite. Box OverDense has many more potential sites for Pop III star formation than the other 1 h^{-1} Mpc boxes, and therefore serves as the best testing ground. It is important to keep in mind that while every parameter we test has an effect on Pop III star formation, the most drastic differences between the simulations are caused by the choice of initial conditions. The role of cosmic variance will be explored more comprehensively in Chapter 3.

Table 2.1: SIMULATION RUNS OF CHAPTER 2

Run Description	Base grid	ℓ_{max}	dx (pc)	$m_{DM}(\mathrm{M}_\odot)$	$n_{H,min}(\mathrm{cm}^{-3})$
Convergence study					
UnderDense−_noSF_7 No star formation	256^3	7	44	5500	-
UnderDense−_noSF_8 No star formation	256^3	8	22	5500	-
UnderDense−_noSF_9 No star formation	256^3	9	11	5500	-
UnderDense+_noSF_7 No star formation	256^3	7	44	5500	-
UnderDense+_noSF_8 No star formation	256^3	8	22	5500	-
UnderDense+_noSF_9 No star formation	256^3	9	11	5500	-
OverDense_noSF_7 No star formation	256^3	7	44	5500	-

OverDense_noSF_8	256^3	8	22	5500	-
No star formation					
OverDense_noSF_9	256^3	9	11	5500	-
No star formation					
OverDense_noSF_HMpc	256^3	8	11	690	-
No star formation, 0.5 h^{-1} Mpc box					
Fiducial runs					
UnderDense−_nH1e4_fid	256^3	8	22	5500	10000
Underdense box with no massive galaxies, fiducial parameters					
UnderDense+_nH1e4_fid	256^3	8	22	5500	10000
Underdense box with one massive galaxy, fiducial parameters					
OverDense_nH1e4_fid	256^3	8	22	5500	10000
Overdense box, fiducial parameters					
Density threshold study					
OverDense_nH1e3	256^3	8	22	5500	1000
Lowest density threshold for Pop III star formation					
OverDense_nH5e3	256^3	8	22	5500	5000
Low density threshold for Pop III star formation					

OverDense_nH2e4	256^3	8	22	5500	20000
High density threshold for Pop III star formation					
Mass resolution & refinement criteria					
OverDense_SL7	256^3	8	22	5500	10000
Super-Lagrangian refinement 0.7^ℓ					
OverDense_SL5	256^3	8	22	5500	10000
Aggressive super-Lagrangian refinement 0.5^ℓ					
OverDense_HiRes	512^3	7	22	690	10000
Higher mass resolution					
OverDense_HMpc	256^3	7	22	690	10000
0.5 h^{-1} Mpc box					
OverDense_HMpc_HiRes	256^3	8	11	690	10000
0.5 h^{-1} Mpc box, higher spatial resolution					
OverDense_HMpc_SL5	256^3	7	22	690	10000
0.5 h^{-1} Mpc box, super-Lagrangian refinement 0.5^ℓ					
OverDense_QMpc	256^3	8	5.5	86	10000
0.25 h^{-1} Mpc box					
Alternative physics					
OverDense_ExtremeSN	256^3	8	22	5500	10000
Extreme PISNe (Section 2.2.3)					

OverDense_ExtremeRad	256^3	8	22	5500	10000
Extreme Pop III radiation field (Section 2.2.3)					
OverDense_LowMass	256^3	8	22	5500	10000
Pop III IMF and feedback mirror Pop II (Section 2.2.3)					

Column 1.) Name of the run;

2.) Base grid, number of DM particles, number of root cells;

3.) Maximum number of additional levels of refinement;

4.) Minimum cell size at the highest level of refinement in comoving pc;

5.) DM particle mass in $M_\odot$;

6.) Minimum H number density for Pop III star formation in cm^{-3};

7.) Further description of the run. sideways

Density threshold for Pop III star formation

First, we choose an appropriate value for the density threshold for creating Pop III stars. In Figure 2.1, we examine the high-density end of the volumetric probability distribution function (PDF) of the hydrogen number density at $a = 0.085$ ($z = 10.8$). In order to test the properties of the primordial gas from which the first stars form, we ran a special set of simulations with no star formation or chemical enrichment (runs OverDense_noSF_8, UnderDense+_noSF_8, UnderDense−_noSF_8, as well as additional versions of each with one more and one fewer maximum level of spatial refinement). Though the total mass of gas in each box is the same, only gas in the most massive halos has collapsed to this density regime, meaning that the PDFs are very sensitive to the number and nature of such halos. The PDFs of Box OverDense and UnderDense+ are offset by a factor of ~5 at all densities, while the UnderDense− box is offset from Box UnderDense+ by another factor of ~5. This difference is also seen in the maximum density achieved in each box. For our fiducial resolution of 8+8 levels, all three boxes are able to reach a density of at least 10000 cm^{-3} by $a = 0.085$, giving us enough time to study star formation in every box before our stopping point of $a = 0.1$.

Based on these results, we chose $n_{H,min} = 10000\,\mathrm{cm}^{-3}$ as our fiducial value for the density threshold. In addition to the constraints obtained from the PDF, other considerations went into this selection. A lower value would suffice to meet our proto-cloud collapse criteria, but would result in all Pop III stellar particles forming before cells are maximally refined. Such an outcome is poor practice in hydrodynamic simulations, as subgrid physics is being invoked on scales where the resolution is still good enough to self-consistently capture relevant physical processes. On the other hand, using a higher threshold would allow the maximally refined cells to reach densities beyond the resolving power of the simulation. When such conditions are reached, either further refinement or subgrid physics should already be in use. In addition, using a higher density threshold in our test runs often led to Pop III stars forming in bursts (in the same timestep, in neighboring cells). We do not speculate here whether such bursts are physically plausible or not, but the scales necessary to model this process properly are certainly unresolved in our simulations. We suspect

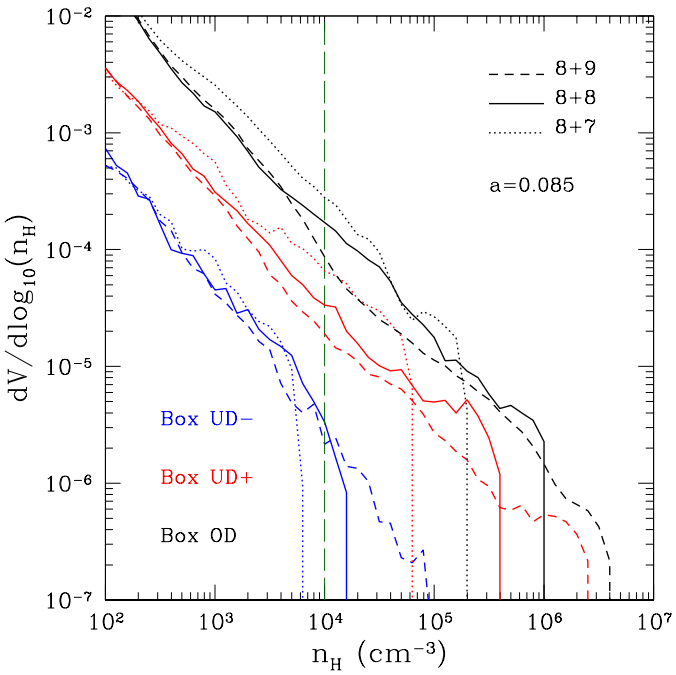

Figure 2.1. The distribution function of hydrogen number density for runs without star formation at $a = 0.085$ ($z = 10.8$). Blue lines are for Box UnderDense–, red for Box UnderDense+, black for Box OverDense. Dotted lines are for 8+7 levels of refinement, solid lines for 8+8, short-dashed lines for 8+9. The long-dashed green line represents our chosen density threshold $n_{H,min} = 10000\,\mathrm{cm}^{-3}$. All runs with at least 8 levels of refinement have sufficiently dense gas to form stars by this epoch.

that higher temporal or spatial resolution would reveal that feedback from the first star in a cell would suppress, or at least delay further clustered star formation, as the H_2 photo-dissociation timescales due to internal Lyman-Werner feedback from a single $100\,\mathrm{M}_\odot$ star within a given star-forming clump are typically shorter than the clump's free-fall timescale (Safranek-Shrader et al., 2012).

To determine the ultimate effect of the density threshold on Pop III star formation, additional runs were performed with $n_{H,min} = 1000$, 5000, and $20000\,\mathrm{cm}^{-3}$ using the Box OverDense initial conditions. Varying this threshold by a factor of 20 changes the scale factor at which the first star forms only from $a = 0.0463$ to $a = 0.0483$, or from redshift $z = 20.6$ to $z = 19.7$. In the $n_{H,min} = 1000\,\mathrm{cm}^{-3}$ run, the density threshold is reached at a lower level of refinement from the other cases considered, allowing the first star to form sooner. The variation in the other three runs is only from $a = 0.0480$ to $a = 0.0483$. Such marginal differences demonstrate that for a given set of initial conditions, our actual density threshold criterion for the formation of Pop III stars has little effect on when and where they

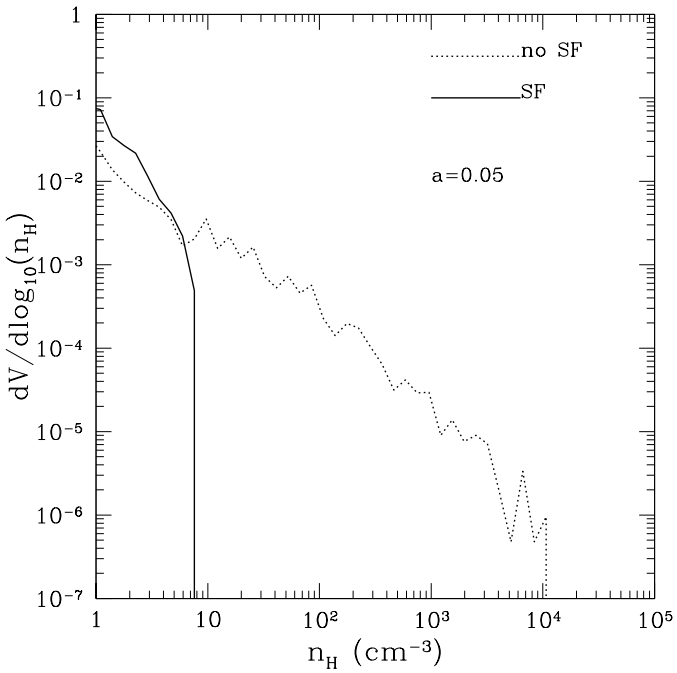

Figure 2.2. PDF for most massive galaxy in Box OverDense at $a = 0.05$ ($z = 19$) with star formation (solid, run OverDense_nH1e4_fid) and without star formation (dotted, run OverDense_noSF_8). The Pop III star that recently formed at the center of this galaxy has temporarily depleted it of the dense gas needed to continue star formation.

form. In each of these runs, only a single Pop III star formed in each box before $a = 0.05$, and the total number of Pop III stars varied between 4 and 5 at $a = 0.055$. Based on these tests, we have determined that the total number of Pop III stars formed had little correlation with the density threshold within the considered range.

After the first star forms in a given halo, the gas density can be significantly reduced near the center, quenching further star formation. Figure 2.2 demonstrates this effect in run OverDense_nH1e4_fid. The PDF of this galaxy is depleted at high density 10 Myr after a PISN explosion. The corresponding galaxy from the run without star formation, OverDense_noSF_8, is also shown for reference. While the galaxy in OverDense_noSF_8 contains some dense gas above $n_H = 10\,\mathrm{cm}^{-3}$, it is depleted in our fiducial run. Since this density is nowhere near any $n_{H,min}$ that we have considered in our tests, we can conclude that Pop III stars will not form in quick succession in this halo.

Molecular Fraction

Another component of the Pop III star formation criterion is the requirement of a minimal fraction of H_2 in the host cell. To determine what value of the H_2 threshold makes sense in the context of these simulations, we examine the molecular fraction of hydrogen as a function of density in the runs without star formation at the epoch ($z \approx 20$) when gas is beginning to reach densities close to $n_{H,min}$. Figure 2.3 shows that the molecular fraction in primordial gas generally increases with density, but saturates above $n_H \approx 10\,\mathrm{cm}^{-3}$. The saturation value of f_{H_2} grows slowly with time in the absence of star formation, and does not appear to depend significantly on spatial or mass resolution. Our fiducial choice of 10^{-3} for the minimal H_2 fraction does not exclude dense gas from forming stars in any runs, as long as little Lyman-Werner radiation is present.

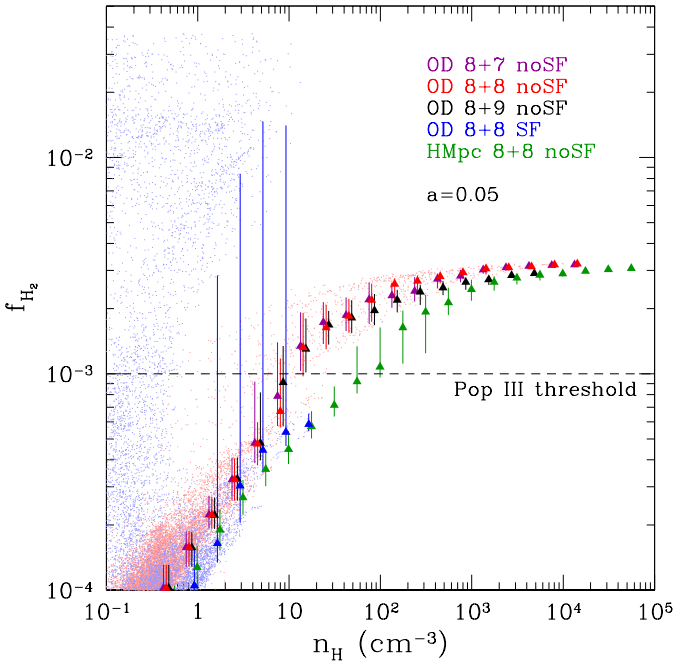

Figure 2.3. Molecular fraction of hydrogen vs. number density for three Box OverDense runs using different resolution without star formation at $a = 0.05$ ($z = 19$, runs OverDense_noSF_9, OverDense_noSF_8, and OverDense_noSF_7 are black triangles, red circles, and purple squares, respectively). The median H_2 fraction in each density bin is indicated by a triangle, while the error bars show 25th and 75th percentile levels. At this epoch, when the first stars would normally be forming, our fiducial resolution of 8 levels has the same molecular fraction as if we were using one more or one fewer level of refinement, and the points actually lie directly on top of each other for $n_H < 10\,\mathrm{cm}^{-3}$. Run OverDense_noSF_HMpc at $a = 0.05$ (green filled triangles) has lower values and wider spread of H_2 fraction for $n_H < 10^2\,\mathrm{cm}^{-3}$ but converges with the other runs at higher densities, demonstrating a lack of dependence on mass resolution. Also shown is run OverDense_nH1e4_fid where star formation has already taken place by $a = 0.05$ (blue filled circles). Since the gas has been enriched by a PISN, molecular gas can form at much lower densities. Light red and light blue points trace out the H_2 fraction in every single cell for run OverDense_noSF_8 and OverDense_nH1e4_fid, respectively.

Super-Lagrangian Refinement

Since we have demonstrated resolution dependence for the maximum density of gas within a given galaxy, it is expected that refinement criteria could play a role in controlling when gas in the simulation first reaches the $n_{H,min}$ threshold. To test this, in some of our runs we employ super-Lagrangian (SL) refinement criteria. This approach dictates that the refinement threshold between subsequent levels is lowered by a constant factor, granting a more effective zoom-in on the densest regions at earlier times. The refinement criteria in a cell can be written as $2 \times m_{DM} \times \frac{\Omega_m}{\Omega_{DM}} \times X^{\ell} \times 0.8$ for the dark matter mass, and $0.3 \times m_{DM} \times \frac{\Omega_m}{\Omega_{DM}} \times X^{\ell} \times 0.8$ for the gas mass, where ℓ is the level of the cell which is to be refined. In this formalism $X = 1$ implies Lagrangian refinement as described at the beginning of Section 2.2. We have tried runs with very aggressive SL refinement ($X = 0.5$, run OverDense_SL5) and less aggressive refinement ($X = 0.7$, run OverDense_SL7). Running these simulations in Box OverDense with $n_{H,min} = 10000\,\mathrm{cm}^{-3}$, we found that the epoch at which Pop III stars first appear is pushed back from $a = 0.0478$ to $a = 0.0456$ with $X = 0.7$, and to $a = 0.0435$ with $X = 0.5$. This demonstrates that the use of SL refinement is an important numerical tool for exploring the earliest epoch of star formation in a given simulation box. However, using SL refinement produces an enormous number of high-level cells: at $a = 0.05$, run OverDense_SL5 has a factor of 4000 more maximally refined cells than run OverDense_nH1e4_fid. This drastic difference makes the SL simulations prohibitively expensive soon after the first stars form.

Therefore, we use these SL runs to study Pop III star formation at the earliest possible epochs, when the mass of the halos that hosted them was low enough for PISNe to have their maximal effect.

Increased Mass Resolution

We test the effects of mass resolution by setting up one run with 512^3 initial grid, giving a DM particle mass of $690\,\mathrm{M}_{\odot}$. We use 7 additional levels of refinement, therefore granting us the same maximum spatial resolution as in the fiducial 256^3 run. Having consistency in spatial resolution allows us to test the effects of mass resolution alone. All other numerical parameters are kept consistent with run OverDense_nH1e4_fid.

The increased mass resolution has several immediate implications. Since we now resolve halos of mass $10^6\,M_\odot$ with over 1000 particles, we can better probe the regime where the very first Pop III stars are expected to collapse in proto-galactic 'minihalos'. Indeed, the epoch of formation of the first star in the box becomes $a = 0.0427$ in a halo of $1.5 \times 10^6\,M_\odot$ (compared to $a = 0.0481$ and $M_h = 7.5 \times 10^6\,M_\odot$ in run OverDense_nH1e4_fid). The higher mass resolution effectively means that there is more power on the small scales responsible for the growth of halos in this mass regime. Due to the high computational cost of this run, we have only advanced it to $a = 0.055$.

The effect of increasing mass resolution is further explored through runs using the H and Q boxes of 0.5 h^{-1} Mpc and 0.25 h^{-1} Mpc in size. Applying the 256^3 base grid to these boxes gives us a DM particle mass of $690\,M_\odot$ and $86\,M_\odot$, respectively. We find that the H box (run OverDense_HMpc_HiRes) produces a Pop III star by $a = 0.0456$ in a halo of mass $1.5 \times 10^6\,M_\odot$, while the Q Mpc box (run OverDense_QMpc) does not produce one until $a = 0.0473$. Using SL refinement in the HMpc box (run OverDense_HMpc_SL) allows us to see a Pop III star forming in a $8 \times 10^5\,M_\odot$ halo.

The earlier formation epochs and lower mass of halos hosting the first stars in the H and Q Mpc boxes, compared to the fiducial 1 h^{-1} Mpc runs, show that it is crucial to have high enough mass resolution to capture Pop III star formation in halos close to $10^6\,M_\odot$. It has been previously shown that halos less massive than this threshold will not achieve significant enough H_2 abundances to trigger Pop III star formation at earlier epochs (Yoshida et al., 2003). The further significance of this mass range will be explained in our Results section. Figure 2.4 shows explicitly how varying refinement criteria, spatial resolution, and mass resolution affected the lowest possible mass for a star-forming galaxy.

Low Mass Pop III IMF

We present one simulation, run OverDense_LowMass, which does not rely on a top-heavy IMF for Pop III stars. The conditions for Pop III star formation in this run are similar to our fiducial top-heavy recipe in that we use the same threshold $n_{H,min}$ to determine which cells are allowed to form stars. However, the stellar particle masses are drawn from the same IMF as for Pop II stars. This run explores the possibility that Pop III stars were ordinary

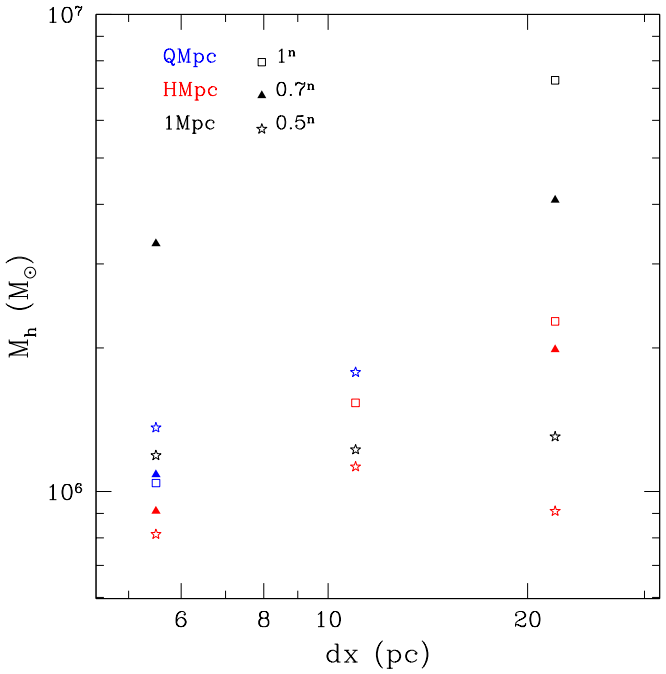

Figure 2.4. The least massive galaxy to host a star in various runs vs. the minimum comoving cell size employed in the run. The colors indicate the simulation box size: 1 h^{-1} Mpc (black), 0.5 h^{-1} Mpc (red), or 0.25 h^{-1} Mpc (blue). The shape of points indicates the refinement criterion employed in the box, with open squares for Lagrangian refinement, filled triangles for 0.7^{ℓ} SL refinement, and open five-pointed stars for 0.5^{ℓ} SL refinement.

low-mass objects. Whenever the density in a given cell exceeds the threshold, the star formation rate is determined according to the following relation:

$$\dot{\rho_*} = \frac{\epsilon_{ff}}{\tau_{SF}} \rho_{gas}, \tag{2.4}$$

where ρ_{gas} is the mass density of all gas in the cell. This relation is similar to equation 2.2, but does not explicitly use molecular hydrogen. This modification is necessary because primordial gas can reach densities above our star formation threshold, but cannot become fully molecular without the presence of dust.

Extreme Pop III Feedback

To isolate the relative impacts of the feedback effects, we ran toy simulations using exaggerated values for the PISN energy and ionizing photon yield. In one run, called OverDense_ExtremeSN, PISNe released 270×10^{51} erg of thermal energy, a factor of 10 larger than in all other runs. The extreme ionizing simulation OverDense_ExtremeRad had in-

stead an additional factor of 10 boost in the ionizing photon flux of Pop III stars, giving the $100\,M_\odot$ and $170\,M_\odot$ stars 388,000 and 345,000 photons per lifetime, respectively. While these models are too strong to be consistent with any published results, using them allows us to explore the most extreme effects of Pop III feedback.

2.3 Results

Pop III stars in our simulations begin to form in halos of mass $M_h \gtrsim 10^6\,M_\odot$ starting at $a \approx 0.045$ ($z \approx 21.2$) in accordance with expectations from prior work (Yoshida et al., 2003). Figure 2.5 shows the mass of host halo in which each Pop III star formed. Pop II stars begin forming in most of these halos shortly thereafter, but are not shown in this plot. In the 1 h^{-1} Mpc runs, the halo mass for first star formation is close to $10^7\,M_\odot$. This mass is an order of magnitude larger than that of the halos hosting the first stars in the simulations of Wise et al. (2012a) and Greif et al. (2011), and those preferred by theoretical considerations (Tegmark et al., 1997). Consequently, those authors also find an earlier epoch for the formation of the first stars. Given that the extra mass resolution granted by the H Mpc box allows us to see star formation in $10^6\,M_\odot$ halos, we infer that our fiducial 1 h^{-1} Mpc runs are not properly resolving the very first star-forming minihalos. Rather, they are more generally simulating Pop III star formation in an early population of galaxies. The fraction of star-forming halos in run OverDense_nH1e4_fid at $a = 0.07$ ($z = 13.3$) was only 1% in the mass range $10^6\,M_\odot < M_h < 10^7\,M_\odot$, but it reached 65% for $M_h > 10^7\,M_\odot$. In run OverDense_HMpc_HiRes, these numbers increase significantly to 15% for $10^6\,M_\odot < M_h < 10^7\,M_\odot$ and 100% for $M_h > 10^7\,M_\odot$. In addition to the resolution effects, the suppression of star formation in the 10^6 to 10^7 $M_\odot$ range is also plausible in the regime of a moderate Lyman-Werner background (e.g. Machacek et al. 2001; O'Shea & Norman 2008; Safranek-Shrader et al. 2012).

It is worth noting that the ratio of star-forming galaxies in the range $10^7\,M_\odot < M_h < 10^8\,M_\odot$ falls off gradually with time in run OverDense_nH1e4_fid. It changes from 65% at $a = 0.065$ to 20% at $a = 0.1$, suggesting that halo mass alone is not a good proxy for determining whether a galaxy can achieve the high density required for our Pop III star formation criteria. One potential cause of the change is the decreased physical spatial

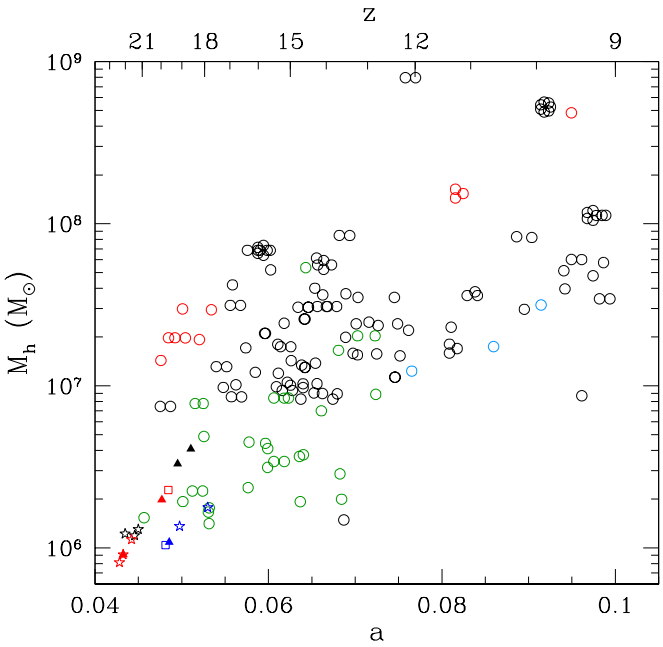

Figure 2.5. Each Pop III star's host halo mass at the time of formation vs. the scale factor at which the star formed for run UnderDense−_nH1e4_fid (blue), run UnderDense+_nH1e4_fid (red), run OverDense_nH1e4_fid (black), and run OverDense_HMpc_HiRes, which did not go past $a = 0.075$ (green). Additional points for the halos hosting the first stars from each of the runs used for Figure 2.4 are also included, with the same color and shape scheme. In our 1 h^{-1} Mpc runs, Pop III star formation happens almost exclusively in halos between $10^7\,M_\odot$ and $10^8\,M_\odot$. The additional mass resolution in run OverDense_HMpc_HiRes makes it possible to see that the first Pop III stars form in halos between $10^6\,M_\odot$ and $10^7\,M_\odot$. The average mass of Pop III star-forming halos increases slightly with time. When multiple Pop III stars form within a galaxy in a very short time interval, points on the plot are grouped into a clustered shape.

resolution at later epochs, but according to our study of the gas in the first star-forming galaxy shown in Figure 2.3, the factor-of-two difference in spatial resolution achieved by using one fewer level of refinement does not preclude gas from reaching the fiducial $n_{H,min} = 10^4\,\mathrm{cm}^{-3}$ threshold. The difference is more likely to be rooted in the evolution of physical density in halos of a given mass. For halos between 10^7 and $10^8\,M_\odot$, the average matter density within the virial radius changes from $2.5 \times 10^{-2}\,M_\odot\,\mathrm{pc}^{-3}$ at $a = 0.065$ to $6.8 \times 10^{-3}\,M_\odot\,\mathrm{pc}^{-3}$ at $a = 0.1$, due to the expansion of the universe. Even the density within the central 100 pc of these halos changes from $0.73\,M_\odot\,\mathrm{pc}^{-3}$ to $0.34\,M_\odot\,\mathrm{pc}^{-3}$ between the same two epochs.

Very few Pop III stars formed in halos with $M_h > 10^8\,M_\odot$, because such halos have already been enriched to metallicities above $\log_{10} Z/Z_\odot = -3.5$, allowing for normal star formation to commence. In many cases, halos with $M_h > 3 \times 10^7\,M_\odot$ had earlier formed

one or more $100\,M_\odot$ Pop III stars, which shut off star formation temporarily but did not enrich the galactic gas, allowing it to remain pristine and continue forming Pop III stars. Another major exception occurs in run OverDense_nH1e4_fid in a galaxy that has already formed a significant number of Pop II stars that have in turn enriched the ISM terminating further Pop III star formation. However at $a = 0.095$, this galaxy undergoes a major merger with another massive halo, causing low-metallicity gas in the outer part of the galaxy to collapse, thereby triggering a burst of Pop III star formation which appears as a cluster of points with $M_h = 5 \times 10^8$ on Figure 2.5. All of these stars form with metallicities around the critical $\log_{10} Z/Z_\odot = -3.5$ value, suggesting that their existence is sensitive to the value of this threshold and therefore should not be treated as a general result.

2.3.1 Effect of Pop III stars on their host galaxies

The strong ionizing flux of Pop III stars and enormous energy injections from PISNe have been shown in previous work to significantly alter the ISM of their host galaxies. Here we explore such effects during time when Pop III stars are the dominant drivers of feedback.

In Figure 2.6 we show that there is a significant variation in the potential effect of Pop III stars on their host galaxies depending on the galaxy mass. In halos with $M_h < 3 \times 10^6\,M_\odot$, Pop III stars can temporarily evacuate the gas from the galaxy, and the metals from PISNe are ejected past the virial radius into the intergalactic medium (IGM). On the other hand in halos with $M_h > 3 \times 10^6\,M_\odot$, the metals are confined within the virial radius, and there is little movement of baryons beyond the virial radius.

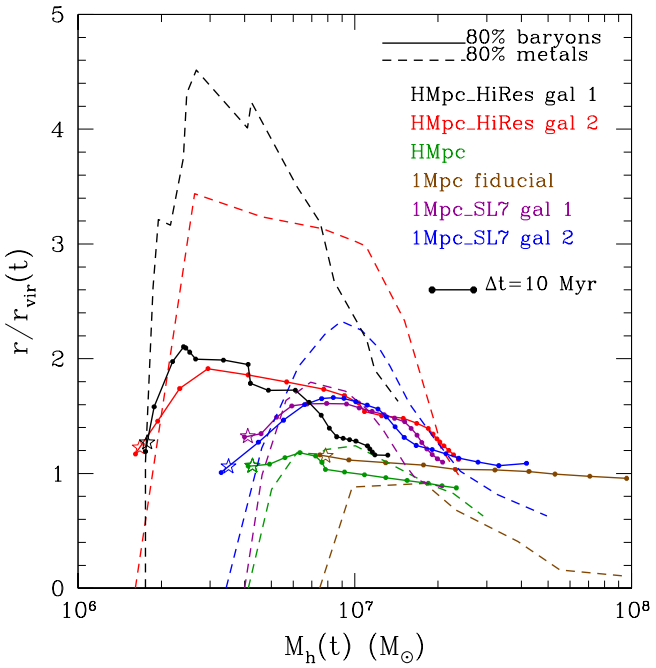

Figure 2.6. Though metals from PISNe are ejected past the virial radius, they do not stay there for long. Plotted here are the radii enclosing 80% of the metals produced in the galaxy, as well as the radii where the enclosed mass of baryons divided by the virial mass equals 80% of the universal baryon fraction. Each line represents a galaxy as it evolves in time, beginning at the epoch when the first star forms. Within 50 Myr of the PISN event (denoted by five-pointed stars), most of the gas and metals have begun to recollapse, or are at least enclosed within the virial radius once again. The maximum extent of metal propagation is strongly regulated by galaxy mass. The distance between two points on each line corresponds to 10 Myr.

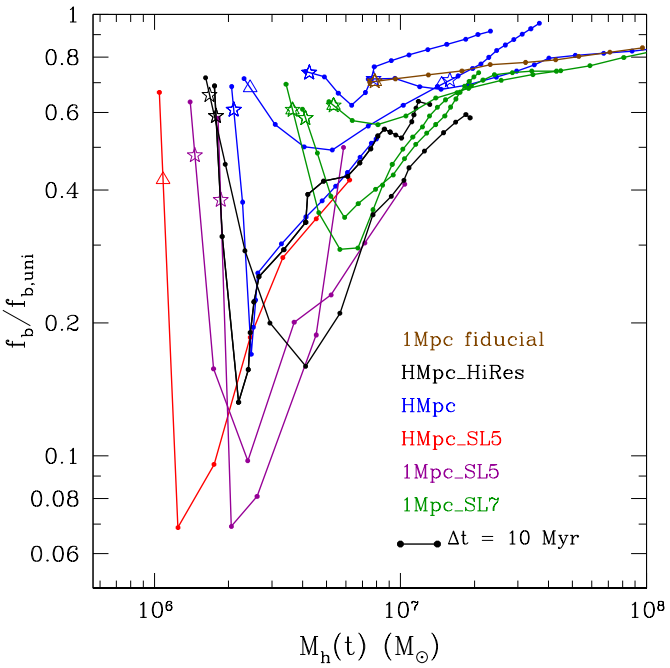

Figure 2.7. The evolution of the baryon fraction vs. the evolving halo mass. Each line traces a galaxy from the time of formation of the first star. The baryon fraction is computed within the virial radius, and is normalized by the universal value $f_{b,uni}$. While supernovae initially cause a depletion of baryons in galaxies of $M_h < 3 \times 10^6\,M_\odot$, this depletion is only temporary. In galaxies of $M_h > 3 \times 10^6\,M_\odot$, there is no strong evidence that PISNe are able to deplete baryon fractions. Notation is the same as in Figure 2.6, with the addition of triangles to represent SN II produced by $100\,M_\odot$ Pop III stars.

This divide can also be seen in Figure 2.7, which shows the baryon fraction computed within the virial radius for thirteen halos taken from a variety of runs. Again, a noticeable threshold at $M_h = 3 \times 10^6\,M_\odot$ distinguishes galaxies that have their gas evacuated by PISNe from those that do not. Galaxies less massive than this threshold typically have their gas content plummet by at least a factor of two within 10-30 Myr after the PISN, with the least massive ones falling below 10% of the universal baryon fraction. In contrast, more massive galaxies lose a much smaller percentage of their gas and end up with baryon fractions in excess of their pre-explosion values within ~100 Myr. This dividing line between "low-mass" and "high-mass" galaxies is therefore a logical choice for parameter that distinguishes different regimes of Pop III feedback. We examine these two regimes separately below.

Dependence on halo mass

To further understand how halo mass can determine the effectiveness of Pop III stellar feedback, we examine the structural evolution of several galaxies in different mass regimes.

First, we examine a relatively low-mass galaxy from run OverDense_HMpc_HiRes, which is shown by the black line that extends to $M \approx 10^7\,\mathrm{M}_\odot$ in Figures 2.6 and 2.7. The first star (of $170\,\mathrm{M}_\odot$) forms when the mass of the halo is $1.7 \times 10^6\,\mathrm{M}_\odot$. Within 20 Myr of its formation, the PISN has blasted metals out beyond 1 kpc from the galactic center (or $4R_{vir}$ at this epoch), and the baryon fraction has dipped to $f_b / f_{b,uni} \approx 0.15$. However, soon after this point the baryon fraction begins to grow again, and the virial radius increases enough to enclose a larger fraction of the expelled gas and metals.

At t = 139 Myr after the first PISN (the halo mass is now $9 \times 10^6\,\mathrm{M}_\odot$), the galaxy has regained ~37% of the PISNe metals. The baryon fraction is over half of the universal value. It will still take more time for this galaxy to completely recover from the explosion, but there is considerable evidence from Figure 2.6 that metals and baryons in general are flowing in rather than out of the galaxy. Another sign of recovery is that Pop II star formation has commenced within the galaxy, as it now contains 4 Pop II stellar particles (which still contribute little to the metal budget).

Figure 2.8 follows the radial distribution of metals in this galaxy from the time of the first PISN until 139 Myr after it has exploded. The Pop II stars have contributed less than 1 $\mathrm{M}_\odot$ to the metal budget, so essentially all of the metals shown here are products of the first PISN, and of a second PISN which happens 15 Myr after the first in a neighboring halo at a distance of approximately 2 kpc. The mass of this galaxy has increased by a factor of ~5 between a = 0.0508 when the star first formed and a = 0.0726 at the final snapshot considered.

While the PISNe do clearly cause baryon depletion and suppress star formation in galaxies such as the one presented here, the rate of growth of these galaxies is high enough that a mixture of ejecta and fresh primordial gas fall in to restore the baryon fraction to at least 50% of the universal fraction within ~150 Myr. This replenishment results from a combination of actual re-accretion of ejected material, accretion of new primordial baryons

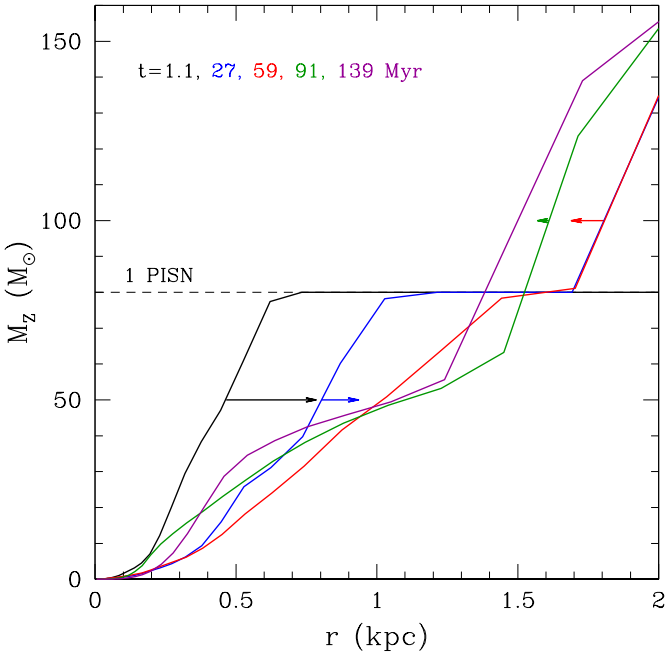

Figure 2.8. The ejecta of PISNe is traced via examining the enclosed mass of metals as a function of galactocentric distance. Lines of different color correspond to 1, 27, 59, 91, and 139 Myr after the first PISN. A second PISN happens 15 Myr after the first in a nearby galaxy. Approximately 60 Myr after the first PISN, more metals are flowing into the galaxy than outwards, as the metal-rich ejecta have mixed with primordial gas accreting onto the galaxy. Arrows show the direction of metal movement at each epoch. The length of each arrow corresponds to the distance traversed by the metals in a 20 Myr interval. The y-axis positions of the arrows show the mass of metals at each epoch used to compute the rate of propagation. This galaxy is depicted by the black line in Figures 2.6 and 2.7.

from filaments, and rapid growth of galaxy virial radius ("gobbling up" of ejecta).

The eventual fate of this low-mass galaxy, and of many such minihalos which were significantly affected by the first PISNe, is to merge with a more massive companion prior to the completion of the metal re-accretion process. The resultant galaxy will ultimately have a baryon fraction close to the universal value, and will contain enough of the PISN ejecta from the progenitors to form Pop II stars. In some cases, we observed that halos in this mass range sustained more long-term damage from PISN, and their baryon fraction stayed below 50% by the end of our simulation, as late as 200 Myr after the explosion. This scenario played out in relatively isolated environments with slow filamentary accretion. Galaxies that underwent such long-term disruption by PISNe had their virial mass increase at an average rate of $0.04\,M_\odot$ yr^{-1} for 100 Myr after the explosion, while all other galaxies that hosted Pop III stars grew at rates ranging from $0.04\,M_\odot$ yr^{-1} to $0.8\,M_\odot$ yr^{-1}.

In run OverDense_HMpc_HiRes, which effectively resolved galaxies in the minihalo regime, 21% of PISNe occurred in underdense environments where the metals were permanently ejected from the host galaxy. Another 25% of PISNe happened in galaxies where the host merged with a separate galaxy prior to the complete gobbling of metals. The remaining 54% of PISNe happened in galaxies where metals were effectively gobbled up by the end of the simulation. These findings suggest that 20-45% of the metals from Pop III supernova ejecta can be observed in the IGM at $z \approx 10$.

Next, we study a galaxy from run OverDense_nH1e4_fid that was the first to form a Pop III star, which also happens to be a PISN progenitor. This galaxy is represented by the brown lines in Figures 2.6 and 2.7. The first star forms when the mass of the halo is $7.5 \times 10^6\,\mathrm{M}_\odot$. About 30 Myr later, 80% of the metals generated in the PISN have propagated as far as 420 parsecs from the core. The injection of metals by the PISN is enough to bring the gas metallicity in some cells hundreds of parsecs away from the galactic center to be as high as $\log_{10} Z/Z_\odot = -1$. After another 50 Myr, the effects of the outflow have subdued. Not only are 80% of generated metals now entirely confined to the innermost 120 parsecs, but the metals have diffused, and the maximum metallicity has decreased by 1 dex. This suggests that the inflow of new primordial gas is playing a greater role in the evolution of the galaxy than the outflow generated by the PISN. The majority of metals produced by PISN do not escape into intergalactic space.

We emphasize that some galaxies in this mass range should have hosted Pop III stars at earlier times than those resolved by our simulations. The apparent ineffectiveness of Pop III feedback demonstrated here shows that simulations which do not resolve halos with $M_h < 3 \times 10^6\,\mathrm{M}_\odot$ are missing a portion of galactic evolution. This omission could mean that Pop III stars should self-terminate at earlier times, hence decreasing their contribution to the cosmic ionizing background. On the other hand, the expulsion of baryons from low-mass halos leads to suppression of Pop II star formation, which implies we may also overestimate the Pop II rates. The balance of these effects will be explored further in Chapter 3.

Figure 2.9 shows how far metals propagate in galaxies relative to the stellar cores. The "gas metal half-mass radius" is calculated as the more familiar stellar half-mass radius, but

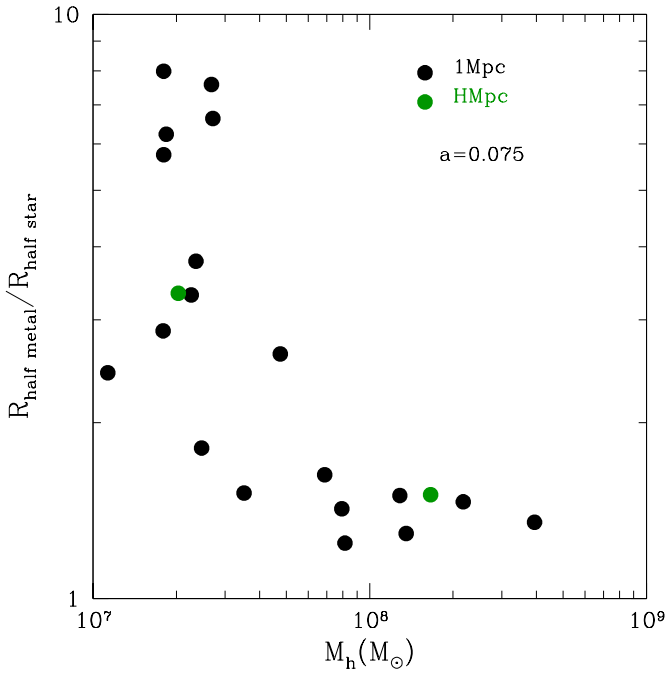

Figure 2.9. The ratio of gas metal half-mass radius to stellar half-mass radius vs. halo mass for galaxies which have had at least one PISN, at $a = 0.075$ ($z = 12.3$) in runs OverDense_nH1e4_fid (black) and OverDense_HMpc_HiRes (green). Metals propagate further relative to the stellar cores in galaxies of lower mass.

tracing the total mass of metals in the gas phase instead of stellar mass. In general, the metals are always able to propagate out beyond the stellar cores, but the extent depends strongly on galaxy mass. For a homogeneous comparison, and potential future probes by observation, we plot all star-forming galaxies from a single epoch, $a = 0.075$ ($z = 12.3$). In galaxies with $M_h \geq 10^8\,M_\odot$, the metals remain within a factor of 2 of the stellar radius. In less massive galaxies, metals are able to propagate further, sometimes by as much as a factor of 10, owing to the lower potential wells of these galaxies. Nonetheless, considering that the stellar half-mass radii of all our galaxies range 5-30 pc, the location of the bulk of metals is still limited to only the innermost regions of galaxies.

2.3.2 Effects of the uncertainty in Pop III feedback and IMF

With our additional runs, we can check if the relatively inefficient feedback is due to the specific fiducial parameters that we adopted. However, even with the extreme Pop III feedback prescriptions described in Section 2.2.3, we find that the baryon fraction within the virial radius is never significantly depleted. At $a = 0.05$ ($z = 19$), 10 Myr after the PISN

explosion in the first star-forming galaxy, we find f_b = 10.3% and 5.8% in runs OverDense_ExtremeSN and OverDense_ExtremeRad, compared to 11.5% in the fiducial run. In the case of the extremely energetic PISN, this is a relatively small resulting difference for a 10-fold increase in the thermal energy and ionizing radiation output. The mass of the galaxy at this epoch is $9 \times 10^6\,M_\odot$, which we have shown to be large enough to withstand standard Pop III feedback. On the other hand, the difference is more pronounced in the case of extreme ionizing feedback, indicating that the disruptive efficacy of supernovae is significantly increased when it explodes in a region where neutral hydrogen has been more effectively ionized and dispersed by radiative feedback.

The effect of both types of extreme feedback on the propagation of metals is stronger. Metals tend to be blown out of galaxies anisotropically, often extending outwards in directions orthogonal to filaments, into lower density regions. The galactocentric radius that encloses 80% of the metals formed in the PISN stretches out to $1.7R_{vir}$ 20 Myr after the explosion in both run OverDense_ExtremeSN and run OverDense_ExtremeRad, compared to just $0.91R_{vir}$ in the fiducial run. These metals do not fully escape the gravitational pull of the galaxy, however, and 90 Myr after the explosion, the galaxies from both extreme feedback runs contain 80% of the metals from the first PISN within the virial radius (in the fiducial run, they are contained within just $0.15R_{vir}$).

Though we increased the feedback effects by a factor of 10, only modest and temporary differences were observed between the runs. Such inefficiency of feedback demonstrates the weak coupling of thermal energy from PISNe to the ISM at the densities and temperatures resolved by our simulations, as almost any amount of energy can be quickly radiated away. This can be seen when considering typical cooling times in the ISM, $\tau_{cool} = k_b T/\Lambda n \approx 3000(T/10^4\,\mathrm{K})(1\,\mathrm{cm^{-3}}/n)\,\mathrm{yr}$, for $\Lambda = 10^{-23}\,\mathrm{erg\ s^{-1}\ cm^{-3}}$ (Hopkins et al., 2011). The cooling time of the dense, filamentary gas surrounding the supernova remnant ($n = 10\,\mathrm{cm^{-3}}$, $T = 10^4$ K) is just ~300 yr, which is comparable to a typical timestep in our simulations (~500 yr). This dense gas mixes with the shock-heated supernova remnant, allowing the entire region to return to the ambient temperature of the ISM within a few Myr.

The impact of extreme feedback is more pronounced in the IGM, particularly in run

OverDense_ExtremeRad. The mass fraction of ionized gas between 1-3 kpc from the galactic center is enhanced by a factor of ~200, compared to the fiducial run, even 40 Myr after the PISN. Within the same distance range, the IGM temperature is a factor of ~10 higher at this epoch. The relatively hot and ionized IGM in turn could affect accretion rates onto galaxies at later times.

The effect of making PISNe ten times more powerful in the H Mpc box was more drastic, as this box sampled lower mass galaxies. The baryon fraction in the first galaxy dropped below 10^{-5} after the first PISN, compared to 1.7% in the standard run OverDense_nH1e4_fid. The radii enclosing 80% of the baryons and metals are twice as large as in the standard run, demonstrating that the added energy in this extreme run coupled with the ISM more efficiently. Even with the standard feedback prescription we would expect a strong blowout in a halo of this mass ($2.7 \times 10^6\,M_\odot$ at this epoch). However, in the fiducial run this galaxy ultimately gobbled up most of the ejected metals. On the other hand, the extreme PISN energy (270×10^{51} erg) is able to completely destroy the high-density gas clouds needed for star formation, prevent re-accumulation of dense gas from filaments, and cause the metal ejecta to travel far enough into the IGM where they may never fall back onto the galaxy.

These tests indicate that given enough energy input, the host halos of Pop III stars can become completely devoid of gas for cosmologically-significant intervals of time, particularly when they are below the mass threshold $\sim 3 \times 10^6\,M_\odot$. However, for the feedback parameters currently considered realistic (our fiducial runs), the feedback of the first stars is limited as illustrated in Figures 2.6 and 2.7.

In the run with low masses of Pop III stars (OverDense_LowMass), without any PISN, metal transport is extremely ineffective. At $a = 0.055$, 80% of the metals that have been generated by stars in the first star-forming galaxy are confined within 75 pc of the galactic center, compared to 420 pc in the fiducial run. This test demonstrates that if Pop III stars did not have a top-heavy IMF, their contribution to enriching the IGM would be further marginalized. These results agree qualitatively with the work of Ritter et al. (2012), who argued that filamentary accretion was never significantly disrupted if Pop III stars had low or moderate characteristic masses and exploded in type II supernovae.

2.4 Discussion and Conclusions

We have presented the results of simulations that implemented primordial star formation in the cosmological code ART. We find that the effects of stellar feedback on the amount of baryons and metals within the first galaxies depend strongly on galaxy mass. For the lowest-mass galaxies (M_h ~$10^6\,M_\odot$) our results are similar to those of Bromm et al. (2003); Whalen et al. (2008); Wise & Abel (2008); Wise et al. (2012a), with gas and metals often being driven well beyond the virial radius of the Pop III star's host galaxy. For more massive galaxies ($M_h \geq 10^7\,M_\odot$), however, a single PISN is not effective in evacuating the galactic ISM, as suggested by Wise & Cen (2009). Feedback from Pop III stars does not typically inject enough energy into the massive halos to permanently photo-evaporate the gas, and drive metal-rich outflows past the virial radius. While Pop III stars can temporarily expel gas and quench star formation, the ISM begins to replenish soon after the SN explosion, as accretion from filaments at this epoch is very fast. All galaxies considered in our analysis with at least $M_h \approx 3 \times 10^6\,M_\odot$, and some which are even less massive, appear to have more than 50% of the universal baryon fraction restored 100 Myr after the first Pop III supernova event. Metals are ejected anisotropically, and can travel relatively longer distances through the diffuse IGM in directions perpendicular to the dense filaments which feed galactic accretion. This means that it typically takes more time for the ejected metals to be re-accreted into the galaxy, but we have demonstrated that this re-accretion does frequently occur, even in low-mass galaxies.

The aforementioned dividing line of $M_h \approx 3 \times 10^6\,M_\odot$ is important for determining whether the energy injection from the supernova at the end of the star's life can expel gas and metals out to a significant distance. The concept of a dividing line between early galaxies that suffer from significant blowout from those that do not has been considered in prior work (e.g. Ciardi et al. 2000; Ricotti et al. 2002b). However, our results point to a considerably lower threshold than what had been expected, as all but the very first galaxies are apparently robust to PISN feedback when continued accretion from filaments and the fallback of ejecta into the growing galaxy is considered. The strength of this conclusion is bolstered by our use of a very strong feedback model for Pop III stars (even in our

fiducial runs). In addition, the first stars may have a lower characteristic mass (Greif et al., 2011), which would make PISNe less frequent and the feedback effects would be further marginalized (Ritter et al., 2012).

In order for simulations to capture the full range of relevant effects from Pop III star formation, resolving halos around $10^6\,\mathrm{M}_\odot$ with a sufficiently large number of particles is critical. With insufficient resolution (less than 1000 DM particles for $10^6\,\mathrm{M}_\odot$ halos), all galaxies seem to reach $M > 10^7\,\mathrm{M}_\odot$ without having yet formed a star. Since these galaxies are already beyond the $M_h \approx 3 \times 10^6\,\mathrm{M}_\odot$ dividing line, they display few disruptive effects from Pop III feedback. Aggressive super-Lagrangian refinement may help resolve star formation in halos of lower mass, but requires a prohibitively large number of computations. A more practical approach is to begin simulations with sufficiently high resolution in the initial conditions.

CHAPTER 3

The epoch of Population III stars

3.1 Introduction

The first stars in the universe formed in gas devoid of metals. This exotic environment may have caused the initial mass function (IMF) for Population III stars to be different from the modern day case. Namely, the high Jeans mass of metal-free gas suggests a top-heavy IMF (Abel et al., 2000; Bromm et al., 1999; Abel et al., 2002; Yoshida et al., 2003). In turn, the feedback processes in the first stars may have been more drastic, prompting the release of extreme amounts of ionizing radiation (Tumlinson & Shull, 2000; Bromm et al., 2001a) and the occurrence of pair instability supernovae (PISNe) (Heger & Woosley, 2002).

To explore the effects of these stars on their host galaxies, we developed a model for Pop III star formation and feedback and implemented it into the adaptive refinement tree (ART) code, as described in a companion paper, (Muratov et al. 2013a, see Chapter 2). Pop III stars were modeled to form in gas that was dense, partially molecular, and of primordial composition. Pop III SNe and ionizing radiation feedback were enhanced relative to their Pop II counterparts, and the first PISNe seeded the ISM with metals. We ran a suite of cosmological simulations with this model, and found that the dynamical impact of Pop III feedback depended strongly on the galaxy mass. In agreement with previous work in the field (e.g. Bromm et al. 2003; Whalen et al. 2008; Wise et al. 2012a), we found that PISNe were able to efficiently expel gas and metals from the $M_h \sim 10^6\,\mathrm{M}_\odot$ halos expected to host the very first stars (Tegmark et al., 1997). However, these effects were often temporary, as cosmological inflows of fresh gas restored the baryon fraction to the universal value. The metals, which had previously escaped past the virial radius, also typically fell back into the growing potential wells of the accreting galaxies, leaving the intergalactic medium (IGM)

mostly pristine. In galaxies with mass $M_h > 10^7\,\mathrm{M}_\odot$, most gas remained bound even after a PISN event, and metals were not ejected past the virial radius.

Since Pop III stars by definition only form in primordial gas, the large amount of metals released in PISNe leads to the 'self-termination' of Pop III star formation (Yoshida et al., 2004). According to our findings in Chapter 2, this self-termination can only be local, as enrichment of the IGM and external halos is rather minimal from single PISNe. Therefore, determining the epoch when Pop III termination becomes universal is a somewhat different question (Tornatore et al., 2007). Pop III star formation could be relevant for a much longer phase of cosmic history if a Pop III star formed in every pristine halo with $M \gtrsim 10^6 - 10^8\,\mathrm{M}_\odot$ prior to reionization, as the abundance of such halos increases considerably with cosmic time.

Population II star formation can commence in galaxies once they are either sufficiently massive to enable the rapid gas cooling by atomic hydrogen lines, or enriched enough to enable efficient metal cooling (Ostriker & Gnedin, 1996). Because the feedback of Pop II stars is weaker than that of Pop III, Pop II star formation should ramp up rapidly in the host galaxy, provided that accretion from filaments continues to bring new supply. However, the relative weakness of the feedback, taken in conjunction with the plethora of primordial sites where Pop III stars may still form, means that the host galaxy, as well as the universe as a whole, are still influenced by Pop III stars for some time after Pop II star formation begins. Though this scenario has already been explored through semi-analytical models(e.g. Scannapieco et al. 2003; Yoshida et al. 2004; Schneider et al. 2006) and numerical simulations (e.g. Tornatore et al. 2007; Maio et al. 2010, 2011; Greif et al. 2010; Johnson et al. 2013; Wise et al. 2012a), understanding this transition quantitatively is relevant for the ability of future observational facilities such as the James Webb Space Telescope (JWST) to observe galaxies dominated by Pop III stars. Studies thus far have shown that the first galaxies generally sit on the brink of detectability by JWST (Pawlik et al., 2011, 2013; Zackrisson et al., 2011, 2012).

In this paper, we follow the evolution of the galaxies described in Chapter 2 through the epoch of dominance of the first stars. This sample of simulated galaxies spans a range of masses and accretion histories, therefore representing a broad variety of cosmic envi-

ronments. We study the transition from Pop III to Pop II star formation, and quantify the duration of this epoch. We also explore the effect of cosmic variance, and determine the prevalence and importance of Pop III stars at various cosmic epochs.

3.2 Simulations

A full description of our simulation setup, including the details of both the Pop III and Pop II star formation recipes, is presented in Chapter 2. Here, we outline the setup only briefly. We perform the simulations with the Eulerian gasdynamics+N-body adaptive refinement tree (ART) code (Kravtsov et al., 1997; Kravtsov, 1999, 2003; Rudd et al., 2008; Gnedin & Kravtsov, 2011). We use a 256^3 initial grid with up to 8 additional levels of refinement. For most of our runs, we apply this grid to a 1 h^{-1} Mpc comoving box with the WMAP-7 cosmology ($\Omega_m = 0.28$, $\Omega_\Lambda = 0.72$, $h = 0.7$, $\sigma_8 = 0.817$, $\Omega_b = 0.046$, $\Omega_{DM} = 0.234$). This gives us a DM particle mass $m_{DM} = 5.53 \times 10^3\,\mathrm{M}_\odot$ and a minimum cell size of 22 comoving pc. We also employ a 0.5 h^{-1} Mpc comoving box, where using the same grid, the DM particle mass is set to $m_{DM} = 690\,\mathrm{M}_\odot$ and the minimal cell size is 11 comoving pc.

Pop III stars formed in the almost pristine gas with the abundance of heavy elements below the critical metallicity $\log_{10} Z/Z_\odot = -3.5$ (Bromm et al., 2001b). Cells were allowed to form Pop III stars if the gas density exceeded a threshold $n_{H,min}$, and a molecular hydrogen fraction threshold $f_{H_2,min}$. Through a series of convergence tests, we found that $n_{H,min} = 10^4\,\mathrm{cm}^{-3}$ and $f_{H_2,min} = 10^{-3}$ were appropriate values for the two thresholds.

The Pop III prescription was designed to test the maximum possible effect of feedback, relying on an IMF that was top-heavy. Half of the Pop III stars formed as $170\,\mathrm{M}_\odot$ particles and were set to explode in PISNe. Each PISN injected 27×10^{51} erg of thermal energy and $80\,\mathrm{M}_\odot$ of metals into the ISM (Heger & Woosley, 2002). The remaining 50% of Pop III stars formed as $100\,\mathrm{M}_\odot$ particles that explode in type II SNe, generating 2×10^{51} erg of energy. All Pop III stellar particles emited a factor of 10 more ionizing photons per second than their Pop II counterparts of the same mass (Schaerer, 2002; Wise & Cen, 2009). A suite of cosmological simulations performed with this model revealed that Pop III stars drastically affected halos with $M_h < 3 \times 10^6\,\mathrm{M}_\odot$, but not halos of higher masses. Extended convergence tests revealed that without sufficient mass resolution, it was easy to miss these

important dynamical effects.

In gas that is enriched beyond the critical metallicity, Pop II star formation is modeled according to the molecular-based star formation recipe presented in Gnedin et al. (2009) and Gnedin & Kravtsov (2011).

In Table 1 we list all of the simulations from the suite which we analyze in this paper. Some of the simulations that were used in Chapter 2 for convergence tests and for determining the best parameters for Pop III star formation are not included here. In addition to the three fiducial 1 h^{-1} Mpc boxes (UnderDense–_nH1e4_fid, UnderDense+_nH1e4_fid, and OverDense_nH1e4_fid), we study runs with extreme feedback, where we increased PISNe energy (run OverDense_ExtremeSN) and ionizing photon emission (run OverDense_ExtremeRad) by a factor of 10. We also include a run with a conservative "low-mass" Pop III IMF that assigns the same feedback parameters to Pop III as for Pop II stellar particles (run OverDense_LowMass) to account for the present uncertainty in the Pop III IMF (e.g. O'Shea & Norman 2007; Greif et al. 2011).

In Chapter 2, we demonstrated that mass resolution was crucial for capturing Pop III star formation in halos of mass M_h~$10^6\,\mathrm{M}_\odot$, and that halos with $M_h < 3\times10^6\,\mathrm{M}_\odot$ were most susceptible to Pop III feedback. Our fiducial 1 h^{-1} Mpc runs lacked the resolution to study these objects effectively, but using the same grid on a smaller box increased the resolution sufficiently while keeping the computational cost down. For this reason, we also use a 0.5 h^{-1} Mpc box (referred to as H Mpc) for run OverDense_HMpc_HiRes, to test the validity of our results in the low-mass regime. Run OverDense_SL7 employs a super-Lagrangian refinement, allowing for the simulation to more effectively zoom in on overdense regions in primordial galaxies, allowing Pop III stars to form at early times in low mass halos, hence extending the mass range of our sample.

Table 3.1: SIMULATION RUNS OF CHAPTER 3

Run	Base grid	ℓ_{max}	dx (pc)	$m_{DM}(\mathrm{M}_\odot)$	$n_{H,min}(\mathrm{cm}^{-3})$
Description					
Fiducial runs					
UnderDense−_nH1e4_fid	256^3	8	22	5500	10000
Underdense box, no high-mass galaxies, fiducial parameters					
UnderDense+_nH1e4_fid	256^3	8	22	5500	10000
Underdense box, one high-mass galaxy, fiducial parameters					
OverDense_nH1e4_fid	256^3	8	22	5500	10000
Overdense box, fiducial parameters					
Alternative refinement & resolution OverDense_SL7	256^3	8	22	5500	10000
Super-Lagrangian refinement					
OverDense_HMpc_HiRes	256^3	8	11	690	10000
0.5 h^{-1} Mpc box, high spatial & mass resolution					
Alternative feedback					
OverDense_ExtremeSN	256^3	8	22	5500	10000
Extreme PISNe (Section 3.3.5)					
OverDense_ExtremeRad	256^3	8	22	5500	10000
Extreme Pop III radiation field (Section 3.3.5)					

OverDense_LowMass	256^3	8	22	5500	10000
Pop III IMF and feedback mirror Pop II (Section 3.3.6)					

Column 1.) Name of the run;

2.) Base grid, number of DM particles, number of root cells;

3.) Maximum number of additional levels of refinement;

4.) Minimum cell size at the highest level of refinement in comoving pc;

5.) DM particle mass in $M_\odot$;

6.) Minimum H number density for Pop III star formation in cm^{-3};

7.) Further description of the run.

3.3 Results

3.3.1 Cosmic Variance

Using the DC mode formalism for the generation of initial conditions (Sirko, 2005; Gnedin et al., 2011) allows us to test several representative regions of the universe without sacrificing resolution, as would be needed were we to simulate a larger cosmological volume. With a single parameter that stays constant over time in a given simulation box, Δ_{DC}, we encode the amplitude of density fluctuations on the fundamental scale of the box. Although many studies have been done to understand the effects of cosmic variance on the dark matter halo mass function (Tinker et al., 2008), studying it in hydrodynamic simulations is significantly more difficult. Here we present an analysis of the variance of the three different 1 h^{-1} Mpc boxes in our study. For reference, the DC mode values are $\Delta_{DC} = -2.57, -3.35$, and 4.04 in Box UnderDense–, UnderDense+, and OverDense respectively. The H Mpc box has a DC mode of $\Delta_{DC} = 5.04$. Large positive values of Δ_{DC} indicate an overdense region.

The first Pop III stars form around scale factor $a \approx 0.047$ ($z \approx 20.3$) in both Box UnderDense+ and Box OverDense. On the other hand, Box UnderDense– stalled significantly, and no star formation occurs until $a \approx 0.073$ ($z \approx 12.7$), translating to a time difference of 165 Myr. This wide range demonstrates immediately that the epoch when Pop III star formation begins is a strong function of local overdensity. If voids like the one represented by Box UnderDense– are indeed not enriched by external sources, Pop III stars could have existed in these voids until the end of cosmic reionization, at epochs that will be probed by JWST (Hummel et al., 2012) and the Large Synoptic Survey Telescope (LSST) (Trenti et al. 2009; but see Pan et al. 2012b).

Figure 3.1 shows the number of halos capable of hosting Pop III stars. Box OverDense clearly dominates over the other two across the entire mass range, and since halo mass strongly correlates with the density of central gas cells, Box OverDense is host to many more star forming galaxies. Despite having fairly similar DC mode values, Box UnderDense– and Box UnderDense+ have disparate halo abundances at the early epochs considered in our study. This is particularly visible at the high-mass end, where the most massive halo in Box UnderDense– has about 10 analogs in Box UnderDense+. The H Mpc

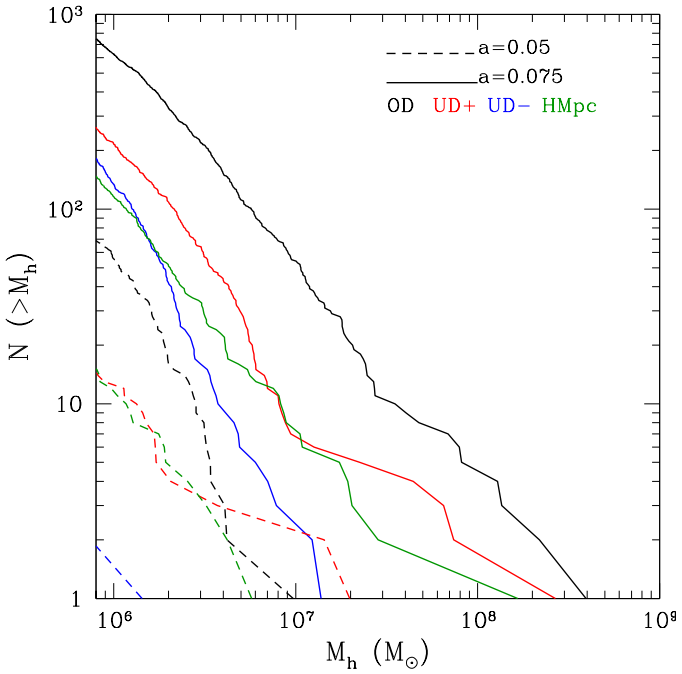

Figure 3.1. Cumulative number of halos vs. mass at two epochs for the three 1 h^{-1} Mpc boxes and one 0.5 h^{-1} Mpc box. The smallest mass plotted corresponds to the earliest halo to form a Pop III star among all of our runs. Box OverDense (black) has an order of magnitude more halos than Box UnderDense+ (red) and Box UnderDense– (blue) across almost the entire range of masses considered. While Box UnderDense+ and Box UnderDense– have similar DC mode values, the former clearly has more massive halos. One Box UnderDense+ galaxy in particular is on par with the most massive galaxies in Box OverDense. The H Mpc box (green, run OverDense_HMpc_HiRes) represents only an eighth of the volume of the other boxes, and therefore hosts fewer massive galaxies while still representing an overdense region.

box contains fewer halos than Box UnderDense+, but per unit volume it contains higher density of massive halos, consistent with its higher Δ_{DC} value. At the time of formation of the first stars (around $a = 0.05$), none of the halos are more massive than $2 \times 10^7\,M_\odot$. Figure 3.2 presents the star formation rate (SFR) density in each box for the runs with fiducial parameters. We can immediately see that SFRs vary by orders of magnitude among the three boxes. Only two galaxies are able to form Pop III stars in the extreme void represented by Box UnderDense–, and the total mass of Pop II is only $2300\,M_\odot$ by $a = 0.1$ ($z = 9$) in this run. Such a narrow margin of error suggests that it is possible that with slightly different parameters for the initial overdensity, galaxies in this box may have never been able to form Pop III stars before all halos were stripped of gas by external ionizing radiation.

Box UnderDense+ and Box OverDense have similar global Pop II SFRs at $a < 0.07$ ($z > 13$), as each box is being dominated by only a few galaxies at early times. However,

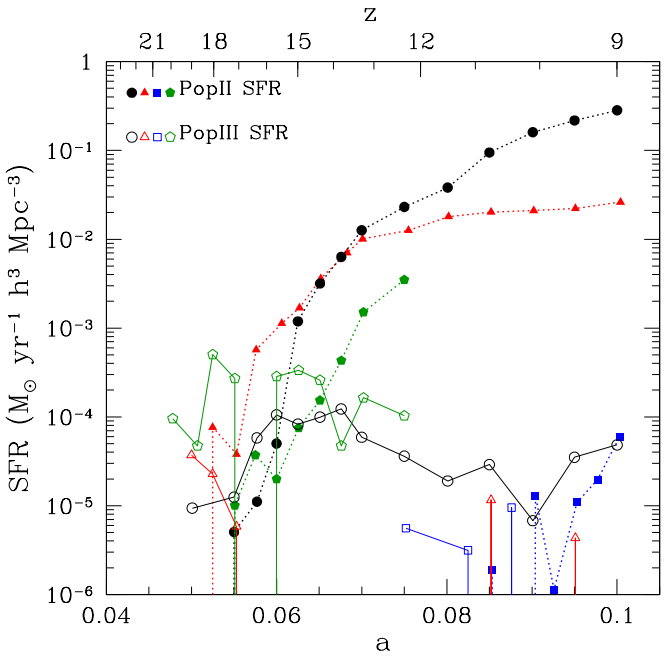

Figure 3.2. The SFR density vs scale factor for runs UnderDense−_nH1e4_fid (blue squares), UnderDense+_nH1e4_fid (red triangles), OverDense_nH1e4_fid (black circles), and OverDense_HMpc_HiRes (green pentagons). Global SFRs are significantly affected by the initial conditions in their respective simulation boxes. Pop III SFRs are denoted by open circles, Pop II SFRs by filled circles.

the disparity between the mean density of the boxes becomes evident at late times, as Box OverDense becomes filled with a population of massive halos able to host dense cores in which Pop III, and subsequently, Pop II star formation is initiated. Box UnderDense+ continues hosting only one such galaxy for a large fraction of the duration of the simulation until $a = 0.085$, when an external halo of primordial composition reaches the Pop III threshold density and later merges with the central galaxy.

In contrast to the large differences between the simulation boxes, we find the variance in the SFR between various realizations within the same box to be low. Differences between the setup of our test runs in Chapter 2 primarily affect Pop III stars, but at $a = 0.055$, the total number of PopIII stars in Box OverDense was the same. Furthermore, Pop II stars constitute most of the star formation at $a > 0.07$. Once a given galaxy transitions to Pop II as the dominant stellar population, details of Pop III star formation do affect the history of that galaxy. The Pop III SFR stays relatively constant between 10^{-5} and $10^{-4}\,M_\odot$ yr^{-1} throughout the entire simulation. The range of values and relative constancy of this SFR up to z~6 are similar to the results found in larger-scale (4 Mpc) SPH simulations by Johnson

et al. (2013).

The H Mpc box, represented here by run OverDense_HMpc_HiRes, samples a very overdense region of the universe. The Pop III SFR density in this run significantly exceeds the other boxes before $a = 0.065$, but the Pop II SFR density never reaches the corresponding values for Box OverDense and Box UnderDense+. This disparity results from the fact that the 1 Mpc boxes ultimately sample more massive galaxies which are able to sustain high Pop II star formation rates. Instead, a large fraction of low-mass galaxies simulated in the H Mpc box have considerable gas blowout after PISNe, delaying Pop II star formation for a cosmologically significant period of time.

3.3.2 The Ejection and Gobbling of Pop III Metals

The initial enrichment of the galactic ISM in our simulations happens almost exclusively through internal Pop III SNe rather than external intergalactic winds. This means that every galaxy will first have a phase during which Pop III stars constitute the entirety of the galactic stellar mass and drive all feedback. As our model considers only PISNe as a source of metal feedback during this Pop III phase, the enrichment required to transition to Pop II star formation is accomplished by at least one energetic PISN explosion. The explosion can be particularly potent, as it occurs in a hot, ionized, and diffuse medium carved out by the Pop III star's enhanced ionizing radiative feedback. The explosion therefore disrupts, or at least displaces, the dense gas necessary for further star formation.

We refer readers to Chapter 2, as well as prior work (e.g. Whalen et al. 2008) for a detailed description of this process. Here, we note that the efficacy of the feedback of these individual massive stars depends significantly on the galaxy mass at the time of Pop III star formation. A $\sim 10^6\,M_\odot$ "minihalo" can lose a significant fraction of its baryons, while a more massive ($\sim 10^7\,M_\odot$) halo can better withstand the explosion. Most galaxies, particularly those in overdense environments, benefit from continued accretion from filaments, and can re-accrete the lost metals and baryons even if they were initially ejected beyond the virial radius. Particularly, many galaxies "gobble up" the metals that they previously ejected through a combination of rapid growth of their virial radius and true gravitational fallback of the ejecta. The virial radius of a given galaxy grows both because of continual

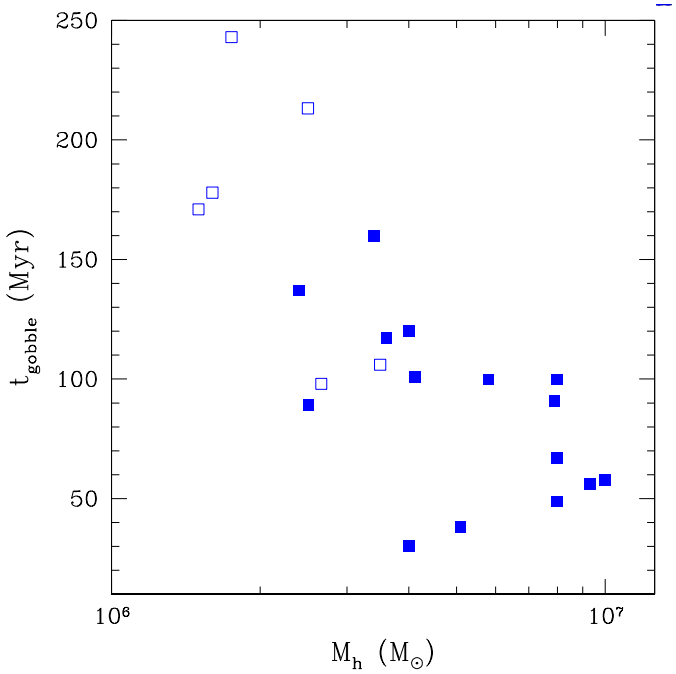

Figure 3.3. The timescale for a galaxy to re-accrete ("gobble up") 80% of the metals ejected by its PISN vs. halo mass at the time of the explosion. Open squares represent instances where the primary galaxy merges with a secondary that has already been enriched by internal star formation, preventing accurate tracing of the metals initially associated with the primary galaxy's PISN. The more massive the galaxy, the shorter is the gobble timescale. Plotted galaxies are from runs OverDense_nH1e4_fid, OverDense_SL7, and OverDense_HMpc_HiRes.

buildup of matter in the outer parts of galaxies, and because the virial radius depends on the critical density of the universe, which evolves with redshift. In Figure 3.3, we show the time it takes each galaxy to gobble up 80% of the mass of metals generated in the initial PISN. We confirm that this timescale is relatively rapid (50-100 Myr) for galaxies more massive than the $3 \times 10^6\,M_\odot$ threshold we discovered in Chapter 2. In galaxies which are sufficiently massive (above ~$10^7\,M_\odot$) at the time of the explosion, the metals rarely travel further than the virial radius, hence the gobble timescale is short.

Galaxies below the $3 \times 10^6\,M_\odot$ threshold have longer gobble timescales, and occasionally will merge with a more massive galaxy prior to the completion of the re-accretion process. Such instances prevent accurate tracing of the metals associated with the original PISN of the minihalo, and are therefore lower limits for the gobble timescales (denoted by open squares in Figure 3.3). In particularly underdense environments with slow filamentary accretion, the metals are permanently ejected from the minihalo. In run Over-

Dense_HMpc_HiRes, which resolves galaxies in the minihalo regime, only 20% of PISNe result in such permanent ejections. The metals that are the result of the permanent ejections do not pollute other halos sufficiently to initiate Pop II star formation over the course of our simulations. Though these metals enrich the IGM, the total fraction of volume that is enriched beyond the critical metallicity, $\log_{10} Z/Z_\odot = -3.5$, has a peak value of 0.05% at $z \approx 10.5$ in our most minihalo-dominated simulation, run OverDense_HMpc_HiRes.

The scatter of the gobble timescale for galaxies of a given mass is largely associated with their diverging mass accretion histories. We investigated the accretion history in both absolute and relative terms. The absolute growth rate is computed as the change in halo mass in the 100 Myr following the supernova. The relative growth is computed as the timescale for the halo to double the virial mass it had at the time of the explosion. We find that the absolute growth rate serves as a better predictor for the gobble timescale: all galaxies with $dM/dt > 0.2\,\mathrm{M}_\odot$ yr^{-1} have a gobble timescale of less than 100 Myr, while those with $0.04 < dM/dt < 0.2\,\mathrm{M}_\odot$ yr^{-1} show a wide range between 30 to 160 Myr. The 20% of PISNe that result in permanent ejections have growth rates under $0.04\,\mathrm{M}_\odot$ yr^{-1} and mass doubling timescales longer than 90 Myr (most longer than 150 Myr).

3.3.3 Transition to Normal Star Formation

For the remainder of the paper, we mainly deal with galaxies that are more massive than $3 \times 10^6\,\mathrm{M}_\odot$, and hence have relatively short gobble timescales. These galaxies typically dominate the star formation rate in the simulation boxes at late times. In particular, in this section, we focus on run OverDense_nH1e4_fid during the first 140 Myr after the first star forms. This corresponds to the range of scale factor $0.048 < a < 0.07$, or equivalently $20 > z > 13$. After that epoch the most massive galaxies in this box all form exclusively Pop II stars. Even more importantly, by $a = 0.07$, all Pop III stars in these galaxies have already exploded as SNe, and no longer provide radiative feedback. We will quantify the duration of the Pop III epoch in Section 3.3.4.

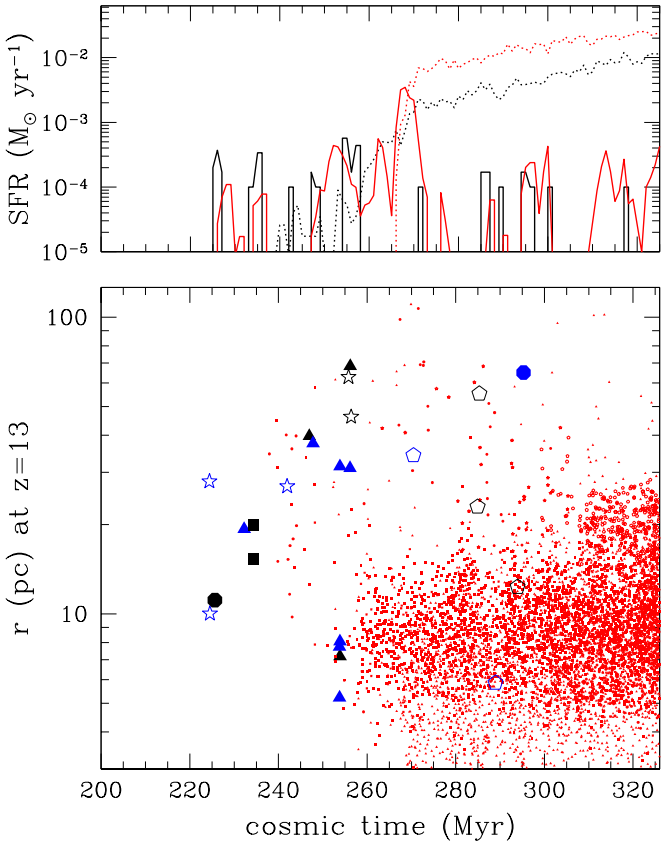

Figure 3.4. Bottom panel: distance of each star from its host galaxy's center at $a = 0.07$ ($z = 13$) vs the cosmic time when the star formed, for five of the most massive star-forming galaxies run OverDense_nH1e4_fid. Pop II stars are red points, the remnants of 100 $M_\odot$ Pop III stars are blue symbols, and 170 $M_\odot$ Pop III remnant tracer particles are black symbols. Individual galaxies are differentiated by different symbols. Every galaxy has a similar history: a few Pop III stars form and temporarily quench further star formation. However, eventually gas recollapses and self-sustained Pop II star formation begins within a 10-30 pc core, enriching the dense gas enough for Pop III star formation to be no longer possible within the galaxy. Pop III remnants often end up somewhat displaced from the Pop II core by as much as 100 pc. Top panel: the sum of SFRs for Pop III (solid) and Pop II (dotted) for these five galaxies, in black. For comparison, these rates are also plotted for run OverDense_LowMass, in red.

Figure 3.4 shows the distance of each stellar particle from the center of its host galaxy at $a = 0.07$ vs the cosmic time at which the stellar particle was formed. Five of the most massive star-forming galaxies are stacked together in this plot, showing the similarity in the assembly history of these galaxies. We see that each of the five galaxies does not host more than a few Pop III stars. If a $100\,M_\odot$ star is the first to form in a given galaxy, it will not release any metals, and further Pop III star formation will be possible after the gas recollapses. If on the other hand a $170\,M_\odot$ star is formed, Pop III star formation is rarely possible in the same galaxy again. The metals generated by a single PISN from a $170\,M_\odot$ star are enough to switch the galaxy's primary mode of star formation to be Pop II by pushing the metallicity of gas above the critical threshold $\log_{10} Z/Z_\odot = -3.5$ (as has been previously demonstrated e.g. Greif et al. 2010; Wise et al. 2012a). A few exceptions to this pattern happen when the sites of Pop III star formation within the galaxy are sufficiently spread apart (i.e. two Pop III stars form in two different halos that subsequently merge) or alternatively, the Pop III stars form sufficiently close together in time. The latter case is possible when many cells within the same giant molecular cloud reach the threshold density for star formation and several Pop III stars form before their feedback disperses the remaining gas in the cloud. This produces Pop III star multiples that are very close to each other in location and age.

The spatial distributions of Pop III stellar remnants and Pop II stars within their host galaxy are significantly different. Pop III stellar remnants can be found out to 70 pc from the center, while most Pop II stellar particles at this epoch are within 20 pc. Note that the centers of the galaxies are defined to only ~5-15 pc, because of the limited number of DM particles and force resolution. High-mass galaxies have better-defined centers than low-mass galaxies. Both stellar populations generally form close to the center of the host galaxy, but Pop III stars typically form at earlier times when the galaxy is not very massive, and the spatial concentration of DM particles is not very high. In this environment, dense star-forming gas has a clumpy structure, which is off-center from the DM cusp. After powerful Pop III feedback ejects gas from this region, the Pop III remnants become less bound, further increasing the apocenter of their orbit. On the other hand, Pop II stars are rapidly produced only once the galaxy becomes relatively massive and concentrated, and

the DM center is more closely associated with the location of the densest gas. Pop II thermal feedback and radiative feedback is not strong enough to shut off further Pop II star formation near the galactic center, ultimately leading Pop II stars to have a very high spatial concentration.

3.3.4 Epoch of Equivalence

The duration of the Pop III epoch has been previously studied (e.g. Trenti & Stiavelli 2009) by comparing the SFRs of the two populations, and finding the epoch where the Pop III SFR drops off dramatically compared to Pop II. Our Figure 3.2 shows that by $a = 0.07$ ($z = 13$), Pop II star formation is more common in all boxes except the void represented in Box UnderDense−. However this approach may not properly account for the impact Pop III stars have on the universe at early times, as Pop III stars are able to generate significantly more thermal, ionizing, and metal feedback per baryon compared to their Pop II counterparts (e.g. Tumlinson & Shull 2000). To more precisely quantify the impact of Pop III stars, we compute the relative "budget" of thermal energy, metals, and ionizing photons contributed in each galaxy by the two stellar populations, integrated over the lifetimes of all stellar particles. For the purpose of this analysis, we define the "epoch of equivalence" as the time when Pop II stars have generated just as much ionizing radiation as their Pop III counterparts. In turn, we define the "duration of the Pop III phase" as the length of time from the formation of the first Pop III star to the epoch of equivalence.

A global picture of the feedback budget is shown in Figure 3.5. Here, we calculate at each epoch the relative contribution of Pop III stars to the budget of ionizing photons, thermal energy from SNe, and metals injected into the ISM in run OverDense_nH1e4_fid. Every budget can be examined individually. Pop II stars have produced as many ionizing photons as Pop III stars by $a = 0.0633$ ($z = 14.8$). The total thermal energy released by SNe in Pop II stars surpasses that produced by Pop III stars at $a = 0.0642$ ($z = 14.6$). The total mass of metals produced in Pop II stars surpasses that of Pop III at $a = 0.0689$ ($z = 13.5$). Though we have chosen to define the epoch of equivalence in terms of ionizing radiation, any one of these feedback quantities leads us to the conclusion that the duration of the Pop III phase is short.

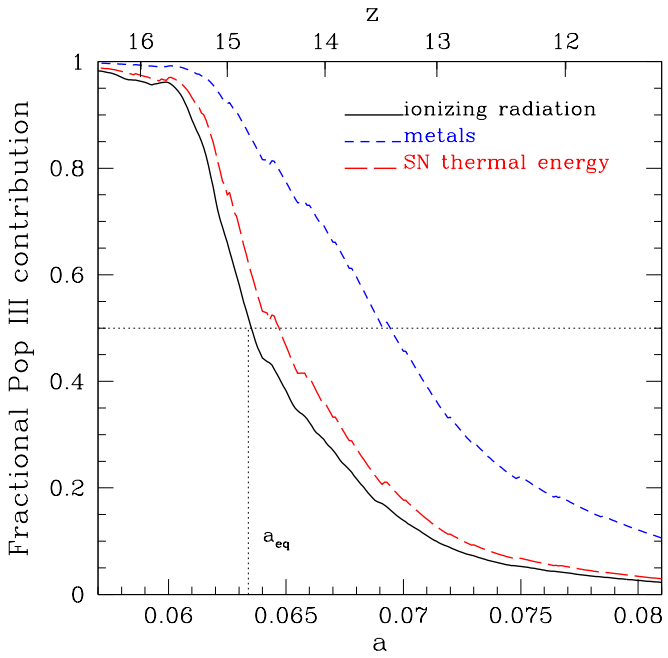

Figure 3.5. The fractional contribution of Pop III stars to the integrated feedback budget of the universe in run OverDense_nH1e4_fid vs scale factor. This fraction is calculated separately for metals (blue), ionizing radiation (black), and thermal radiation due to SNe (red).

The exact timing of the epoch of equivalence can be affected by the simulation's ability to resolve low-mass halos that are susceptible to the deleterious effects of Pop III feedback, and hence have longer delays for the onset of Pop II star formation. The resolution necessary to probe such low-mass galaxies is not attained in run OverDense_nH1e4_fid. Run OverDense_SL7 probes the same volume as run OverDense_nH1e4_fid, but Pop III stars form at earlier times, when the host halos are less massive. In this run, the epochs of equivalence shifts to $a = 0.0633$, $a = 0.0674$ and, $a = 0.0710$ for ionizing radiation, SNe thermal energy, and metals, respectively. For run OverDense_HMpc_Hires, which resolves many galaxies in the minihalo regime, the epoch of equivalence for the three types of feedback is reached at $a = 0.0721$, $a = 0.0735$, and $a = 0.0764$. These consistently small shifts in the duration of Pop III dominance give us no reason to believe that significant contributions of feedback from Pop III stars are missed in simulations that fail to resolve minihalos.

We did not include Box UnderDense+ and UnderDense− in this analysis, as they sample few star-forming galaxies and do not make significant contributions to the budget of either Pop III or Pop II feedback.

To understand the transition to normal star formation in individual galaxies, we also budget the feedback from stars within a given galaxy. All numbers quoted in this section reflect the integrated total feedback contributions over the lifetime of stellar particles up to the epoch considered. As a case study, we take the third galaxy to form a Pop III star in run OverDense_nH1e4_fid. The first star forms at $a = 0.055$ ($z = 17.2$), when the halo mass and gas mass of this galaxy are $9.8 \times 10^6\,\mathrm{M}_\odot$ and $1.2 \times 10^6\,\mathrm{M}_\odot$, respectively. At $a = 0.06$ ($z = 15.7$, 23 Myr later) this galaxy has formed a total of $118\,\mathrm{M}_\odot$ of Pop II stars. The first and only Pop III star that formed in the galaxy was a $170\,\mathrm{M}_\odot$ PISN progenitor. Although the total mass of each stellar population is comparable, the one Pop III star still has had dominant influence on the galaxy by way of its stronger feedback. The Pop III star contributed 98% of the metals, 88% of the ionizing photons, and 93% of the thermal energy. This galaxy is clearly in the phase of its evolution where the effects of Pop III star formation are most likely to be seen. However, we again note that despite this, the galaxy has not been significantly perturbed by the Pop III star's injection of energy. Only 1.7% of its hydrogen is ionized, and all of the metals are confined in the gas well within the virial radius.

By $a = 0.0675$ (80 Myr after the first star formed), one more Pop III star has formed, giving $270\,\mathrm{M}_\odot$ of Pop III mass formed in the galaxy, but now the total Pop II mass has increased to $1260\,\mathrm{M}_\odot$. This second Pop III star has a metallicity of $\log_{10} Z/Z_\odot = -4.7$, suggesting it formed in gas that was significantly enriched, but below the critical metallicity of $\log_{10} Z/Z_\odot = -3.5$. Therefore, whether the star should be Pop III or Pop II is rather sensitive to the particular setup of the model and stochastic effects. These two Pop III stars still contribute 91% of the metals, 61% of the ionizing photons and 71% of the thermal energy. We see that the energy contributions from Pop II stars are catching up with those of Pop III stars, but the metal contribution of Pop III stars remains dominant. This confirms the trend of Figure 3.5 that the large yield of metals produced by PISNe is the longest lasting contribution of Pop III star formation.

By $a = 0.075$ ($z = 12.3$, 134 Myr after the first star formed), the epoch of Pop III stars is clearly over in this galaxy despite the fact it has merged with another galaxy, which also hosted a recent PISN progenitor Pop III star. The combined initial mass of Pop III stars is

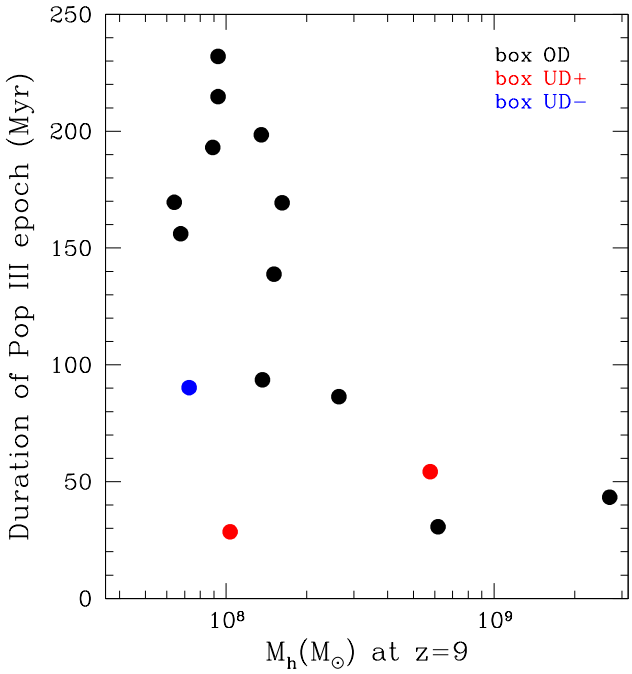

Figure 3.6. The duration of the epoch of Pop III dominance vs halo mass of host galaxies at $a = 0.1$ ($z = 9$). Plotted are galaxies in run OverDense_nH1e4_fid (black), run UnderDense+_nH1e4_fid (red), and run UnderDense−_nH1e4_fid (blue). More massive galaxies have shorter Pop III epochs.

now 440 $M_\odot$, while Pop II stars contribute 48000 $M_\odot$. The Pop III stars still contribute 26% of the metal budget, but a mere 5% of the ionizing photon budget and 8% of the thermal energy budget. By $a = 0.08$, the Pop III metal contribution has finally dwindled away to only 8%.

The epoch when Pop III stars have affected the evolution of the galaxy is thereby confined to between the time when the first star formed (around $a \approx 0.055$) and the time when the budget of energy and metals of Pop II stars begins to dominate over Pop III stars ($a \approx 0.07$). The window of Pop III dominance only corresponds to $\Delta a \approx 0.015$, or 100 Myr.

Figure 3.6 shows the duration of Pop III stars in this galaxy as well as in others, plotted against halo mass at $a = 0.1$ ($z = 9$). We note that the length of the Pop III epoch is generally between 20-200 Myr, confirming that the galaxy discussed in the above paragraphs is not unique. The figure also reveals a trend where more massive galaxies exit the Pop III stage earlier. These galaxies form in more biased regions, where the infall of fresh gas overcomes the negative feedback generated by the thermal and ionizing radiation of Pop III stars, and

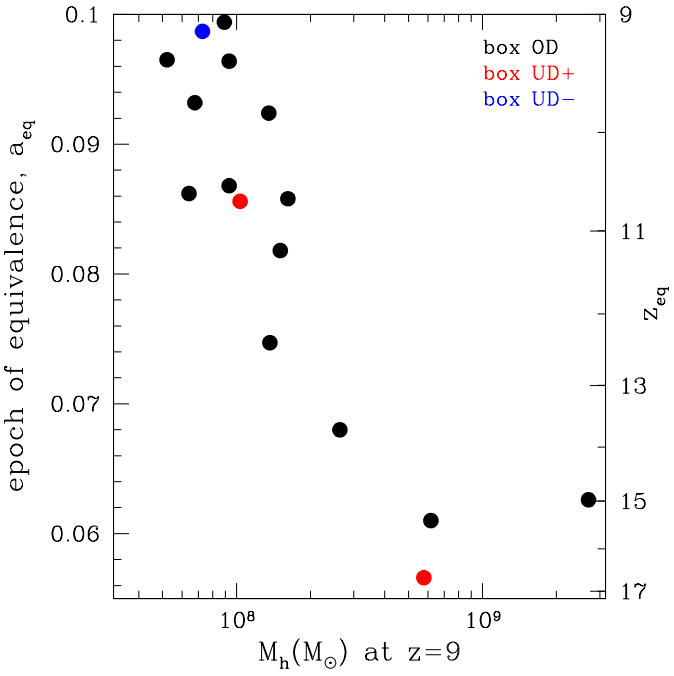

Figure 3.7. The epoch at which the contributions of Pop III and Pop II stars to the cumulative ionizing photon budget are equal vs. halo mass of host galaxies at $a = 0.1$ ($z = 9$). Plotted are galaxies in run OverDense_nH1e4_fid (black), run UnderDense+_nH1e4_fid (red), and run UnderDense−_nH1e4_fid (blue). The epoch of equivalence is earlier in more massive galaxies.

Pop II star formation erupts in the newly metal-enriched gas. Galaxies with higher mass at the time of Pop III star formation can also radiate the PISN energy more efficiently.

Figure 3.7 shows that the epoch of equivalence for individual galaxies is also strongly correlated with the mass of the halo at $a = 0.1$ ($z = 9$). This trend is perhaps not surprising, as the sites of formation for the first stars have to be the most biased regions of the simulation box, which would also produce the most massive galaxies. It is impressive, however, that the trend is so definitive that it suggests we can predict the epoch at which Pop III stars became subdued based simply on the halo mass.

We have also calculated the duration of the Pop III epoch by using the metal and thermal energy budgets instead of the ionizing photon budget. The thermal energy dominance of Pop III stars tends to last on average about ~10 Myr longer than the ionizing photon epoch, with little variance, while Pop III metals remain dominant for ~20-50 Myr longer. These differences are significant, and can be understood by looking at the relative efficiency of metal production of 170 $M_\odot$ Pop III stars compared to Pop II counterparts. While

our Pop II stellar particles only generate an amount of metals equivalent to 1.1% of their initial mass, 170 $M_\odot$ Pop III stars release almost 47% of their mass in metals following a PISNe, meaning that Pop III metal feedback is ~40 times more effective. In comparison, the number of ionizing photons per stellar baryon per lifetime is 6,600 for Pop II stars vs. 34,500 for 170 $M_\odot$ Pop III stars respectively, or only a factor of ~5 difference in efficiency. The variation between individual galaxies can be caused by the number of non-enriching 100 $M_\odot$ Pop III stars that form there, and by the rate at which Pop II stars form, as their feedback is released gradually over a longer interval of time. Despite this variation associated with using different feedback budget quantities, our conclusion on the brevity of the Pop III epoch within each galaxy remains robust.

3.3.5 Extreme Pop III Feedback

The duration of the phase of Pop III dominance may depend on the details of the implementation of stellar feedback. Here, we consider two additional runs with artificially enhanced Pop III feedback. By $a = 0.0625$ ($z = 15$), the fiducial run OverDense_nH1e4_fid has 34 Pop III stars, while run OverDense_ExtremeSN has 28 stars and the OverDense_ExtremeRad run has 19 stars. At least part of this variation can be explained by the stochastic effect that the OverDense_ExtremeRad run happened to form 170 $M_\odot$ PISNe progenitors more frequently than the other runs, hence quenching further Pop III star formation in their hosts. However, other evidence suggests that the extreme radiation does indeed quench star formation very efficiently. This can be discerned by examining Pop II star formation, as 20000 $M_\odot$ formed in run OverDense_nH1e4_fid, 7000 $M_\odot$ formed in run OverDense_ExtremeSN, and 4000 $M_\odot$ in run OverDense_ExtremeRad at $a = 0.0625$. While we showed in Chapter 2 that the effects of baryon expulsion in individual galaxies are temporary in both extreme feedback runs as well as the fiducial run, the cumulative effect of many more extreme feedback events appears to have suppressed of star formation at early times. More specifically, the relatively high IGM temperature in run OverDense_ExtremeRad (discussed in Chapter 2) means that newly accreted gas in this run will take longer to cool and condense to the densities required for star formation. In our models, radiative feedback is more effective than SN thermal feedback as a means of regulating star formation

and modifying the structure of the ISM and IGM.

In both extreme feedback runs, the epoch of equivalence is pushed back relative to the fiducial run, however it is always reached before $z = 13$. The epoch of equivalence for ionizing radiation increased by $\delta a = 0.0073$ in run OverDense_ExtremeRad, while the epoch of equivalence for SN thermal energy increased by $\delta a = 0.0068$ in run OverDense_ExtremeSN. Even the drastic modifications to Pop III feedback output causes only marginal shifts in the duration of their epoch of dominance.

3.3.6 Low Mass Pop III IMF

Using a standard Miller & Scalo (1979) IMF for Pop III stars causes the duration of the Pop III epoch to be somewhat shorter than in the top-heavy fiducial case. The epoch of equivalence for ionizing radiation is reached in run OverDense_LowMass at $a = 0.0622$, a mere 90 Myr after the first star forms.

However, the more important distinction in run OverDense_LowMass is the relative lack of suppression of star formation within individual galaxies. By $a = 0.065$ ($z = 14.4$), the total stellar mass formed in run OverDense_LowMass is a factor of ~3 larger than in the fiducial run OverDense_nH1e4_fid (see top panel of Figure 3.4). The majority of stellar mass in both runs is contained in the three most massive star-forming galaxies, each of which has more stellar mass in the run OverDense_LowMass. In this run, nearly all of the stellar particles that formed in these three galaxies are within 10 pc of their galaxy's center, enriching the dense gas slowly and steadily to perpetuate further star formation. A similar Pop II core is seen in the fiducial run, however some Pop II stars and Pop III stellar remnants are found up to 100 pc from their galactic centers (see bottle panel of Figure 3.4).

3.4 Discussion and Conclusions

We find that Pop III star formation can be expected to extend until at least $z \approx 10$. However, Pop III stars are not the dominant source of any form of global feedback past $a \approx 0.07$ ($z \approx 13$) in our models, and only dominate individual galaxies for 20 to 200 Myr.

The UV radiation of Pop III stars and their powerful supernovae can temporarily evacuate neutral gas from the first galaxies, creating an opportunity for ionizing flux from the

subsequent generations of Pop II stars to escape into the IGM. However, it is unlikely that the escape fraction was high because, as we showed in Chapter 2, accretion from filaments brings back most of the expelled gas on roughly the same timescale (~150 Myr) as the duration of the Pop III phase. This means that by the time star formation resumes at the center, the entire galaxy will again be surrounded by neutral gas, making it difficult for the ionizing photons to escape.

We also showed in Chapter 2 that most of the metals generated in PISN would eventually fall back into the host galaxy rather than remaining in the IGM. This fact, combined with the relative insignificance of the Pop III metal budget, leads us to conclude that Pop III stars did not play much of a role in enriching the IGM, and that instead enrichment happened by normal stellar populations during and after the epoch of reionization. Though Pop III star formation is self-limiting in individual galaxies, the absence of universal enrichment implies that Pop III stars could form in underdense regions long after the universe has primarily transitioned to normal star formation. It is in the least massive galaxies that formed in underdense regions where we might expect to find the most recent signatures of Pop III stars, suggesting that ultra-faint dwarf spheroidal galaxies and their disrupted remnants within the Milky Way stellar halo may be the best place to look. As the chemical signature of PISNe metals is expected to be different from that of Type II SNe, and the metal budget of Pop III stars is the last feedback tracer to be overtaken by Pop II, studying the spectra of old metal-poor stars with peculiar chemical abundances may be one of the most promising ways to observationally constrain the chemical history of the first galaxies. The first efforts in this field of 'stellar archeology' have already yielded interesting results (e.g. Frebel et al. 2007; Frebel & Bromm 2012), but there is more work to be done to increase the sample of metal-poor stars, and to improve our knowledge of their abundances, many of which are currently too faint for detailed spectroscopy. The next generation of spectroscopic and photometric surveys, such as Gaia-ESO and SkyMapper, will significantly improve the sample, while the light collecting area of upcoming large ground-based optical telescopes will enable detailed follow-ups of significantly fainter candidates (Frebel, 2011).

According to mock observational analysis for JWST by Pawlik et al. (2013), galaxies with SFR in excess of $0.1 M_{\odot}$ yr^{-1} may be detectable at $z > 10$. Only two galaxies in

our simulation box meet this criterion in the final snapshot at $z = 9$, and both are evolved far beyond the point where Pop III stars made a significant contribution to their feedback budgets. According to a separate analysis by Zackrisson et al. (2012), JWST surveys that target gravitationally lensed fields are much more likely to discover primordial galaxies than ultra-deep unlensed fields. These authors suggest that galaxies in which 0.1% of baryons are in the form of Pop III stars are detectable through direct starlight if they are in the lensed field of the galaxy cluster J0717.5+3745. However, none of our galaxies meet this criterion at any given time.

A more promising way to detect galaxies that are still dominated by Pop III stars is by looking for PISN events directly. The light curve of a PISN is thought to have a similar peak luminosity to SN Ia, but can remain bright for hundreds of days, producing a distinct signature (e.g. Scannapieco et al. 2005). This idea has been explored and shown to be feasible for both LSST (Trenti et al., 2009) and JWST (Hummel et al., 2012; Pan et al., 2012a; Whalen et al., 2012). Again, considering our findings that underdense regions of the universe may remain pristine until late times, we conclude that the best approach for observing campaigns is to focus on areas where the spatial clustering signal is low, such as cosmic filaments that connect two bright galaxies, to maximize the number of bright, low-redshift detections.

If Pop III stars instead formed with a normal IMF (e.g. Greif et al. 2011), they would have even less impact on the universe, as demonstrated by our low-mass IMF run. On the other hand, while our extreme feedback runs showed that the Pop II transition could be delayed with stronger Pop III feedback parameters, the transition inevitably happened before $z = 13$ and takes less than 250 Myr for even the least massive galaxies. In both cases, our modifications to the feedback output were simplistic ways to test the range of potential impact of Pop III stars without detailed modeling of additional physical processes. However, it is becoming clear that missing physics in galaxy formation simulations can have a wide variety of effects that are not accounted by our limited approach. For example, momentum-driven winds from radiation pressure can significantly stir the ISM increasing its susceptibility to other forms of feedback (Wise et al., 2012b; Agertz et al., 2013). The presence or absence of dust can affect the shielding capabilities of molecular clouds from

UV radiation, thereby modifying fragmentation and potentially the IMF of stars (Aykutalp & Spaans, 2011). Dark matter annihilation (Smith et al., 2012), cosmic rays (Jasche et al., 2007; Stacy & Bromm, 2007; Uhlig et al., 2012), magnetic fields (Turk et al., 2012), and X-ray feedback from the first black holes (Haiman, 2011; Jeon et al., 2012) could also be potentially important. We plan to explore them further in the next generation of simulations.

CHAPTER 4

Modeling the Metallicity Distribution of Globular Clusters

4.1 Introduction

A self-consistent description of the formation of globular clusters remains a challenge to theorists. A particularly puzzling observation is the apparent bimodality, or even multimodality, of the color distribution of globular cluster systems in galaxies ranging from dwarf disks to giant ellipticals (reviewed by Brodie & Strader, 2006). This color bimodality likely translates into a bimodal distribution of the abundances of heavy elements such as iron. We know this to be the case in the Galaxy as well as in M31, where relatively accurate spectral measurements exist for a large fraction of the clusters. In this paper we will interchangeably refer to metal-poor clusters as blue clusters, and to metal-rich clusters as red clusters.

Bimodality in the globular cluster metallicity distribution of luminous elliptical galaxies was proposed by Zepf & Ashman (1993), following a theoretical model of Ashman & Zepf (1992). The concept of cluster bimodality became universally accepted because the two populations also differ in other observed characteristics. The system of red clusters has a significant rotation velocity similar to the disk stars whereas blue clusters have little rotational support, in the three disk galaxies observed in detail: Milky Way, M31, and M33 (Zinn, 1985). In elliptical galaxies, blue clusters have a higher velocity dispersion than red clusters, both due to lack of rotation and more extended spatial distribution. Red clusters are usually more spatially concentrated than blue clusters (Brodie & Strader, 2006). All of these differences, however, are in external properties (location and kinematics), which reflect *where* the clusters formed, but not *how*. The internal properties of the red and blue clusters are similar: masses, sizes, and ages, with only slight differences. Even the metal-

licities themselves differ typically by a factor of 10 between the two modes, not enough to affect the dynamics of molecular clouds from which these clusters formed. Could it be then that both red and blue clusters form in a similar way on small scales, such as in giant molecular clouds, while the differences in their metallicity and spatial distribution reflect when and where such clouds assemble?

All scenarios proposed in the literature assumed different formation mechanisms for the red and blue clusters, and most scenarios envisioned the stellar population of one mode to be tightly linked to that of the host galaxy (e.g., Forbes et al., 1997; Cote et al., 1998; Strader et al., 2005; Griffen et al., 2010). The other mode is assumed to have formed differently, in some unspecified "primordial" way. This assumption only pushed the problem back in time but it did not solve it. For example, Beasley et al. (2002) used a semi-analytical model of galaxy formation to study bimodality in luminous elliptical galaxies and needed two separate prescriptions for the blue and red clusters. In their model, red clusters formed in gas-rich mergers with a fixed efficiency of 0.007 relative to field stars, while blue clusters formed in quiescent disks with a different efficiency of 0.002. The formation of blue clusters also had to be artificially truncated at $z = 5$. Strader et al. (2005), Rhode et al. (2005), and Griffen et al. (2010) suggested that the blue clusters could instead have formed in very small halos at $z > 10$, before cosmic reionization removed cold gas from such halos. This scenario requires high efficiency of cluster formation in the small halos and also places stringent constraints on the age spread of blue clusters to be less than 0.5 Gyr. Unfortunately, even the most recent measurements of relative cluster ages in the Galaxy (De Angeli et al., 2005; Marín-Franch et al., 2009; Dotter et al., 2010) cannot detect age differences smaller than 9%, or about 1 Gyr, and therefore cannot support or falsify the reionization scenario. Dotter et al. (2010) also show that the red clusters have larger dispersion of ages (15%, or about 2 Gyr) and those located outside 15 kpc of the Galactic center tend to show measurably lower ages, by as much as 50% (or 6 Gyr). In addition, Strader et al. (2009) find that the red clusters in M31 have lower mass-to-light ratios than the blue clusters, possibly indicating an age variation.

In this paper we set out to test whether a common mechanism could explain the formation of both modes and produce an entire metallicity distribution consistent with the

observations. We begin with a premise of the hierarchical galaxy formation in a ΛCDM universe. Hubble Space Telescope observations have convincingly demonstrated one of the likely formation routes for massive star clusters today – in the mergers of gas-rich galaxies (e.g., Holtzman et al., 1992; O'Connell et al., 1995; Whitmore et al., 1999; Zepf et al., 1999). We adopt this single formation mechanism for our model and assume that clusters form only during massive gas-rich mergers. We follow the merging process of progenitor galaxies in a Galaxy-sized environment using a set of cosmological N-body simulations. We need to specify what type and how many clusters will form in each merger event. For this purpose, we use observed scaling relations to assign each dark matter halo a certain amount of cold gas that will be available for star formation throughout cosmic time and an average metallicity of that gas. In order to keep the model transparent, we choose as simple a parametrization of the cold gas mass as possible. Finally, we make the simplest assumption that the mass of all globular clusters formed in the merger is linearly proportional to the mass of this cold gas.

Although such model appears extremely simplistic, we have some confidence that it may capture main elements of the formation of massive clusters. Kravtsov & Gnedin (2005) used a cosmological hydrodynamic simulation of the Galactic environment at high redshifts $z > 3$ and found dense, massive gas clouds within the protogalactic disks. If the high-density regions of these clouds formed star clusters, the resulting distributions of cluster mass, size, and metallicity are consistent with those of the Galactic metal-poor clusters. In that model the total mass of clusters formed in each disk was roughly proportional to the available gas mass, $M_{GC} \propto M_g$, just as we assume here.

We tune the parameters of our semi-analytical model to reproduce the metallicity distribution of the Galactic globular clusters, as compiled by Harris (1996). This distribution is dominated by the metal-poor clusters but is also significantly bimodal. We attempt to construct a model without explicitly differentiating the two modes and test if bimodality could arise naturally in the hierarchical framework.

We adopt a working definition of red clusters as having $[\mathrm{Fe/H}] > -1$ and blue clusters as having $[\mathrm{Fe/H}] < -1$. This definition should also roughly apply to extragalactic globular cluster systems. We use the concordance cosmology with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, $h = 0.7$.

4.2 Prescription for Globular Cluster Formation

4.2.1 Cold Gas Fraction

We follow the merging process of protogalactic dark matter halos using cosmological N-body simulations of three Milky Way-sized systems described in Kravtsov et al. (2004). The simulations were run with the Adaptive Refinement Tree code (Kravtsov et al., 1997) in a $25h^{-1}$ Mpc box. Specifically, we use merger trees for three large host halos and their corresponding subhalo populations. The three host halos contain $\sim 10^6$ dark matter particles and have virial masses $\sim 10^{12}\,\mathrm{M}_\odot$ at $z = 0$. Two halos are neighbors, located at 600 kpc from each other. The configuration of this pair resembles that of the Local Group. The third halo is isolated and is located 2 Mpc away from the pair. All three systems experience no major mergers at $z < 1$ and thus could host a disk galaxy like the Milky Way.

In addition to the host halos, the simulation volume contains a large number of dwarf halos that begin as isolated systems and then at some point accrete onto the host halo. Some of these satellites survive as self-gravitating systems until the present, while the rest are completely disrupted by the tidal forces. We allow both the host and the satellite systems to form globular clusters in our model.

We adopt a simple hypothesis, motivated by the hydrodynamic simulation of Kravtsov & Gnedin (2005), that the mass in globular clusters, M_{GC}, that forms in a given protogalactic system is directly proportional to the mass of cold gas in the system, M_g. We define the corresponding mass fraction, f_g, of cold gas that will be available for star cluster formation in a halo of mass M_h as

$$f_g \equiv \frac{M_g}{f_b M_h}, \tag{4.1}$$

where $f_b \approx 0.17$ is the universal baryon fraction (Komatsu et al., 2011).

The gas fraction cannot exceed the total fraction of baryons accreted onto the halo, which is limited by external photoheating and depends on the cutoff mass $M_c(z)$:

$$f_{in} = \frac{1}{(1+M_c(z)/M_h)^3}. \tag{4.2}$$

We use an updated version of the cutoff mass as a function of redshift (originally defined

by Gnedin 2000), based on our fitting of the results of recent simulations by Hoeft et al. (2006), Crain et al. (2007), Tassis et al. (2008), and Okamoto et al. (2008):

$$M_c(z) \approx 3.6 \times 10^9 \, e^{-0.6(1+z)} \, h^{-1} \, \mathrm{M}_\odot. \tag{4.3}$$

Given the scatter in simulation results and the numerical limitations of the modeling of gas physics, a reasonable uncertainty in this mass estimate is of the order 50%. However, the resulting cluster mass and metallicity distributions are not very sensitive to the exact form of this equation. Note that Orban et al. (2008) provided an earlier revision of the equation for $M_c(z)$; our current form is more accurate. If $M_c(z)$ falls below the mass of a halo with the virial temperature of 10^4 K, we set $M_c(z)$ equal to that mass:

$$M_{c,\mathrm{min}}(z) = M_4 \equiv 1.5 \times 10^{10} \Delta_{\mathrm{vir}}^{-1/2} \frac{H_0}{H(z)} \, h^{-1} \, \mathrm{M}_\odot, \tag{4.4}$$

where $\Delta_{\mathrm{vir}} = 180$ is the virial overdensity and $H(z)$ is the Hubble parameter at redshift z. This criterion ensures that we only select halos able to cool efficiently via atomic hydrogen recombination lines.

Some of the baryons accreted onto a halo may be in a warm or hot phase (at $T > 10^4$ K) unavailable for star formation, thus $f_g < f_{in} < 1$. We assume that only the gas in cold phase ($T \ll 10^4$ K) is likely to be responsible for star cluster formation. The cold gas fraction f_g is calculated by combining several observed scaling relations. From the results of McGaugh (2005), the average gas-to-stellar mass ratio in nearby spiral and dwarf galaxies can be fitted as

$$\frac{M_g}{M_*} \approx \left(\frac{M_*}{M_s(z)} \right)^{-0.7}, \tag{4.5}$$

where M_s is a characteristic scale mass, which we found to be $M_s(z=0) \approx 4 \times 10^9 \, \mathrm{M}_\odot$. This relation saturates at low stellar masses, where f_g cannot exceed f_{in}. At higher redshift the only information on the gas content of galaxies comes from the study by Erb et al. (2006) of Lyman break galaxies at $z = 2$. These authors estimated the cold gas mass by inverting the Kennicutt-Schmidt law and using the observed star formation rates. These estimates are fairly uncertain and model-dependent. Within the uncertainties, their results can be fitted

by the same formula but with a different scale mass: $M_s(z=2) \approx 2 \times 10^{10}\,\mathrm{M}_\odot$. To extend this relation to all epochs, we employ a relation that interpolates the two values:

$$M_s(z) \approx 10^{9.6+0.35z}\,\mathrm{M}_\odot. \tag{4.6}$$

An additional scaling relation is needed to complement equation (4.5) with a prescription for stellar mass as a function of halo mass. We compile it by combining the observed stellar mass–circular velocity correlation with the theoretical circular velocity–halo mass correlation. Woo et al. (2008) found that the stellar mass of the dwarf galaxies in the Local Group correlates with their circular velocities, which are taken as the rotation velocity for the irregular galaxies or the appropriately scaled velocity dispersion for the spheroidal galaxies. In the range $10^7\,\mathrm{M}_\odot < M_* < 10^{10}\,\mathrm{M}_\odot$, appropriate for the systems that may harbor globular clusters, the correlation is $V_c \propto M_*^{0.27\pm0.01}$. This can be inverted as $M_* \approx 1.6 \times 10^9\ \mathrm{M}_\odot\ (V_c/100\ \mathrm{km\ s^{-1}})^{3.7}$.

Cosmological N-body simulations show that dark matter halos, both isolated halos and satellites of larger halos, exhibit a robust correlation between their mass and maximum circular velocity (e.g., Fig. 6 of Kravtsov et al., 2004): $V_{\rm max} \approx 100\ (M_h/1.2 \times 10^{11}\,\mathrm{M}_\odot)^{0.3}$ km s^{-1}. This maximum circular velocity of dark matter is typically lower than the rotation velocity of galaxies because of the contribution of stars and gas. To connect the two velocities, we apply the correction $V_c = \sqrt{2}V_{\rm max}$, which reflects the observation that the mass in dark matter is approximately equal to the mass in stars over the portions of galaxies that contain the majority of stars. Then the equations in the last two paragraphs lead to $M_* \approx 5.5 \times 10^{10}\,(M_h/10^{12}\,\mathrm{M}_\odot)^{1.1}\ \mathrm{M}_\odot$.

We also need to extend this local relation to other redshifts. Conroy & Wechsler (2009) matched the observed number densities of galaxies of given stellar mass with the predicted number densities of halos of given mass from $z = 0$ to $z \sim 2$, averaged over the whole observable universe. They find that the stellar fraction f_* peaks at masses $M_h \sim 10^{12}\,\mathrm{M}_\odot$ and declines both at lower and higher halo masses. The range of masses of interest to us is below the peak, where we can approximate f_* dependence on halo mass as a power-law. The results from Fig. 2 of Conroy & Wechsler (2009) are best fit by a steeper relation

than we derived for the Local Group and also show significant variation with redshift at lower halo masses $\sim 10^{11}\,\mathrm{M}_\odot$: $M_* \propto M_h^{1.5}(1+z)^{-2}$ (there is much less variation with time around the peak at $10^{12}\,\mathrm{M}_\odot$, implying only a weak evolution in the total stellar density at $z < 1$). We adopt the same redshift dependence to our local relation, while using the shallower slope derived from Woo et al. (2008) because it deals with a population of halos in the mass range corresponding to the Milky Way progenitors. The stellar mass fraction of isolated halos is thus

$$f_* = \frac{M_*}{f_b M_h} \approx 0.32 \left(\frac{M_h}{10^{12}\,\mathrm{M}_\odot} \right)^{0.1} (1+z)^{-2}. \tag{4.7}$$

This relation steepens at low masses because of two additional limits on the gas and stellar fractions, which we impose to constrain the range of equations (4.5–4.7) to be physical.

First, the sum of the gas and stars ("cold baryons") cannot exceed the total amount of accreted baryons in a halo:

$$f_* + f_g \leq f_{in}. \tag{4.8}$$

At each redshift, there is a transition mass $M_{h,\mathrm{cold}}$, below which $f_* + f_g = f_{in}$ and above which $f_* + f_g < f_{in}$. For masses $M < M_{h,\mathrm{cold}}$ (but not too low, see next paragraph), we set $f_{g,\mathrm{revised}} = f_{in} - f_*$, with f_* still given by equation (4.7). We consider the baryons that are not included in f_g or f_* to be in the warm-hot diffuse phase of the interstellar medium.

Second, the ratio of stars to cold baryons, $\mu_* \equiv f_*/(f_* + f_g)$, is not allowed to increase with decreasing halo mass. For massive halos ($M_h > M_{h,\mathrm{cold}}$) μ_* monotonically decreases with decreasing mass because of the condition (4.5). At some intermediate masses $M_{h,\mu} < M_h < M_{h,\mathrm{cold}}$, μ_* continues to decrease but the gas fraction is reduced by the condition (4.8). At $M_h < M_{h,\mu}$, μ_* would reverse this trend and increase with decreasing halo mass because the cold gas is almost completely depleted. Such a reversal is unlikely to happen in real galaxies, which would not be able to convert most of their cold gas into stars. Therefore, for all masses $M_h < M_{h,\mu}$ we fix μ_* to be equal to the minimum value reached at $M_{h,\mu}$. This affects both f_* and f_g.

We expect our stellar mass prescription to apply in the range of halo masses from 10^9 to $10^{12}\,\mathrm{M}_\odot$, at least for the Local Group. However, this relation breaks when a halo becomes a

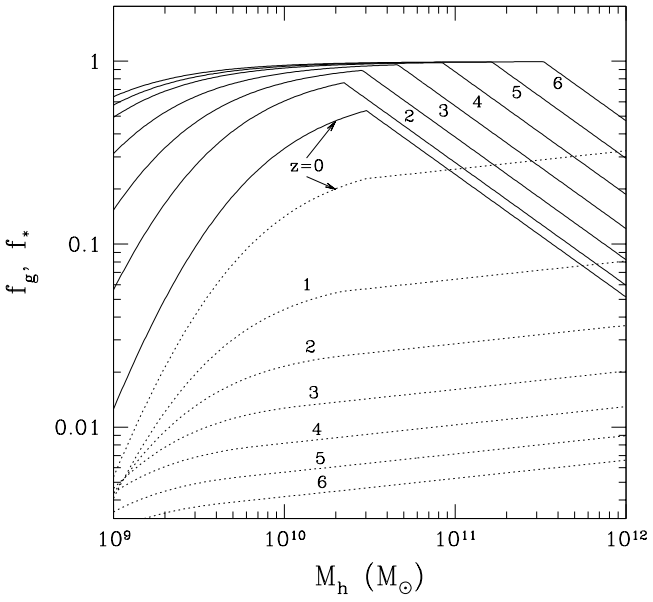

Figure 4.1. Gas mass fraction (*solid lines*) and stellar mass fraction (*dotted lines*) vs. halo mass in our model at redshifts $z = 0, 1, 2, 3, 4, 5, 6$. Stellar fractions monotonically increase with time while gas fractions decrease with time. The kink in the curves is due to our restriction on the maximum stellar fraction via eq. (4.8).

satellite of a larger system. Satellite halos often have dark matter in the outer parts stripped by tidal forces of the host, while the stars remain intact in the inner parts. Unless the satellite is completely disrupted, we keep its stellar mass fixed at the value it had at the time of accretion, even though the halo mass may subsequently decrease.

The simultaneous effects of the above scaling relations are difficult to understand as equations. Figure 4.1 illustrates graphically the values of the gas and stellar fractions used in our model at various cosmic times. At $z = 0$ the gas fraction peaks for halos with $M_h \sim 3 \times 10^{10}$ $M_\odot$. At lower masses it is reduced by the amount of accreted baryons (eq. 4.8), while at higher masses it is reduced by the gas-to-stars ratio (eq. 4.5). The stellar fraction follows equation (4.7) at high masses but drops faster at low masses because of the constraint (4.8). For a Galaxy-mass halo, $M_h \approx 10^{12}\,M_\odot$, our model gives $M_* \approx 5.5 \times 10^{10}\,M_\odot$ and $M_g \approx 9 \times 10^9\,M_\odot$. These numbers are consistent with the observed amount of the disk and bulge stars and the atomic and molecular gas in the Galaxy, from Binney & Tremaine (2008).

At earlier epochs at all masses of interest, the gas fraction is higher and the stellar

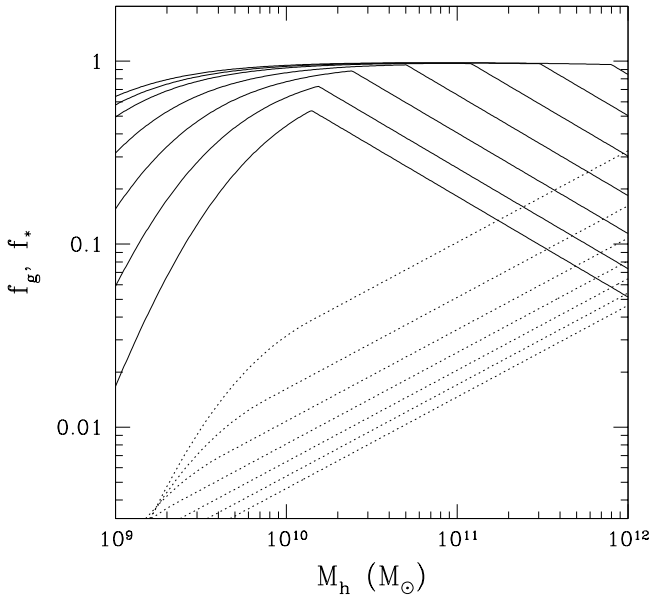

Figure 4.2. Same as Fig. 4.1, but for an alternative stellar mass prescription, eq. (4.9).

fraction is lower. There is a range of halos with $M_h \gtrsim 10^{10}$ $M_\odot$, which have an almost 100% gas fraction at redshifts $z > 3$. Such halos should be most efficient at forming massive star clusters.

We realize that our adopted relations for the evolution of the stellar and gas mass are not unique, as we are basing each fit on two data points. In order to test the sensitivity of our results to these assumptions, we consider alternative functional forms for these fits in Section 4.4.5. In particular, we give the stellar fraction a steeper dependence on halo mass and weaker dependence on cosmic time:

$$f_{*,\mathrm{alt}} = 0.32 \left(\frac{M_h}{10^{12}\,\mathrm{M}_\odot} \right)^{0.5} (1+z)^{-1}. \tag{4.9}$$

Such a slower evolution of the stellar mass is consistent with the observational studies of Borch et al. (2006), Bell et al. (2007), and Dahlen et al. (2007). The corresponding gas and stellar fractions are shown in Figure 4.2. Note that the amount of cold gas available for cluster formation is not strongly affected by this change (compare Figs. 4.1 and 4.2).

4.2.2 Rate of Cluster Formation

Having fixed the parametrization of the available cold gas, we then relate the gas mass of a protogalaxy to the combined mass of all globular clusters it can form within $\sim 10^8$ yr (the timescale of the simulation output). We based it on the rate derived in Kravtsov & Gnedin (2005):

$$M_{GC} = 3 \times 10^6 \, \mathrm{M}_\odot \, (1 + p_2) \, \frac{M_g / f_b}{10^{11} \, \mathrm{M}_\odot}. \tag{4.10}$$

An additional factor, $1 + p_2$, allows us to boost the rate of cluster formation. Such a boost may be needed because we form new clusters only at arbitrarily chosen epochs corresponding to the simulation outputs. Unresolved mergers between the outputs may require $p_2 > 0$. In our model we find the best fit to the Galactic metallicity distribution for $p_2 \sim 3$ (see Table 4.1).

Note also that equation (4.10) imposes the minimum mass of a halo capable of forming a globular cluster. Based on dynamical disruption arguments (Section 4.3) we track only clusters more massive than $M_{\min} = 10^5 \, \mathrm{M}_\odot$. Since we always have $M_g < f_b M_h$, in order to form even a single cluster with the minimum mass, the halo needs to be more massive than $10^9 \, \mathrm{M}_\odot$. For gas-rich systems at high redshift, $M_{GC} \sim 10^{-4} \, M_g / f_b \sim 10^{-4} \, M_h$.

Given the combined mass of all clusters to be formed in an event, M_{GC}, our procedure for assigning masses to individual clusters is as follows. We first draw the most massive cluster, which we call the nuclear star cluster, even though we do not have or use the information about its actual location within the host galaxy and it is not important for our current study. The mass assigned to the nuclear cluster, $M_{\max}$, is derived from the assumed initial cluster mass function, $dN/dM = M_0 M^{-2}$:

$$1 = \int_{M_{\max}}^{\infty} \frac{dN}{dM} dM, \tag{4.11}$$

which gives $M_{\max} = M_0$. This normalization is constrained by the integral cluster mass:

$$M_{GC} = \int_{M_{\min}}^{M_{\max}} M \frac{dN}{dM} dM = M_{\max} \ln \frac{M_{\max}}{M_{\min}}. \tag{4.12}$$

The power-law initial mass function agrees both with the observations of young star clus-

ters and the hydrodynamic simulations. After the nuclear cluster is drawn, the masses of smaller clusters are selected by drawing a random number $0 < r < 1$ and inverting the cumulative distribution: $r = N(< M)/N(< M_{\rm max})$:

$$M = \frac{M_{\rm min}}{1 - r\,(1 - M_{\rm min}/M_{\rm max})}. \tag{4.13}$$

We continue generating clusters until the sum of their masses reaches M_{GC}.

The formation of clusters is triggered by a gas-rich major merger of galaxies, which includes mergers of satellite halos onto the main halo as well as satellite-satellite mergers. New clusters form when the halo mass at i-th simulation output exceeds the mass at the previous output by a certain factor, and at the same time the cold gas fraction exceeds a threshold value:

$$\texttt{case}-1: \quad M_{h,i} > (1+p_3)\, M_{h,i-1} \quad \text{and} \quad f_g > p_4. \tag{4.14}$$

Also, we require that the maximum circular velocity does not decrease in this time step, to ensure that the mass increase was real rather than a problem with halo identification. We have experimented with a more relaxed criterion for the main halo than for satellite halos, with $p_{3,\rm main} < p_{3,\rm sat}$, but did not find a significantly better fit to the mass or metallicity distributions. We therefore keep a single value of p_3 for all halos.

For some model realizations, we consider an optional alternative channel for cluster formation without a detected merger, if the cold gas fraction is very high:

$$\texttt{case}-2: \quad f_g > p_5, \tag{4.15}$$

where the threshold p_5 is expected to be close to 100%. This channel allows continuous cluster formation at high redshift when the galaxies are extremely gas-rich. High-redshift galaxies are probably in a continuous state of major and minor merging, but because of their lower masses it is more difficult to detect such mergers in the simulation. Additional motivation for this channel follows from some nearby starburst galaxies that are forming young massive clusters despite appearing isolated. `Case-2` formation is allowed only for

Table 4.1. FIDUCIAL MODEL PARAMETERS

Parameter	Value	Effect
$\sigma_{\rm met}$	0.1	log-normal dispersion of mass-metallicity relation
p_2	3.0	boost of the rate of cluster formation
p_3	0.2	minimum merger ratio
p_4	0.04	minimum cold gas fraction for `case-1` formation
p_5	0.98	minimum cold gas fraction for `case-2` formation

isolated halos before they are accreted into larger systems and become satellites. The epoch of accretion is defined by the last timestep before the orbit of the subhalo falls permanently within the virial radius of its host.

Our model sample combines clusters formed in the main halo and in its satellites, either surviving or disrupted. We exclude clusters from the satellites that have a galactocentric distance at $z = 0$ greater than 150 kpc, which is the largest distance of a Galactic globular cluster. We apply the criteria for cluster formation at every timestep of the simulation (every $\sim 10^8$ yr) for each of the three main halos and their satellite populations. The rate of cluster formation per every merger event is therefore approximately $M_{GC}/10^8$ yr.

In order to compare the distribution of clusters obtained from our analysis to the distribution of Galactic globular clusters, we normalize the total number of model clusters by the ratio of the Galaxy mass to the simulated halo masses at $z = 0$:

$$N_{\rm normalized} = N_{\rm model}\,\frac{M_{MW}}{M_{h1}+M_{h2}+M_{h3}}. \tag{4.16}$$

We take $M_{MW} = 10^{12}\,{\rm M}_\odot$ and use $M_{h1} = 2.37\times10^{12}\,{\rm M}_\odot$, $M_{h2} = 1.77\times10^{12}\,{\rm M}_\odot$, and $M_{h3} = 1.70\times10^{12}\,{\rm M}_\odot$ from Kravtsov et al. (2004).

4.2.3 Metallicity

The iron abundance is assigned to each model cluster according to the estimated average metallicity of its host galaxy. The latter we obtain from the mass-metallicity relation for dwarf galaxies of the Local Group at $z = 0$ as formulated by Woo et al. (2008):

$$[{\rm Fe/H}]_0 = -1.8 + 0.4\log\left(\frac{M_*}{10^6\,{\rm M}_\odot}\right). \tag{4.17}$$

In fact, the same fit is valid for the smallest, ultrafaint dwarfs studied by Kirby et al. (2008). Thus we apply this relation to all protogalactic systems in our simulation volume.

We also include the evolution of this relation with cosmic time, based on the available observations of Lyman-break galaxies at $z \approx 2$ (Erb et al., 2006), Gemini Deep Deep Survey galaxies at $z \approx 1$ (Savaglio et al., 2005), and cosmological hydrodynamic simulations that provide the average metallicity of galaxies (Brooks et al., 2007; Davé et al., 2007):

$$[\mathrm{Fe/H}](t) \approx [\mathrm{Fe/H}]_0 - 0.03 \left(\frac{t_0 - t}{10^9 \text{ yr}} \right). \tag{4.18}$$

While this temporal evolution is probably real, it can change the metallicity for the same stellar mass by at most 0.36 dex in 12 Gyr. This amount is smaller than the 0.4 dex change of [Fe/H] due to the stellar mass variation by a factor of 10. In our model, globular clusters form in protogalaxies with a range of stellar masses of several orders of magnitude (see Fig. 4.9 below).

The observed mass-metallicity relation for a large sample of galaxies observed by the SDSS has an intrinsic scatter of at least 0.1 dex (e.g., Tremonti et al., 2004). We account for it, as well as for possible observational errors, by adding a Gaussian scatter to our calculated [Fe/H] abundances with a standard deviation of $\sigma_{\mathrm{met}} = 0.1$ dex. The exact value of this dispersion is not important and can go up to 0.2 dex without affecting the results significantly.

Using equations (4.17) and (4.18) along with the procedures of Section 4.2.2, we can generate a population of star clusters with the corresponding masses and metallicities. The model contains two random factors: the scatter of metallicity and the individual cluster masses assigned via equation (4.13). We sample these random factors by creating 11 realizations of the model with different random seeds. Each realization combines clusters in all three main halos. Taking into account that the halos are about twice as massive as the Milky Way, the expected number of clusters in each model realization is $\sim 150\,(\text{the observed number}) \times (2.37 + 1.77 + 1.7) \approx 870$. The total set of all 11 realizations includes ~ 9500 clusters. For the purpose of conducting statistical tests on the distributions of cluster mass and metallicity, we consider each realization separately and then take the

Table 4.2. SUMMARY OF MODEL EQUATIONS

Equation	Section	Description
4.2	4.2.1	fraction of baryons accreted onto a halo
4.3	4.2.1	cutoff mass for baryon accretion
4.5	4.2.1	cold gas mass relative to stellar mass
4.7	4.2.1	stellar mass relative to halo mass
4.10	4.2.2	mass in globular clusters relative to gas mass
4.17	4.2.3	stellar mass-metallicity relation for halos
4.24	4.3	evolution of cluster mass

median value of the calculated statistic.

For convenience, we provide a list of the most important equations we used in the model in table 4.2.

4.3 Dynamical Disruption

Star clusters are prone to gradual loss of stars, and in some cases, total disruption by internal and external processes. It is expected that the mass function of globular clusters has evolved through cosmic time, from an initial (probably, a power law) distribution to the approximately log-normal distribution that is observed today. Since the main focus of this paper is on the observable properties of the Galactic population, we evolve all of our model clusters dynamically from their time of formation until the present epoch. We adopt the evaporation via two-body relaxation and stellar evolution as the mechanisms for mass loss. Tidal shocks are ignored for simplicity. Cluster mass changes because of the decrease of the number of stars, $N_*(t)$, by evaporation and the decrease of the average stellar mass, $\overline{m}(t)$, by stellar evolution:

$$\frac{1}{M}\frac{dM}{dt} \equiv \frac{1}{N_*}\frac{dN_*}{dt} + \frac{1}{\overline{m}}\frac{d\overline{m}}{dt} = -\nu_{\rm ev}(M) - \nu_{\rm se}(t)\frac{\overline{m}(0)}{\overline{m}}. \tag{4.19}$$

We have assumed, as done in the recent literature, that the evaporation rate depends only on cluster mass. The time t for each cluster is measured from the moment of its formation.

We adopt the calculation of Prieto & Gnedin (2008) for the time-dependent mass-loss rate due to stellar evolution, $\nu_{\rm se}(t)$ (see their Fig. 7). That calculation uses the relation

between star's initial mass and remnant mass from Chernoff & Weinberg (1990) and the main-sequence lifetimes from Hurley et al. (2000). Over time, stellar evolution reduces the cluster mass by up to 40%, for a Kroupa (2001) IMF. This implies that no clusters are disrupted by stellar evolution alone, and the net effect is only a shift in the mass distribution towards the lower end.

We now need to derive the evaporation rate, $\nu_{\rm ev}(M)$, as a function only of cluster mass. We begin by writing down the standard approximation (Spitzer, 1987) using the half-mass relaxation time, t_{rh}:

$$\nu_{\rm ev} = \frac{\xi_e}{t_{rh}} = \frac{7.25\,\xi_e\,\overline{m}\,G^{1/2}\,\ln\Lambda}{M^{1/2}\,R_{\rm h}^{3/2}}, \tag{4.20}$$

where ξ_e is the fraction of stars that escape per relaxation time, $R_{\rm h}$ is the half-mass radius, and $\ln\Lambda$ is the Coulomb logarithm. We take $\overline{m} = 0.87\,{\rm M}_\odot$ for a Kroupa IMF, and $\ln\Lambda = 12$, which is a common value used for globular clusters (Spitzer, 1987).

We then assume that at the time of formation $R_{\rm h}$ depends only on cluster mass, as $R_{\rm h} \propto M^{\delta_0}$, and not on the position in the host galaxy. As a fiducial model, we use a constant density model where $\delta_0 = 1/3$ (Kravtsov & Gnedin, 2005; Prieto & Gnedin, 2008). The relation for the initial size is normalized with respect to the median observed mass of Galactic clusters, $2\times10^5\,{\rm M}_\odot$, and their median size of 2.4 pc:

$$R_{\rm h}(0) = 2.4\ {\rm pc}\left(\frac{M(0)}{2\times10^5\ {\rm M}_\odot}\right)^{\delta_0}. \tag{4.21}$$

A similar relation extends to other dynamically hot stellar systems: nuclear star clusters and ultracompact dwarf galaxies (Kissler-Patig et al., 2006). The mass-size relation may change over the course of the cluster evolution. We consider a power-law relation with a potentially different slope, so that the half-mass radius responds to changes in the cluster mass as

$$\frac{R_{\rm h}(t)}{R_{\rm h}(0)} = \left(\frac{M(t)}{M(0)}\right)^{\delta}. \tag{4.22}$$

Our preferred value is again $\delta = 1/3$, but we also discuss results for other choices of δ_0 and δ. Recent N-body models of cluster disruption are consistent with $\delta \approx 1/3$ (Trenti et al., 2007; Hurley et al., 2008). Note that cluster sizes are only used as an intermediate step

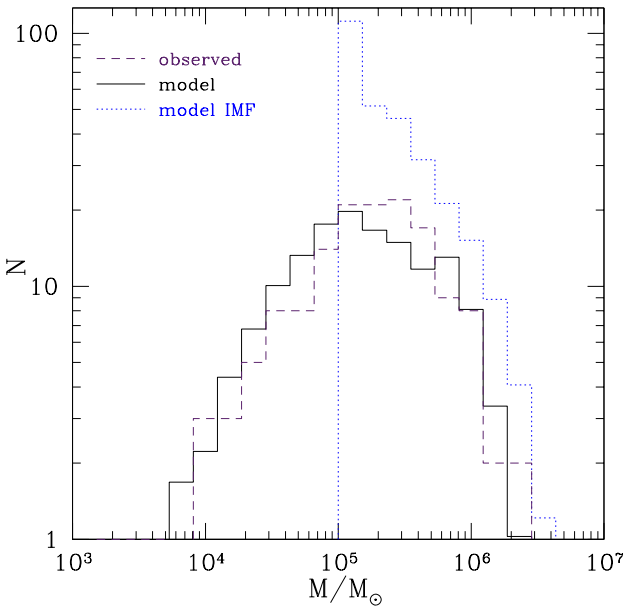

Figure 4.3. Dynamically evolved clusters at $z = 0$ in the fiducial model with $\xi_e = 0.033$, $\delta = \delta_0 = 1/3$ (*solid histogram*), compared to the observed distribution of Galactic globular clusters (*dashed histogram*). Dotted histogram shows the combined initial masses of model clusters formed at all epochs, including those that did not survive until the present. In the model we do not follow clusters with the initial masses below $10^5\,\mathrm{M}_\odot$.

in the derivation of $\nu_{\rm ev}(M)$ and can be subsequently ignored. The evaporation time thus becomes

$$\nu_{\rm ev}^{-1} \approx 10^{10}\,\mathrm{yr}\left(\frac{\xi_e}{0.033}\right)^{-1}\left(\frac{M(0)}{2\times10^5\,\mathrm{M}_\odot}\right)^{\frac{1+3\delta_0}{2}}\left(\frac{M(t)}{M(0)}\right)^{\frac{1+3\delta}{2}}. \qquad (4.23)$$

The fraction ξ_e is not well constrained. The lower limit on ξ_e is achieved in isolated clusters, for which $\xi_e = 0.0074$ (Ambartsumian, 1938; Spitzer, 1940). Tidally-truncated clusters lose stars at a faster rate, as first calculated by Hénon (1961) and Spitzer & Chevalier (1973). Using orbit-averaged Fokker-Planck models of cluster evolution, Gnedin et al. (1999) found ξ_e varying between 0.02 and 0.08 depending on time and cluster concentration (their Fig. 4 and Table 2). More recently, realistic direct N-body models became possible (e.g., Baumgardt, 2001; Baumgardt & Makino, 2003). These calculations revealed that the gradual escape of stars through the tidal boundary, which is not spherical as in the Fokker-Planck calculations, breaks the linear scaling of the disruption time with the relaxation time. Baumgardt (2001) suggested that the evaporation time scales as $\nu_{\rm ev}^{-1} \propto t_{rh}^{3/4}$. Gieles & Baumgardt (2008) verified this relation and found almost no dependence on the cluster

half-mass radius. Instead, they proposed an explicit dependence on the Galactocentric distance R_G and velocity V_G, to reflect the strength of the local tidal field: $\nu_{\rm ev}^{-1} \propto \omega^{-1} \equiv R_G/V_G$. This gives $\nu_{\rm ev}^{-1} \propto M^{3/4}\omega^{-1}$. Their formula is similar to the empirical estimates of the disruption time by Lamers et al. (2005): $\nu_{\rm ev}^{-1} \propto M^{0.65}$.

Since the calculation of the local tidal field is currently beyond our simple model, we ignore the dependence on the Galactocentric distance but argue that we can incorporate the result of Gieles & Baumgardt (2008) for the disruption timescale by using a lower value of $\delta_0 = \delta = 1/9$. With this choice of the exponents, our equation (4.23) gives $\nu_{\rm ev}^{-1} \propto M^{2/3}$. We discuss these alternative models in Section 4.4.6.

For consistency with Prieto & Gnedin (2008), we adopt $\xi_e = 0.033$ for the fiducial model.

With the above ingredients, we can now compute the cluster mass at time t after formation by inverting $\nu_{\rm ev}(M)$ in equation (4.19) and assuming that most of the stellar evolution mass loss happens much faster than the evaporation:

$$M(t) = M(0)\ \left[1 - \int_0^t \nu_{\rm se}(t')\,dt'\right]\ \left[1 - \frac{1+3\delta}{2}\nu_{\rm ev,0}\,t\right]^{2/(1+3\delta)}, \tag{4.24}$$

where $\nu_{\rm ev,0} \equiv \nu_{\rm ev}(M = M(0))$.

The initial mass function of globular clusters is evolved from the time of formation until the present epoch and is shown in Figure 4.3 for the fiducial model. The observed mass function in the Milky Way is well represented by a log-normal distribution. We derive the masses of the Galactic clusters by taking their absolute V-band magnitudes from the Harris (1996) catalog and assuming a constant mass-to-light ratio $M/L_V = 3\,{\rm M}_\odot/{\rm L}_\odot$. The functional form of a Gaussian built around $\log M$ for the observed sample is given by

$$\frac{dN}{d\log M} = \frac{1}{\sqrt{2\pi}\sigma_M}\exp\left[-\frac{(\log M - \overline{\log M})^2}{2\sigma_M^2}\right], \tag{4.25}$$

with the mean $\overline{\log M} = 5.22$ and standard deviation $\sigma_M = 0.61$, in solar masses. The predicted mass function in the fiducial model with $\xi_e = 0.033$ and $\delta = \delta_0 = 1/3$ is consistent with the observations. The Kolmogorov-Smirnov (KS) test probability of the two mass

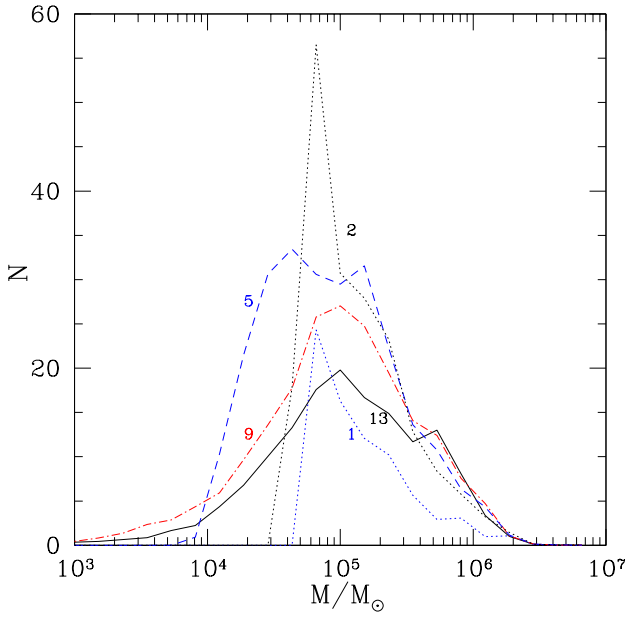

Figure 4.4. The mass function of clusters in the fiducial model at different epochs corresponding to the cosmic times of 1 Gyr ($z \approx 5.7$, *dotted*), 2 Gyr ($z \approx 3.2$, *dotted*), 5 Gyr ($z \approx 1.3$, *dashed*), 9 Gyr ($z \approx 0.5$, *dot-dashed*), and 13.5 Gyr ($z = 0$, *solid*).

functions being drawn from the same distribution is $P_{KS,M} = 7.4\%$. This value is the median of the KS probabilities for the 11 random realizations of the model. The model distribution is also well fit by a Gaussian, with $\overline{\log M} = 5.14$ and $\sigma_M = 0.65$.

The mean of the model distribution is slightly lower than observed, implying that the disruption process needs to be stronger to fully reconcile with the data. Clusters that start out with low mass but are not disrupted effectively over their lifetime over-populate the low end of the present-day model mass function. Old and intermediate-age clusters that started with initial mass $5 < \log M < 5.4$ and survived until the present era appear to be the main cause of this discrepancy.

Figure 4.4 illustrates the evolution of the mass function over cosmic time as an interplay between the continuous buildup of massive clusters ($M > 10^5\,\mathrm{M}_\odot$) and the dynamical erosion of low-mass clusters ($M < 10^5\,\mathrm{M}_\odot$). Since we do not track the formation of clusters below $M_{\rm min}$, the low end of the mass function was built by a gradual evaporation of more massive clusters. The strongest bout of cluster formation happens between a cosmic time of 1 and 5 Gyr, and a peak of the mass function forms at $M \sim 3 \times 10^4\,\mathrm{M}_\odot$. The peak moves to larger masses, $\sim 10^5\,\mathrm{M}_\odot$ by $t = 9$ Gyr, while a power-law tail develops at low masses. A

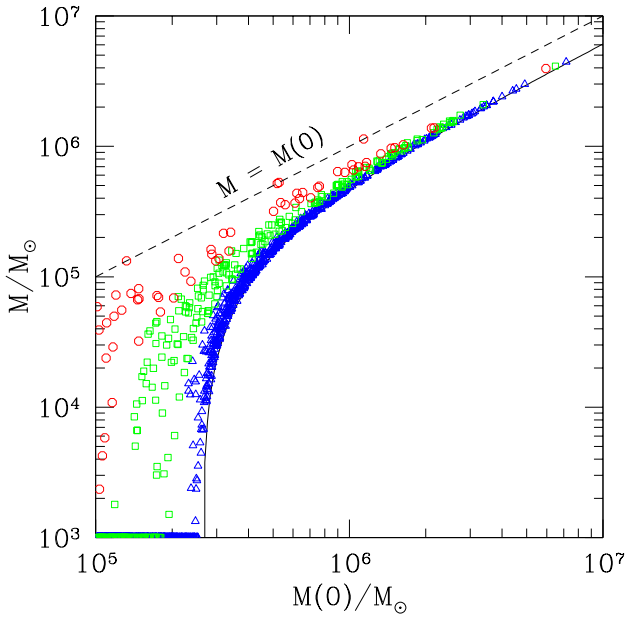

Figure 4.5. Final mass of model clusters vs. their initial mass, for a single realization of the fiducial model. Clusters are divided into three age groups: *triangles* represent old clusters (age > 10 Gyr), *squares* represent intermediate age clusters (5 Gyr < age < 10 Gyr), and *circles* represent young clusters (age < 5 Gyr). All disrupted clusters are placed at the bottom of the plot, to illustrate the range of their initial mass. The birthline of clusters, $M = M(0)$, is plotted as a dashed line for reference.

significant fraction of low-mass clusters is disrupted between 9 Gyr and the present, as few new clusters are produced.

The relation between the cluster initial and final masses is shown in Figure 4.5. Old clusters that have undergone significant amounts of dynamical and stellar evolution form a tight sequence on this plot. The lower boundary with a dense concentration of points corresponds to the expression $M = 0.63\,[M(0) - 2.6 \times 10^5\ \mathrm{M}_\odot]$, which reflects 13 Gyr of stellar and dynamical evolution according to equation (4.24) with the fiducial values of the parameters. Thus an old cluster must have an initial mass of at least $2.6 \times 10^5\,\mathrm{M}_\odot$ to survive dynamical disruption. Clusters in the younger age groups fill the space between their birthline and this boundary. The youngest clusters have the shallowest slope at low mass, as few of them have had enough time to undergo significant disruption. The mean final mass for all three age groups is about the same, implying that some of the oldest globular clusters could have been more massive at the time of their formation than clusters that have formed recently in the local universe.

Fall & Zhang (2001) suggested that a low-mass end of the mass function should ap-

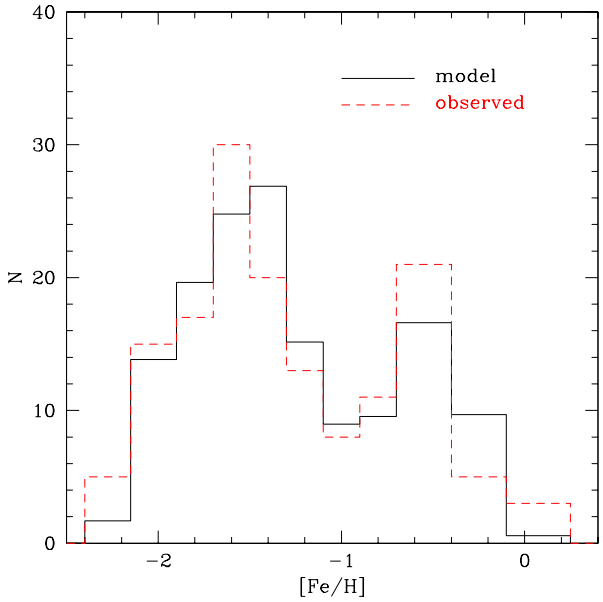

Figure 4.6. Metallicities of model clusters formed at all epochs that have survived dynamical disruption by $z = 0$ in the fiducial model (*solid histogram*), compared to the observed distribution of Galactic globular clusters (*dashed histogram*).

proach $dN/dM \approx const$ as a result of dominant disruption by two-body evaporation. Our mass function in the range $3.5 < \log(M/\mathrm{M}_\odot) < 5.0$ is consistent with a power law $\log(dN/d\log M) = 0.89\log M - 3.04$, or $dN/dM \propto M^{-0.11}$, in good agreement with the expectation.

4.4 Results

4.4.1 Exploration of the Parameter Space

Overall, our model has five adjustable parameters (Table 4.1). To explore possible degeneracies among these parameters, and to find the parameter set that produces the best-fitting metallicity distribution, we set up a grid of models in which each of the parameters was varied within a finite range of values. The range was taken to be large enough to explore all physically relevant values of each parameter.

The boost for cluster formation, p_2, varied from 0 to 5. For consistency with the rate derived in the hydrodynamic simulation of Kravtsov & Gnedin (2005), we aimed to keep this parameter at low values.

The minimum mass ratio for mergers, p_3, varied between 0.15 and 0.5. It is consistent

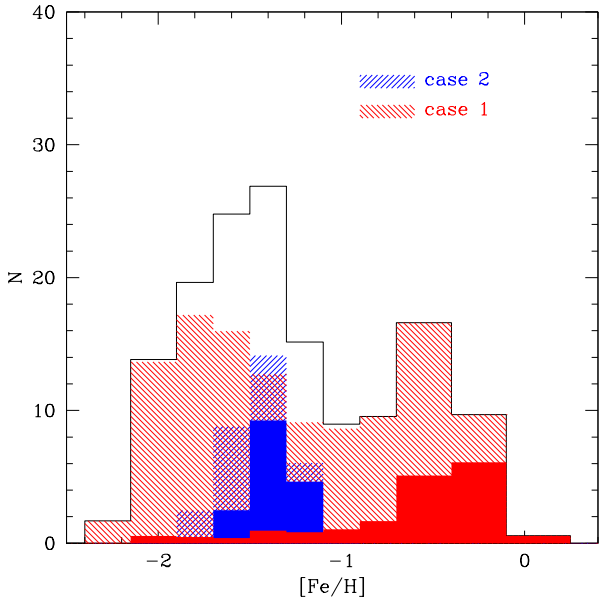

Figure 4.7. Metallicity distribution in the fiducial model, split by the formation criterion: major mergers (`case-1`) and early mergers (`case-2`). Solid histograms show the clusters formed in the main Galactic halo.

with typical major merger criteria used in the literature (e.g. Beasley et al. 2002 use $p_3 = 0.3$).

The cold gas fraction required for cluster formation during a merger, p_4, could be relatively low but non-zero, so that we considered $0 < p_4 < 0.2$. This threshold parameter accounts for why disk galaxies like the Milky Way are still forming stars despite a low gas fraction, while ellipticals are not.

The gas fraction for `case-2`, p_5, has to be very high – above 90%, as our prescription predicts that many halos have a very high gas fraction at high redshift and could overproduce blue globular clusters (as was the case in the Beasley et al. 2002 model).

We considered several values for $\sigma_{\rm met}$ but found that a value of 0.2 or higher smeared out the peaks in the metallicity distribution, while a value of 0 failed to fill the extreme ends of the distribution. We therefore include only three values in our search, $\sigma_{\rm met} = 0, 0.1, 0.2$.

We find the best-fit model by searching through the multi-parameter space and maximizing the KS probabilities of the metallicity distribution, $P_{KS,Z}$, and the mass function, $P_{KS,M}$, being consistent with observations. The likelihood function also contains additional

factors that force the parameters towards the values that we consider ideal. We require the model to produce the observed number of clusters, $N \approx 150$, scaled by the host galaxy mass as in equation (4.16). We wish to maximize the fraction of clusters formed in the main disk, $f_{\rm disk}$, to be consistent with the observed spatial distribution (Section 4.5). We penalize the likelihood function for large values of p_2 and for any young clusters formed after t = 10 Gyr, $N_{\rm after10}$. We also wish to minimize the fraction of clusters formed through the `case-2` channel, $f_{\rm case2}$, for simplicity of the model. Finally, we want to increase the likelihood of the metallicity distribution being bimodal, as characterized by the Dip test, $P_{\rm dip}$, which we discuss later in Section 4.4.4. The actual likelihood function that we maximize is given by

$$\begin{aligned} \log \mathcal{L} \;\; = \;\; & \log P_{KS,Z} + 0.3 \log P_{KS,M} - [(N-150)/30]^2 + \log f_{\rm disk} \\ & -0.15 p_2 - 20 N_{\rm after10}/N - 0.4 f_{\rm case2} + 3 \log P_{\rm dip}. \end{aligned} \tag{4.26}$$

The coefficients for each term were adjusted heuristically until we found that their relative weights matched our expectation to select acceptable distributions. The "best-fit" distribution that maximizes $\mathcal{L}$ is therefore a subjective fiducial model that we use to illustrate how the bimodality may arise. We then look at how many model realizations are similar to the "best-fit" for other possible values of the parameters.

4.4.2 Age and Metallicity Distributions

Figure 4.6 shows the predicted best-fit metallicity distribution of model clusters and the observed distribution of Galactic globular clusters, both metal-poor and metal-rich. Note that we require our model to have the same formation criteria for both cluster populations; we do not explicitly differentiate between the two modes. The only variable is the gradually changing amount of cold gas available for star formation. Yet, the model predicts two peaks of the metallicity distribution, centered on [Fe/H] = –1.54 and [Fe/H] = –0.58, in remarkable agreement with the observations. The standard deviation of the red peak is 0.24 dex and of the blue peak is 0.32 dex.

The probability of KS test of the model and data samples being drawn from the same distribution is $P_{KS,Z}$ = 80%, that is, they are fully consistent with each other. The number

of surviving clusters is $N = 147$, also matching the observations. Even though our current model is extremely simple, this bimodality is reproduced naturally, without explicit assumptions about truncation of the production of metal-poor clusters at some early epoch or about the formation of metal-rich clusters in a merger of two spiral galaxies.

We find that the main halo contributes more significantly to the red peak than it does to the blue peak (Figure 4.7). In particular, clusters with the highest [Fe/H] appear to have been formed primarily by late merging into the main halo.

The fraction of clusters formed via `case-2` channel is $f_{\rm case2} = 22\%$. These clusters produce a single-peaked distribution of blue clusters. In contrast, clusters formed in major mergers contribute to both red and blue modes, in about equal proportions. We return to this point in the discussion of globular cluster systems of elliptical galaxies in Section 4.7.

Clusters that formed after $z = 2$ constitute the bulk of the red peak and contribute little to the blue peak in the metallicity distribution (Figure 4.8). The strength of this result implies that the gas reservoir and the rate of hierarchical merging at intermediate redshifts is conducive to the creation of red clusters. This result lends itself well to the idea that the simulation of Kravtsov & Gnedin (2005) was only able to reproduce the metal-poor population of globular clusters because the simulation was stopped at $z \approx 3$.

Our prescription links cluster metallicity to the average galaxy metallicity in a one-to-one relation, albeit with random scatter. Since the average galaxy metallicity grows monotonically with time, clusters forming later have on the average higher metallicity. The model thus encodes an age-metallicity relation, in the sense that metal-rich clusters are younger by several Gyr than their metal-poor counterparts. This relation is required in the model to reproduce the observed metallicity distribution, because very old galaxies cannot produce high enough metallicities. However, Figure 4.8 shows that clusters of the same age may differ in metallicity by as much as a factor of 10, as they formed in the progenitors of different mass.

Available observations of the Galactic globular clusters do not show a clear age-metallicity relation, but instead indicate an age spread increasing with metallicity (De Angeli et al., 2005; Marín-Franch et al., 2009; Dotter et al., 2010; Forbes & Bridges, 2010). Red clusters have younger mean age overall and may be as young as $\tau \approx 7$ Gyr. Our model does

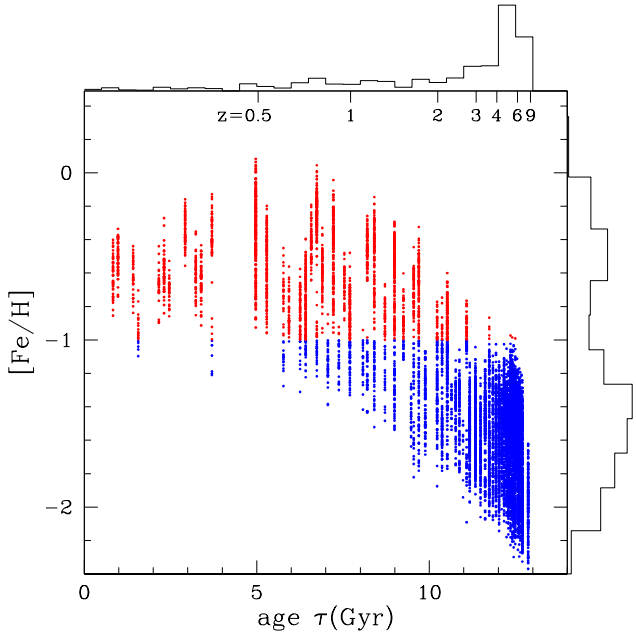

Figure 4.8. Age-metallicity relation in all 11 realizations of the fiducial model ($\approx$ 9500 clusters). The build-up of massive halos drives the steep slope of this relation at early epochs. Outer histograms show marginalized distributions on linear scale. Notice an order-of-magnitude spread in metallicities of clusters forming at a given epoch.

not appear to be in an obvious conflict with this trend. We define cluster age as $\tau \equiv t_0 - t_f$, where t_f is the time of formation. We find the mean age of 11.7 Gyr for the blue population and 6.4 Gyr for the red population, with the standard deviation of 1.3 Gyr and 2.7 Gyr, respectively. More accurate dating of the Galactic and extragalactic clusters is needed to falsify the predicted age-metallicity trend.

Distributions of the cluster formation time and environment in the fiducial model are shown in Figure 4.9. The age distribution, which peaks strongly between 11 and 13 Gyr, demonstrates that the majority of our clusters is still very old and falls in line with the observed perception of globular clusters. However, the distribution of formation redshift appears remarkably flat in the range $1 < z < 7$, emphasizing that the clusters were not formed in a single event but rather through the continuous process of galaxy formation. Few clusters were formed prior to the era of reionization, as sufficiently large quantities of gas could not be condensed to meet the mass threshold for cluster formation at redshifts $z > 9$. The distributions of the total and stellar mass of the host galaxies extend over three orders of magnitude. Their extended high-mass tails contribute to the strength of the red

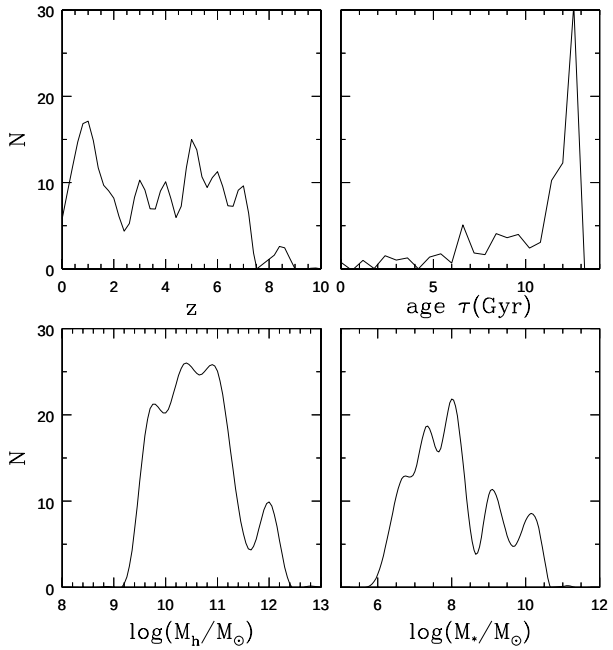

Figure 4.9. Number of clusters in the fiducial model as a function of the environment: redshift of formation (*top left panel*), present age (*top right panel*), host halo mass at the time of formation (*bottom left panel*), and host stellar mass (*bottom right panel*).

peak, as the most massive halos would form most metal-rich clusters.

Globular clusters form much earlier than the majority of field stars. Figure 4.10 shows the fraction of galaxy stellar mass locked in massive star clusters, normalized for convenience as $10^3 M_{GC}/M_*$. To calculate this ratio, we summed over all protogalactic systems that would end up within 150 kpc of the galaxy center at $z = 0$, regardless of their location at earlier times. Thus it represents a global cluster formation efficiency in a Milky Way-sized environment. Specific realizations of the model differ in detail in the three host halos, by as much as a factor of 2. This scatter is shown by the shaded region on the plot. The globular cluster mass includes their continuous formation and the mass loss due to the dynamical evolution. A striking prediction of the model is a very high cluster fraction at early times, near $t = 1$ Gyr, of $M_{GC}/M_* \approx 10-20\%$. Star cluster production may have been a dominant component of galactic star formation at $z > 3$. By $t = 3$ Gyr ($z \approx 2$), the cluster fraction drops to only a few percent, as expected for a galaxy undergoing active star formation. At the current epoch, massive star clusters make up less than 0.1% of the stellar mass. The predicted ratio is progressively more uncertain at higher redshift because it relies on our

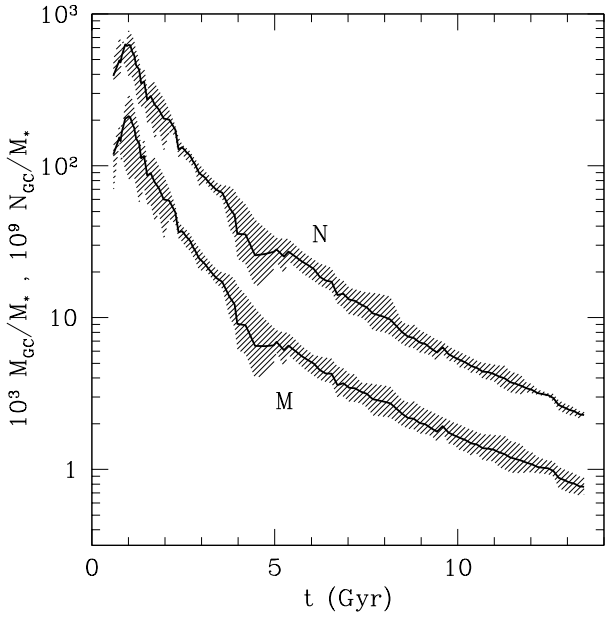

Figure 4.10. Ratios of total cluster mass and number to the galaxy stellar mass, summed over all systems that are located within 150 kpc of the center at $z = 0$. Lines represent the average over all three host halos and their corresponding subhalo populations. Shading represents the scatter in these ratios, given by the lowest and highest values among the three hosts. Massive star cluster formation was a much more dominant component of galactic star formation at early times than it has been for the last 10 Gyr.

extrapolated prescription for the galactic stellar mass. The low-redshift prediction should be robust. We also show a variant of the specific frequency parameter related to the number of clusters, $T \equiv N/(M_*/10^9\,\mathrm{M}_\odot)$, introduced by Zepf & Ashman (1993). It shows a similar decline with time, reaching $T \approx 2$ at the present.

These global cluster formation efficiencies agree with many observations across galaxy types. Rhode et al. (2005) find $T \sim 1$ for both red and blue clusters in the field and group spiral galaxies. This parameter increases with the galaxy mass. In the Virgo cluster, Peng et al. (2008) find $T \sim 5$ for galaxies in the mass range appropriate for the Milky Way. McLaughlin (1999) estimated the cluster mass fraction in both spiral and elliptical galaxies to be $M_{GC}/(M_* + M_g) \approx 0.0026 \pm 0.0005$. This is larger than what we find by a factor of several, but we count in M_* all stars out to 150 kpc, which includes some satellite galaxies as well as the host. Therefore, both predicted cluster efficiencies at $z = 0$ are reasonable. Their rise at high redshift is an interesting prediction of the model.

The model also shows that the globular cluster system overall is more metal-poor than

the stars in disrupted satellites, which are expected to form a stellar spheroid of the Galaxy. We calculated the mass-weighted metallicity of stars formed in the disrupted satellites of all three main halos (using eqs. 4.7, 4.8, 4.17, 4.18). This calculation bears all the uncertainty of our extrapolated time evolution of the stellar fraction and mass-metallicity relation, but nevertheless provides a useful estimate. We find the tail of halo star metallicities as low as the most metal-poor globular clusters, but the overall stellar distribution peaks around $[\mathrm{Fe/H}] \approx -0.3$. A very similar situation is observed in NGC 5128 and discussed by Harris (2010). In our model, majority of globular clusters form before the bulk of field stars and therefore acquire lower metallicities. For comparison, the metallicity of stars in surviving satellite galaxies peaks around $[\mathrm{Fe/H}] \approx -0.8$ and forms an intermediate population between the clusters and the field.

Despite our attempts to incorporate it as a major penalty in the likelihood statistic, we were unable to completely eliminate the phenomenon of young massive star clusters. Interestingly, these clusters did not originate in the main galactic disks. All clusters younger than 5 Gyr formed in satellite halos in the mass range $\sim 10^{10} - 10^{11}\,\mathrm{M}_{\odot}$, at distances $40 - 100$ kpc from the center. Although the proper sample of the Galactic globular clusters does not contain any young clusters, there are several young massive clusters in M31 whose ages were confirmed both from the visual and UV colors (Fusi Pecci et al., 2005; Rey et al., 2007) and from the integrated-light spectroscopy (Puzia et al., 2005). The actual analogs of young model clusters may be found in the LMC, which hosts globular clusters with a wide range of ages and continues to form clusters now. There may even exist young star clusters with masses $\sim 10^5\,\mathrm{M}_{\odot}$ in the Galactic disk, hidden behind tens of visual magnitudes of extinction but revealing themselves through free-free emission of their ionization bubbles (Murray & Rahman, 2010). Massive star cluster formation at late times thus paints a picture consistent with the idea that today's super star clusters are destined to become observationally equivalent to globular clusters, as envisioned by Ashman & Zepf (1992) and Harris & Pudritz (1994).

A separate criterion for the formation of clusters in extremely gas-rich systems (`case-2`) is not necessary for achieving a good fit to the observed metallicity distribution. Though we feel that the inclusion of `case-2` formation channel in the model is both useful and

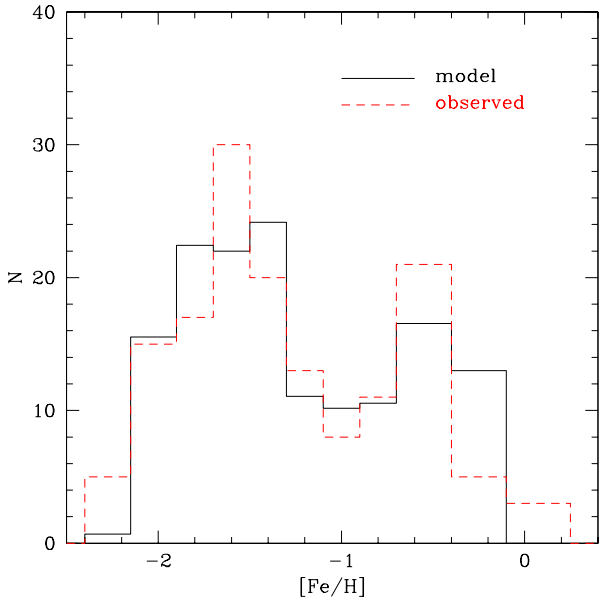

Figure 4.11. Metallicity distribution at $z = 0$ in the model without `case-2` formation (*solid histogram*), compared to the observed distribution of Galactic globular clusters (*dashed histogram*).

physically motivated, it takes away from the elegance of using only resolved mergers as a lone formation mechanism. It turns out that the main benefit of allowing clusters to form via `case-2` is seen in the mass function of surviving clusters. The high-mass end of the mass function matches the observations better if massive halos (primarily the main halo) are allowed to form as many clusters as possible at early times.

We searched the model grid without `case-2`, by setting $p_5 = 1$, and found an almost equally good metallicity distribution as in the fiducial model. Figure 4.11 shows that this distribution also appears bimodal and completely consistent with the data. The KS probability is $P_{KS,Z} = 92\%$. In fact, even the mass function is only marginally less consistent, $P_{KS,M} = 2.0\%$ vs. 7.4% in the fiducial model. The parameters used to obtain this distribution were: $p_2 = 2.85$, $p_3 = 0.16$, $p_4 = 0.04$, $p_5 = 1$, $\sigma_{\rm met} = 0.1$.

4.4.3 Sensitivity to Model Parameters

The fiducial distribution discussed above is not unique among our results in its ability to match the observations. Significant degeneracy exists among combinations of the model parameters that produce metallicity distributions consistent with the Galactic sample. Many

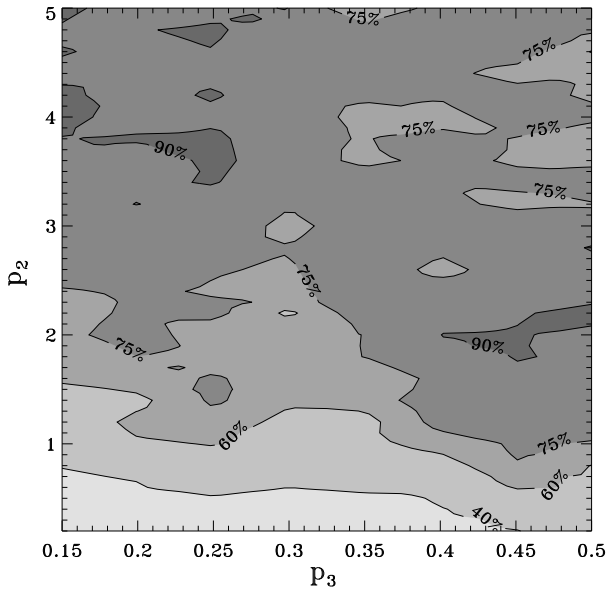

Figure 4.12. Contour plot of the KS probability for the metallicity distributions in the plane of parameters $p_2 - p_3$. Contour labels are the actual probability values, $P_{KS,Z}$. This plot shows that KS test alone cannot rule out any region of the parameter space from being statistically consistent with the data.

models within the grid have sufficiently high KS probabilities. In this section we explore which regions of the parameter space produce models similar to our best fit.

First, let us motivate the use of the likelihood function given by equation (4.26) as opposed to using a standard statistical test to select the best fit. In the early stages of development of our model, we relied on KS test alone to help us understand the range of parameters that produce metallicity distributions that match the observations. However, once the model was completed, it became apparent that KS test alone was not powerful enough for analysis of the results. This is clearly demonstrated in Figure 4.12, which shows the value of the KS probability $P_{KS,Z}$ as a function of p_2 and p_3 across their respective ranges in the grid. Each point represents the maximum possible $P_{KS,Z}$ for the given values of p_2 and p_3 with the other parameters free to vary within the grid. This is done to best represent the full extent of the 5-dimensional parameter space within a 2-dimensional slice. Statistically, any distribution with $P_{KS,Z} > 10\%$ cannot be ruled out with confidence, implying that almost the entire range of our parameters can produce statistically consistent distributions! In addition, although some regions of the parameter space have higher values of $P_{KS,Z}$ than

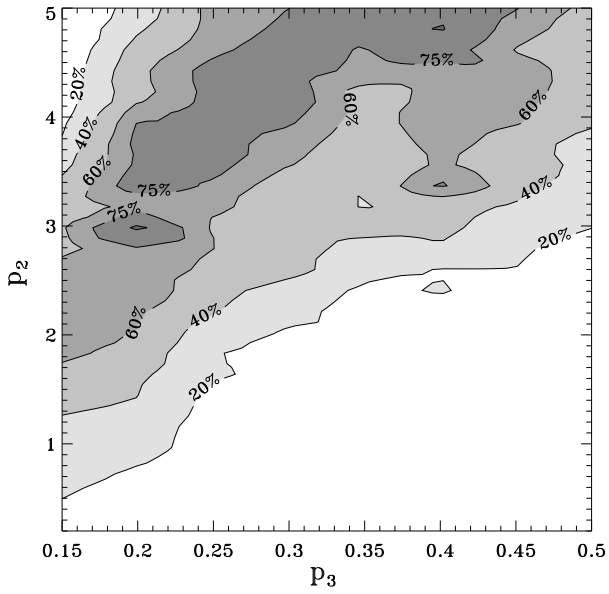

Figure 4.13. Contour plot of the likelihood statistic $\mathcal{L}$ in the plane of parameters $p_2 - p_3$. Contour labels show percentages of the maximum. The highest-value region is a degeneracy along the line $p_2 = 19p_3 - 0.91$.

others, there is no clear pattern in the contours to help us understand the required physics of star cluster formation within our semi-analytical recipe.

In comparison, Figure 4.13 shows contours of the value of the likelihood function from equation (4.26), using the same scheme described above to maximize the value at each point. The shape of these contours demonstrates that p_2 and p_3 are degenerate in their ability to produce good distributions. The degeneracy can most easily be understood by noting that these parameters directly affect the total number of clusters: p_2 controls the cluster formation rate per merger, while p_3 selects eligible mergers. It is therefore expected that the contours show a correlation at high levels of the likelihood function, as the statistic depends sensitively on the total number of clusters.

Figure 4.14 shows the same type of contours as Fig. 4.13, but only for distributions with $p_5 = 1$. Disallowing the `case-2` channel reduces the range of the parameter space where good distributions are found. In particular, compared to the previous plot, Fig. 4.14 lacks any viable models with $p_3 > 0.3$. Given the tight and steep correlation in this plot, it is likely that larger values of p_3 would require very high $p_2 > 5$, which may violate current observational constraints on the cluster formation efficiency. However, Fig. 4.11

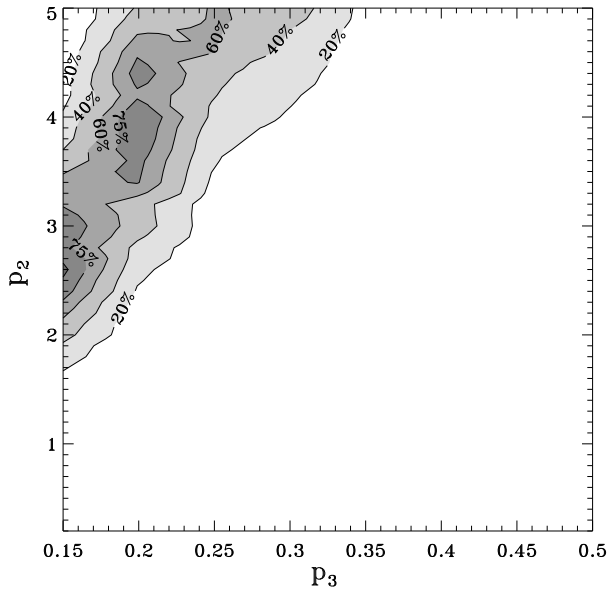

Figure 4.14. Same as Figure 4.13 but for the models without `case-2`. The highest-value region is a degeneracy along the line $p_2 = 24p_3 - 1.1$.

demonstrates that a good model can still be found with reasonably small values of p_2 and p_3, without the `case-2` channel.

To understand the sensitivity of the likelihood function to individual parameters, we also considered one-dimensional slices of the parameter space around the fiducial model, this time allowing for only one parameter to vary at a time. Figure 4.15 illustrates how the sharp peaks of $\mathcal{L}$ allow us to select the best model more accurately than on the basis of $P_{KS,Z}$ alone. Particularly as a function of p_2 and p_3, $P_{KS,Z}$ varies slowly over the entire range of the grid. On the other hand, p_4 and p_5 must stay within a small range of their fiducial values in order to achieve acceptable values of either $P_{KS,Z}$ or $\mathcal{L}$.

Figure 4.16 shows variation of the metallicity distribution when individual parameters deviate from their fiducial values. Each parameter can change the shape of the metallicity function and the number of clusters. The effects of varying p_2 and p_3 are almost opposite, reflecting the degeneracy in the likelihood contours. In particular, smaller p_3 accommodates more minor mergers, which allow massive hosts form more metal-rich clusters as well as some metal-poor clusters. Decreasing p_5 allows more clusters form through the `case-2` channel; most of such clusters are metal-poor. The major role of p_4 appears

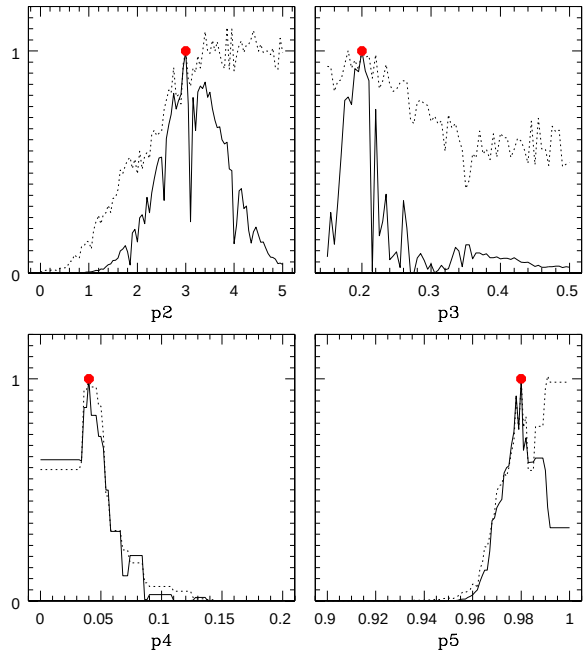

Figure 4.15. Marginalized single-parameter likelihood distributions around the fiducial model, $\mathcal{L}/\mathcal{L}_{\rm max}$ (*solid lines*). Dashed lines show the metallicity probability $P_{KS,Z}$ normalized to the fiducial model value. Compared with $P_{KS,Z}$ alone, the likelihood function $\mathcal{L}$ significantly tightens the constraints on the best values of the parameters. Filled circles show the fiducial model.

to govern the extent of the most metal-rich clusters – lower threshold gas fraction allows clusters to form in the later, more enriched environments of massive hosts.

Figure 4.17 illustrates the response of the metallicity distribution to simultaneous variations of model parameters. First, we plot two distributions where we changed p_3 to 0.15 and 0.3 while keeping the other parameters fixed. The width of the metal-poor peak broadens as p_3 is lowered, indicating that a wider range of halos in the early universe were able to produce clusters. Raising p_3 has the opposite effect. Note that the locations of the two peaks are remarkably robust to these changes. Staying at $p_3 = 0.15$, we set $p_5 = 1$ to eliminate the `case-2` channel and set $p_4 = 0$ to allow even gas-poor massive halos at low redshift to form clusters. The result (long-dashed line) is a distribution with a much broader metal-rich peak, which extends well past the maximum metallicity of the fiducial model. The dot-dashed line represents a corresponding change to $p_4 = 0.08$ and $p_5 = 0.96$ for the $p_3 = 0.3$ model. In this case, the metal-rich peak is severely depleted and remains only as an extended tail of a single-peaked, metal-poor distribution dominated by `case-2` clusters.

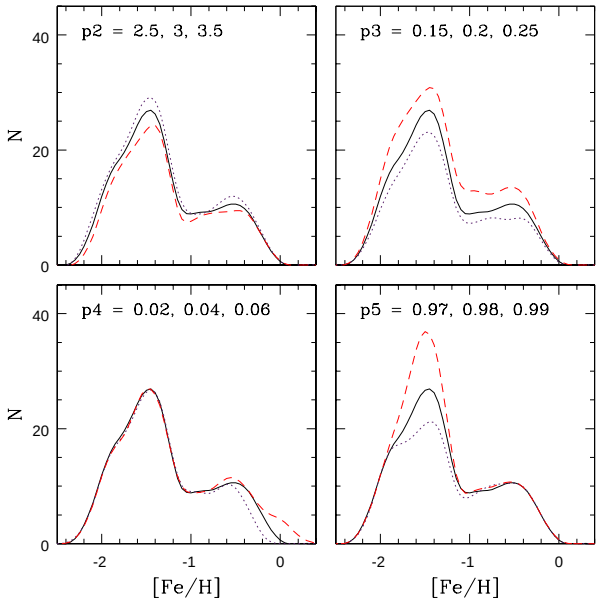

Figure 4.16. The effects of varying individual model parameters on the metallicity distribution. In each panel a single parameter is increased (*dotted line*) and decreased (*dashed line*) relative to the fiducial model (*solid line*). The parameter values are indicated inside the panels.

These distributions are just some of the realizations of our model that were rejected due to their low values of $\mathcal{L}$. All of them have features that conflict with the observed metallicity distribution in the Galaxy.

4.4.4 Origin of the Metallicity Bimodality

The KS statistic measures the overall consistency of the model and observed metallicity distributions, but not specifically bimodality or multimodality within the distributions. In order to address the particular issue of modality, we employ two additional statistical tests, described in the appendix of Muratov & Gnedin (2010).

The Gaussian Mixture Modeling test indicates that the fiducial distribution is bimodal at a high level of significance (better than 0.1%). The peak metallicities of both modes and their widths are close to the observed values and agree with them within the errors. Both samples easily appear bimodal to the eye because the modes are well separated, with the dimensionless peak separation ratio $D > 3$. However, as discussed in the appendix of Muratov & Gnedin (2010), the GMM test is sensitive to the assumption of Gaussian modes. It may indicate highly statistically significant split into two modes when the distribution is

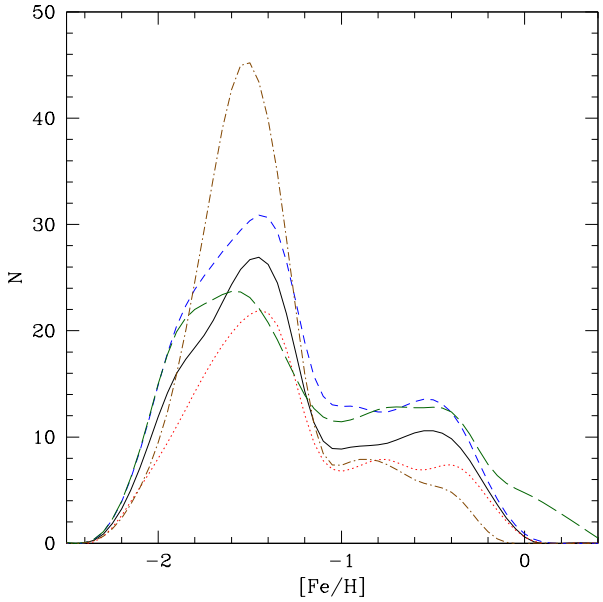

Figure 4.17. The effects of simultaneous variation of several model parameters. The fiducial model (*solid line*) is plotted alongside four distributions that illustrate other outcomes of our model. *Short-dashed* and *dotted lines* correspond to respectively lowering (to 0.15) and raising p_3 (to 0.3) away from its fiducial value (0.2). *Long-dashed line* corresponds to the model with $p_3 = 0.15$, $p_4 = 0$, $p_5 = 1$. *Dot-dashed line* corresponds to the model with $p_3 = 0.3$, $p_4 = 0.08$, $p_5 = 0.96$.

truly unimodal but skewed. For faster and more robust model selection we consider another test of multimodality.

The Dip test compares the cumulative input distribution with the best-fitting unimodal distribution. The maximum distance between the two corresponds to a dip in the differential distribution. The Dip test for the observed Galactic clusters indicates that the distribution is 90% likely to not be unimodal. When applied to our fiducial model, the Dip test implies it is 99% likely to not be unimodal. However, there is a caveat that the probability of the Dip test depends on the number of objects in the sample, similarly to KS test. The higher significance of the model result does not mean that the model is actually more bimodal than the data, because we used all 11 random realizations of the model as a combined sample to evaluate the Dip test. While this is not a fair comparison to the data, it allows us to differentiate efficiently among alternative models.

We ran the Dip test for all models on the grid in a manner similar to the likelihood statistic. The most interesting result of the Dip statistic comes from one-dimensional slices

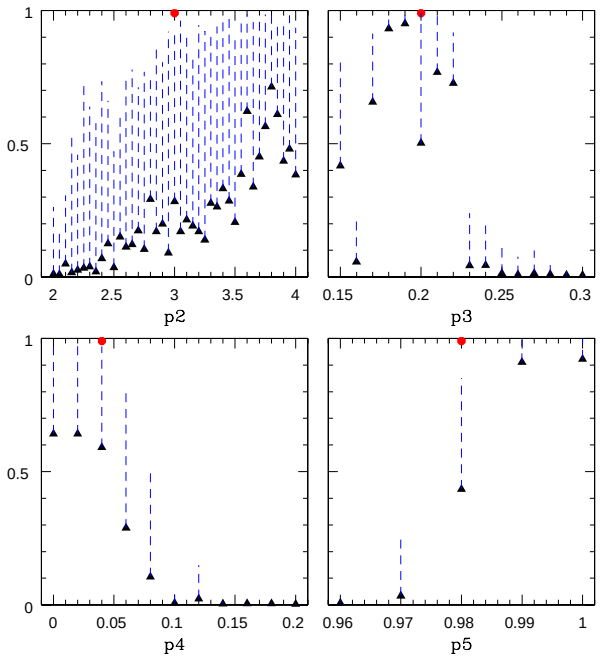

Figure 4.18. Median values of the Dip probability (*triangles*) among distributions with $140 < N < 160$, binned according to each parameter. Dashed lines extend to the 75% quartiles of $P_{\rm dip}$. The fiducial model is shown by a red dot.

of the parameter space. Considering only models with the normalized number of clusters in the range $140 < N < 160$, we binned the distributions according to the values of the four parameters and found the median and quartiles of $P_{\rm dip}$ in each bin. Figure 4.18 shows several trends. (i) Distributions with low formation rate p_2 are unlikely to be bimodal. The 75th percentile of $P_{\rm dip}$ increases systematically with p_2 in the range $2 < p_2 < 3$, but plateaus for $p_2 > 3$. (ii) The most bimodal distributions require p_3 to be small enough to allow for merger ratios 1:5 to trigger cluster formation. Between $p_3 = 0.2$ and $p_3 = 0.5$, the lower p_3 the better. However, mass ratios lower than 1:6 may dilute bimodality. (iii) The gas fraction threshold p_4 should be under 10% for ideal bimodality, to include mergers of massive galaxies. (iv) The fraction p_5 has to be close to 1, implying that `case-2` negatively affects bimodality. A conclusion from this plot is that bimodality appears in a significant number of model realizations, for a wide range of parameters. At the same time, a similarly large number of realizations are unimodal.

The metallicity distribution is bimodal if metal-rich clusters constitute a significant subset of all clusters. Thus, the fraction of red clusters, $f_{\rm red}$, is a simple proxy for bimodality.

Indeed, we find a strong correlation between $f_{\rm red}$ and $P_{\rm dip}$. The red fraction follows similar, but weaker, trends with model parameters to those shown in Fig. 4.18. The median red fraction correlates most strongly with p_5, increasing from $f_{\rm red}$ = 16% for p_5 = 0.96 to $f_{\rm red}$ = 32% for p_5 = 1. The red peak is significantly stronger without `case-2` clusters.

We note that the Dip test, unlike KS test, does not depend on comparing the model distribution to the Galactic sample. Therefore, the trends for bimodality derived from the Dip test should apply to other globular cluster systems. We anticipate that bimodality would likely arise if we applied our model with the parametric constraints stated above to any cosmological *N*-body simulation that follows the mass assembly history of a large galaxy. Further discussion of applying our model in different galactic environments follows in Section 4.7.

In order to investigate the underlying cause of bimodality, we examined various properties of merger events. A merger event is defined as any time in a halo's track when it meets the criteria for `case-1` formation. An important requirement here is the minimum mass of cold gas needed to produce a cluster that would survive dynamical disruption. Through equation (4.10), a cluster mass $M > 2 \times 10^5\,{\rm M}_\odot$ requires $M_g > 3 \times 10^8\,{\rm M}_\odot$. This constraint significantly reduces the number of eligible mergers. We considered the distributions of halo mass, lookback time, and metallicity (without additional dispersion) for all relevant merger events. We find that relatively few mergers happen in the space of high metallicity, high mass, and late time. Almost half of the mergers (44%) take place before τ = 12 Gyr, and only 24% of the mergers happen in the last 10 Gyr. If we also counted the events that led to now-disrupted clusters, these numbers would spread even further to 53% and 17%, respectively. Nevertheless, the recent mergers stand out for two reasons: each such event creates more clusters, and these younger clusters have better chance of surviving the dynamical disruption than the older clusters. Since the number of clusters formed in each merger is positively correlated with the galaxy mass, the few stochastic super-massive mergers with high metallicity are likely to produce a significant number of clusters, which would separate the red peak from the blue peak.

We also considered that cluster bimodality may be linked to the mass ratios in the merger events. "Major" and "minor" mergers have been proposed to play different roles in

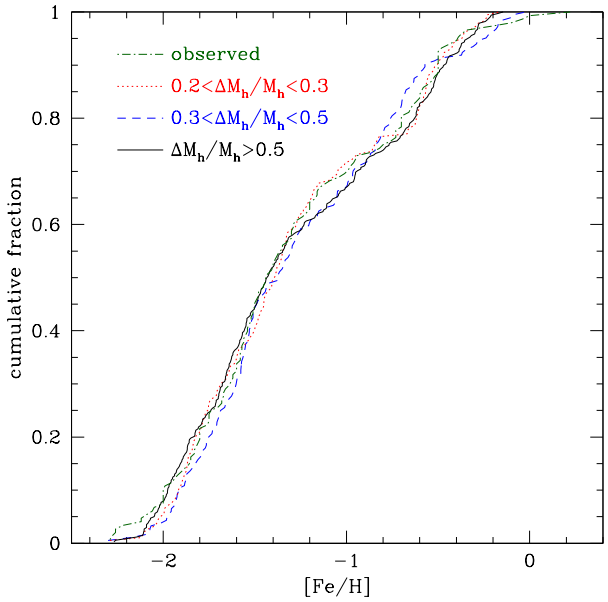

Figure 4.19. Cumulative metallicity distributions of the fiducial model clusters split by the range of merger mass ratios of their formation event: $0.2 < \Delta M_h/M_h < 0.3$ (*dots*), $0.3 < \Delta M_h/M_h < 0.5$ (*dashes*), and $\Delta M_h/M_h > 0.5$ (*solid line*). The similarity of all three distributions implies that metallicity bimodality is not caused by mergers with any particular range of mass ratios. The observed distribution is plotted in *dot-dashes* for comparison.

galaxy formation, so it is conceivable that different types of star clusters may be formed depending on the merger ratio. In Figure 4.19 we plot cumulative metallicity distributions for clusters grouped according to the mass ratios in the cluster-forming merger event. Mergers with the closest masses, $\Delta M_h/M_h > 0.5$, contribute 48% of `case-1` clusters, while the lower two mass ranges each contribute equal portions of the rest. Running KS test on these distributions revealed that they formally represent statistically different populations. However, there is no clear-cut range of metallicities where one type of merging is exclusively producing all of the clusters, and the overall shapes of the distributions are similar. This uniformity suggests that bimodality is a natural consequence of hierarchical cluster formation regardless of the exact definition of a "major" merger.

Figure 4.20 shows how many models from the grid fall into particular ranges of the Dip probability and the ratio of `case-2` clusters to `case-1` clusters, N_2/N_1. The models are restricted to have the normalized number of clusters $140 < N_1 + N_2 < 160$. The region with the highest density of models is in the lower-right corner of the plot, corresponding

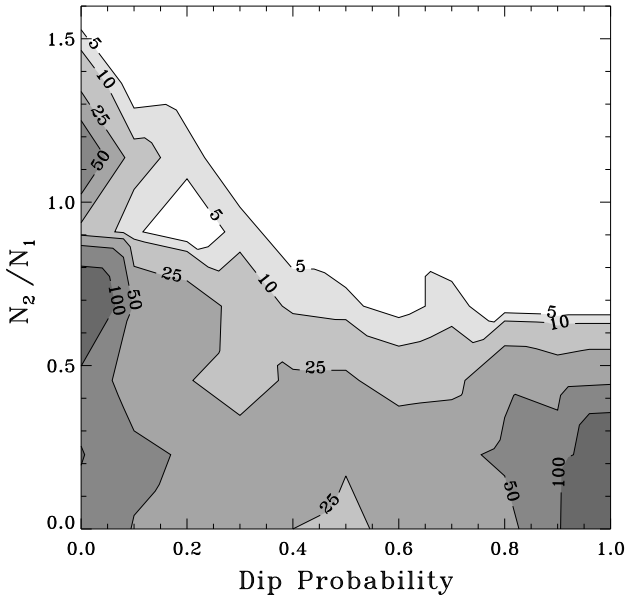

Figure 4.20. Number of models resulting in particular values of the Dip probability and the ratio of `case-2` to `case-1` clusters, for all realizations of the parameter grid with the normalized number of clusters $140 < N < 160$.

to high $P_{\rm dip}$ and low N_2/N_1. Low values of $P_{\rm dip}$ are not significant since they cannot reject a unimodal distribution. Effectively, bimodality requires $N_2/N_1 \lesssim 0.5$. At the significance level of $P_{\rm dip}$ = 90%, corresponding to the observed distribution, 38% of the grid models are bimodal if $N_2/N_1 < 0.3$. This fraction drops to only 15% for $0.3 < N_2/N_1 < 1$, and then further to 9% for $N_2/N_1 > 1$. These statistics confirms that bimodal populations appear only when the `case-2` channel is a secondary formation mechanism.

Another part of the explanation of bimodality of the surviving clusters is due to the dynamical evolution. Most of the disrupted clusters were old and blue. If we add these disrupted clusters to the metallicity histogram in the fiducial model, the blue peak rises by a factor of 2. The red peak remains virtually unaffected, since the more recently formed red clusters are less subjected to dynamical disruption. We ran the Dip test on all grid distributions, including both surviving and disrupted clusters. Among the models with the number of surviving clusters in the range $100 < N < 200$, few distributions have $P_{\rm dip} > 50\%$ and none has $P_{\rm dip} > 80\%$. This means virtually no bimodality. Indeed, the distributions appear almost entirely unimodal, with the peak in the blue end and nothing more than

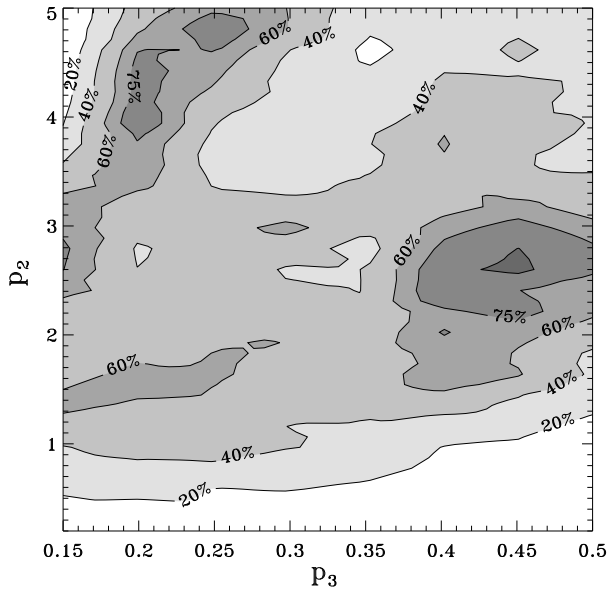

Figure 4.21. Same as Figure 4.13 but for the models with an alternative prescription for M_* as a function of M_h, given by equation (4.9).

a tail in the red end. This leads to a prediction that late-type galaxies, which have more continuous cluster formation than early-type galaxies, may be less likely to exhibit bimodal cluster populations.

4.4.5 Alternative Formation Prescriptions

As alluded to in Section 4.2, some equations that we used in the prescriptions for the stellar mass and the cold gas fraction were based on only a few observed points. Currently there is limited observational or theoretical understanding of how these functions should behave at high redshift, which is the period of primary interest for our study. Below we consider some alternatives for these prescriptions.

The stellar fraction that we adopted from Woo et al. (2008) is well motivated by Milky Way dwarf galaxies at the present epoch, but the abundance-matching models such as the one by Conroy & Wechsler (2009) predict a steeper dependence on halo mass in the range $10^8\,\mathrm{M}_\odot < M_* < 10^{10}\,\mathrm{M}_\odot$. Additionally, the redshift dependence of this relation is uncertain, and recent observational surveys (Borch et al., 2006; Bell et al., 2007; Dahlen et al., 2007) have advocated a slower evolution than $(1+z)^{-2}$ adopted in our model. To accommodate

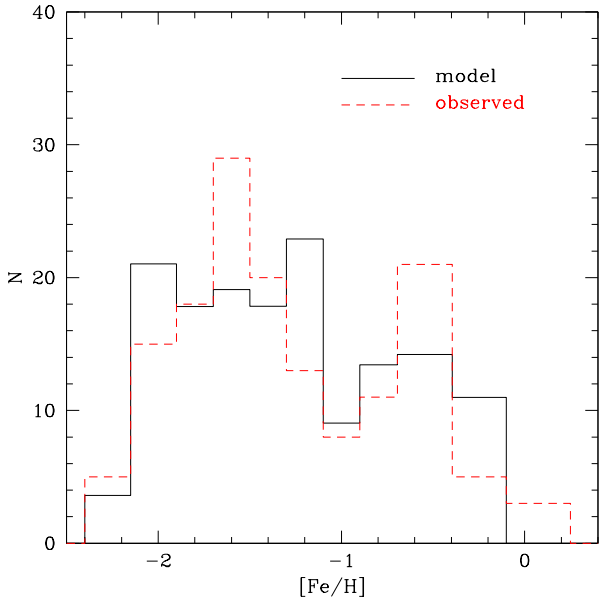

Figure 4.22. Metallicity distribution at $z = 0$ in the best-fit model with an alternative prescription for M_* as a function of M_h, given by equation (4.9) (*solid histogram*), compared to the observed distribution of Galactic globular clusters (*dashed histogram*).

this uncertainty, we re-ran the entire parameter grid with equation (4.9) instead of equation (4.7). The corresponding contour plot is shown in Figure 4.21 and the best-fit metallicity distribution is shown in Figure 4.22. This best fit is capable of reproducing the observed metallicity ($P_{KS,Z}$ = 49%) and mass distributions ($P_{KS,M}$ = 9.5%), similar to our fiducial model. The acceptable range of the parameter space is narrower and shifted towards higher values of p_2, as the steeper mass slope otherwise prevents low-mass halos from forming a sufficient number of clusters. Nevertheless, this alternative prescription still leads to a significant chance of a bimodal metallicity distribution.

Current observational constraints on the gas fraction at high redshift are even more uncertain. We adjusted the fit by altering the scale mass in equation (4.6). As alternatives to the fiducial model, we considered a power law of redshift, $M_s = M_{s,0}(1+z)^2$, and an inverse time dependence, $M_s = M_{s,0}(t/t_0)^{-1}$. Both of these relations resulted in lower gas fractions during the high-redshift epoch when most globular clusters should form. The gas fractions were too low for `case-2` formation for any reasonable choice of p_5. More importantly, the low gas masses did not allow even the most massive halos to form star

clusters until intermediate redshifts. Therefore, none of these fits is a viable alternative to the fiducial model.

The same problems manifested if we took the simplistic approach of holding the gas fraction constant for all halo masses at all times. This idea was initially considered to see if we could generate simple results based only on halo merger histories without speculation on the baryonic physics. We quickly realized that this approach was not going to work. Setting f_g too low effectively prevents cluster formation at high redshift, when blue globular clusters are expected to form, as most halos cannot build up sufficient mass to overcome the minimum mass required to form a single massive star cluster (discussed in Section 4.2.2). A constant gas fraction that is too high, on the other hand, presents obvious unphysical predictions at low redshift, and in particular would drastically over-predict the number of young clusters, forcing us to arbitrarily truncate their formation.

We also considered an alternative parametrization of the gas fraction, suggested by Stewart et al. (2009). They took the same observational constraints as us, but fitted them as

$$\frac{M_g}{M_*} = 0.04 \left(\frac{M_*}{4.5 \times 10^{11}\,\mathrm{M}_\odot} \right)^{-0.59(1+z)^{0.45}} . \tag{4.27}$$

This formula predicts so much cold gas at high redshift that many low-mass halos would be able to form clusters via `case-2` channel for any $p_5 < 1$. If we completely disable `case-2` formation and use the above prescription for the gas mass, we find many model realizations consistent with the observed metallicity distribution. This prescription differs from our fiducial choice in that it produces considerably more young clusters and achieves less clear metallicity bimodality. The maximum value of the likelihood function attainable with this prescription is approximately half of the value for the fiducial prescription. Nevertheless, it could still produce acceptable globular cluster results.

In addition to changing the formulation of the fits, we investigated the effect of adding a random Gaussian dispersion with standard deviation σ_{fits} to the right hand sides of equations (4.3), (4.5), and (4.7) to reflect their intrinsic scatter as well as observational uncertainty. Different random values for the three scatters are generated for each halo at each timestep, but we always force the condition (4.8) on the total baryon content. For simplicity, we used

the same magnitude of $\sigma_{\rm fits}$ for the scatter added to all three equations simultaneously. We re-ran the parameter search grid using $\sigma_{\rm fits} = 0.1, 0.2$, and 0.3 dex. For each value of $\sigma_{\rm fits}$, we were still able to find models with high values of $P_{KS,Z}$ and overall likelihood statistic, although these values decline with the increasing amount of scatter. The metallicity probability varies from 49% to 16% to 6%, for $\sigma_{\rm fits} = 0.1, 0.2, 0.3$ dex respectively.

As an alternative to scatter in the cutoff mass (eq. 4.3) with a fixed functional form for the gas fraction (eq. 4.2), we tried adding scatter to equation (4.2) while keeping fixed equation (4.3). Adding scatter to $f_{\rm in}$ allows the gas fraction to exceed the threshold p_5 much more easily and produce too many `case-2` clusters. To avoid unphysical results, we analyze only results for `case-1` formation. In this case we find the best-fit models with $P_{KS,Z} = 47\%, 13\%$, and 5%, for $\sigma_{\rm fits} = 0.1, 0.2, 0.3$ dex respectively. These models are still consistent with the observed Galactic distribution.

The addition of scatter as described above has two systematic effects on any individual realization of the metallicity distribution: the high metallicity tail is extended even further and the height of the blue peak is damped relative to the case of no scatter. The former effect is due to the possibility of drawing higher values of M_* and hence higher [Fe/H]. The latter effect arises from the enforcement of equation (4.8), which prevents gas-rich halos at high-redshift from gaining any extra gas from the positive scatter in equation (4.5); on the other hand, negative scatter can prevent some of these halos from being eligible for `case-2` formation. Accordingly, the best distributions with higher values of $\sigma_{\rm fits}$ were found for models with low values of p_5. We note that the Dip probability in most realizations is not strongly affected by the new scatter, implying that the actual smearing of the peaks is not significant and bimodality is preserved.

4.4.6 Alternative Dynamical Disruption

In Section 4.3 we noted that the expression for the evaporation rate (eq. [4.23]) contains some inherent parameters. Here we explore alternative disruption models with different values of ξ_e and δ within the fiducial formation prescription.

The effect of decreasing ξ_e is simply to reduce the number of clusters that are completely disrupted by $z = 0$. In the fiducial model with $\xi_e = 0.033$, about 60% of the orig-

inal sample is disrupted. (Note that this implies that roughly $5 \times 10^7\,M_\odot$ worth of stars in the Galactic stellar halo could be remnants of the disrupted clusters.) With the factor of $\xi_e = 0.02$, only $\sim 30\%$ are disrupted. With the factor of $\xi_e = 0.01$, almost all clusters survive.

The effect of decreasing δ and δ_0 is to shift the peak of the mass function to a lower mass.

We repeated the grid parameter search for the best metallicity distribution for two alternative prescriptions, one with $\xi_e = 0.02$, $\delta = \delta_0 = 1/3$, and the other with $\xi_e = 0.033$, $\delta = \delta_0 = 1/9$. We found that in both cases our model could produce an observationally-consistent metallicity distribution. However, lowering either ξ_e or δ significantly alters the mass function away from the data, by allowing too many low-mass clusters to survive. Raising ξ_e and δ may improve the mass function, but steers away from recent constraints on the two parameters (Baumgardt & Makino, 2003; Gieles & Baumgardt, 2008). Therefore, we ultimately conclude that our fiducial prescription ($\xi_e = 0.033$, $\delta = \delta_0 = 1/3$) works best.

For illustration, we list below the properties of the best models in the two alternative prescriptions. For $\xi_e = 0.02$, $\delta = \delta_0 = 1/3$, we find a peak model that has a metallicity distribution with $P_{KS,Z} = 69\%$ and mass distribution with $P_{KS,M} = 0.02\%$. The parameters of this model are $p_2 = 4.4$, $p_3 = 1.4$, $p_4 = 0$, $p_5 = 0.99$.

For $\xi_e = 0.033$, $\delta = \delta_0 = 1/9$, we find a peak model that has a metallicity distribution with $P_{KS,Z} = 19\%$ and mass distribution with $P_{KS,M} = 12\%$. The parameters of this model are $p_2 = 4$, $p_3 = 1.4$, $p_4 = 0$, $p_5 = 0.97$. Its metallicity distribution is shown in Figure 4.23. The overabundance of metal-poor clusters is clear.

The latter alternative prescription ($\delta = 1/9$) predicts the disruption time to scale with cluster mass in a manner nearly identical to Gieles & Baumgardt (2008), if we take that all model clusters have a mean galactocentric frequency $\omega = (2.4 \times 10^7\,\mathrm{yr})^{-1}$. Based on the orbit calculations discussed in the next section, we find a similar median value of ω for the sample of model clusters in the fiducial model. We also find no correlation between ω and cluster mass, which confirms our assertion in Section 4.3 that in such a model the disruption time of the average cluster would scale with mass as $\nu_{\mathrm{ev}}^{-1} \propto M^{2/3}$.

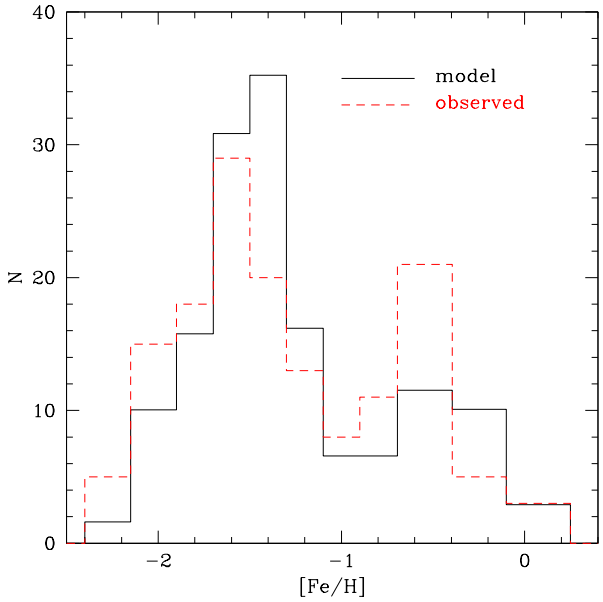

Figure 4.23. Metallicity distribution at $z = 0$ in the best-fit model with an alternative dynamical disruption prescription, $\delta = \delta_0 = 1/9$ (*solid histogram*), compared to the observed distribution of Galactic globular clusters (*dashed histogram*).

4.5 Spatial Distribution

Using the same *N*-body simulation as in this paper, Prieto & Gnedin (2008) investigated the present spatial distribution of model clusters that formed in high-redshift (metal-poor) galactic systems. They calculated the orbits of clusters from the time when their host galaxies accreted onto the main galaxy and identified three distinct populations. *Disk clusters* formed in the most massive halo that eventually hosts the present Galactic disk. These clusters, found within the inner 10 kpc, are scattered into eccentric orbits by the perturbations from accreted galactic satellites. *Inner halo clusters*, found between 10 and 60 kpc, came from the now-disrupted satellite galaxies. Their orbits are inclined with respect to the Galactic disk and are fairly isotropic. *Outer halo clusters*, beyond 60 kpc from the center, are either still associated with the surviving satellite galaxies, or were scattered away from their hosts during close encounters with other satellites and consequently appear isolated.

The azimuthally-averaged space density of metal-poor globular clusters is consistent with a power law, $n(r) \propto r^{-\gamma}$, with the slope $\gamma \approx 2.7$. Since all of the distant clusters originate in progenitor galaxies and share similar orbits with their hosts, the distribution of

the clusters is almost identical to that of the surviving satellite halos. This power law is similar to the observed slope of the metal-poor globular clusters in the Galaxy. However, the model clusters have a more extended spatial distribution (larger median distance) than observed. In the model it is largely determined by the orbits of the progenitor galaxies and the epoch of formation. Moore et al. (2006) showed that the early-forming halos are more spatially concentrated and in order to match the Galactic distribution, globular clusters would need to form at $z \sim 12$. However, such an early formation may be inconsistent with the requirement of high mass and density of the parent molecular clouds.

In this work, we have retraced some of these steps to attempt to reproduce the spatial distribution of the whole Galactic globular cluster system.

The clusters that formed in the disk of the main halo are assigned radial positions according to the exponential profile, $dN/dR \propto Re^{-R/R_d}$, with the observed scale length of the Galactic disk, R_d = 3 kpc. The azimuthal angles are assigned randomly. The vertical position in the disk is also assigned randomly, with the scale-height of one fifth of the scale-length. The clusters are limited to the radial range $0.6 < R < 10$ kpc, where the observed disk globular clusters are located. The distances are also given a random Gaussian scatter of 10% to replicate observational distance uncertainties.

Clusters that formed in satellite halos that survived until $z = 0$ are assigned the present position of the host, with a small displacement analogous to the distribution in the main disk. Clusters that formed in subhalos that did not survive until $z = 0$ are initially assigned the last known position and velocity of the host in the simulation, with the same displacement as above. We then follow the orbits of these stray clusters until $z = 0$ using a leap-frog integration scheme with fixed time step.

The orbit integration follows Prieto & Gnedin (2008). The main halo and the satellite halos contribute their NFW potentials, while the disks within the halos contribute the Miyamoto-Nagai potentials with the total mass of gas and stars computed from equations (4.5) and (4.7). The total gravitational potential is computed by linearly interpolating the masses of halos and subhalos between the simulation snapshots. Positions of subhalos at each timestep are computed with cubic splines between the snapshots. We also include the acceleration on the clusters that results from the use of the splines, as described in Prieto

& Gnedin (2008). Cosmological dark energy contributes an additional component to the acceleration in physical coordinates: $\mathbf{a}_\Lambda = \Omega_\Lambda H_0^2 \mathbf{r}$.

Just as in the previous study, we find a more extended spatial distribution of the globular cluster system than that observed in the Galaxy. Clusters that formed in surviving satellites (about 24% of the sample in the fiducial model) are the most distant from the center, as forced by the location of the satellites. The orbit integration for the clusters formed in disrupted satellites (about 52% of the sample) shows that these clusters also do not migrate in r far from the last known position of their host. Such coupling to the dark matter halos is the main reason for the overextended cluster system.

Clusters that formed in the disk of the main halo (the remaining 24% of the sample) most closely resemble the spatial properties of the Galactic clusters. They are confined to the inner 10 kpc and would be referred to as the bulge or disk clusters. However, this group should contain more than 50% of the sample to be consistent with observations. A recent paper by Griffen et al. (2010) similarly investigated the formation of red clusters by major mergers in the Aquarius simulation and concluded that such clusters must have formed in the central disk.

Note that our orbit calculations, as well as those by Griffen et al. (2010), use the gravitational potential derived in collisionless cosmological simulations. Stars and cold gas would deepen the gravitational potential in the inner regions of the main halo and bring the satellites closer to the center. Dense stellar nuclei of the satellites should also survive against tidal disruption longer than pure dark matter halos. The hydrodynamic simulations of Naab et al. (2009) show that the combined effect of baryons may be to deposit half of stellar remnants of the disrupted satellites, including their globular clusters, within 10 kpc of the center. This would effectively reconcile the predicted cluster distribution with the Galactic sample, since over 50% of our clusters formed in disrupted satellites. More detailed hydrodynamic simulations of galaxy formation are needed to verify either hypothesis.

An observational test of the cluster orbits would be possible when proper motions are measured for a large fraction of the Galactic clusters. Such measurements could be achieved with the planned SIM-Lite space observatory.

4.6 Globular Cluster Colors

We attempted another direct comparison of the model predictions with the observed sample, by constructing single stellar population models using Bruzual & Charlot (2003) code GALAXEV. As input for GALAXEV, we used the age and metallicity for each globular cluster created with the fiducial model.

The distribution of the model $B-V$ color was considerably less bimodal than the metallicity distribution discussed in previous sections. The main cause of the smearing of the two peaks appears to be the younger age of metal-rich clusters predicted by the model. The metal-poor clusters constitute a clearly defined peak at $B-V$ = 0.67, which corresponds well to the blue peak of the Galactic sample. But the metal-rich clusters, which are expected to make up the red peak, have a mean $B-V$ color of 0.77, while the observed red mean is close to 0.85. The standard deviation of the red model clusters is 0.08, implying that the result is consistent within one sigma of the observed, but the bimodality of the distribution is not evident to the eye.

To test the hypothesis that the smearing of the color peaks compared was due to the relative age of the populations, we ran the population synthesis models again, this time using a constant age of 12.1 Gyr for all clusters. The resulting distribution indeed appeared to constitute two peaks, with a blue peak at a mean of $B-V$ = 0.67 and a red peak at $B-V$ = 0.84, with a clearly defined gap between them. It should be noted that a known discrepancy exists between the $B-V$ colors predicted by all major population synthesis codes and those of observed globular clusters (Conroy & Gunn, 2009). All models predict colors that are too blue at high metallicity, which would directly play into smearing bimodality in our result.

In addition to the colors, we examined the cluster luminosities (absolute V-band magnitudes) calculated by GALAXEV. These allow a more direct comparison with the observations than the mass function presented in Section 4.3, for which we were required to assume a constant mass-to-light ratio for all observed clusters. This constant M/L_V ratio has been a traditional approach, but has come under recent scrutiny by Kruijssen (2008) who argued that M/L_V may vary as a function of cluster mass. However, the distribution

of V-band magnitudes for our fiducial model has a KS probability of 4.3%, which is not a significant departure from the 7.4% for the mass function. When we tried the same exercise as above by setting the ages of all model clusters to 12.1 Gyr, the KS probability jumped to 25%. This improvement likely happened because we converted the magnitudes of observed clusters into masses using $M/L_V = 3$, while GALAXEV typically predicted $M/L_V < 3$ for the 12.1 Gyr isochrones. This brought the mean luminosity of model clusters closer to the observed value than the average mass of model clusters was to its observed counterpart.

Even though these population synthesis results are interesting, we believe that the mass and metallicity distributions presented in previous sections are more reliable. Population synthesis modeling adds an extra layer of empirical uncertainty to our results, as the specific nature of horizontal-branch evolution remains an issue that has not been completely resolved.

4.7 Summary and Implications for Galaxy Formation Models

We have presented a model for the origin of the metallicity distribution of globular clusters. In our scenario, bimodality results from the combination of the history of galaxy assembly (rate of mergers) and the amount of cold gas in protogalactic systems. Early mergers are frequent but involve relatively low-mass protogalaxies, which produce preferentially blue clusters. Late mergers are infrequent but typically involve more massive galaxies. As the number of clusters formed in each merger increases with the progenitor mass, just a few late massive mergers can produce a significant number of red clusters. The concurrent growth of the average metallicity of galaxies between the late mergers leads to an apparent "gap" between the red and blue clusters.

The peak metallicities of the red and blue populations are remarkably robust to variations of the model parameters. The peaks encode the mass-metallicity relation in galaxies and do not depend strongly on the rate or timing of cluster formation. The exact definition of a major merger is also not important for our result, as long as the merger mass ratio is at least 1:5.

Our conclusions on the origin of metallicity bimodality are not significantly affected by the large uncertainties in our knowledge of the stellar mass and cold gas mass in high-

redshift galaxies. We considered alternative prescriptions for the stellar fraction, gas-to-stars ratio, and even dynamical disruption, but in all cases found a metallicity distribution consistent with the observations. Such robustness indicates that most external factors are not as important as the internal mass-metallicity relation in host galaxies.

We find that dynamical disruption over the cosmic history naturally converts an initial power-law cluster mass function into an observed log-normal distribution. A continuous formation of clusters in the first several Gyr help to replenish the depleted low-mass end. Dynamical disruption also helps establish metallicity bimodality by preferentially depleting old clusters in the metal-poor peak.

Our prescription links cluster metallicity to the average galaxy metallicity in a one-to-one relation, albeit with random scatter. Since the average galaxy metallicity grows monotonically with time, the cluster metallicity also grows with time. Our model thus encodes an age-metallicity relation, in the sense that metal-rich clusters are somewhat younger than their metal-poor counterparts. Observations of the Galactic globular clusters indicate an age spread that ranges from 1 Gyr for the inner blue clusters to 2 Gyr for the inner red clusters to 6 Gyr for the outer clusters, which is generally consistent with the predicted spread. However, the model may be marginally inconsistent with the observation that some of the metal-rich clusters appear as old as the metal-poor ones. Note that our model is still simplistic and does not include metallicity gradients within protogalaxies, which may dilute the predicted age-metallicity relation.

Our model demonstrates that star cluster formation during gas-rich mergers of protogalactic systems is a single mechanism that successfully reproduces many observed properties of the Galactic globular clusters. It may avoid the need for two separate formation mechanisms for the red and blue clusters invoked in the model of Beasley et al. (2002). Their model relied on constant cluster formation efficiency relative to the field stars, but required different efficiencies for the two modes. While the red clusters in their models continued forming through galaxy mergers, the blue clusters needed to be arbitrarily shut off at $z = 5$. This was the main cause of the bimodal metallicity distribution in their model, as the blue clusters did not have much overlap with the red clusters that formed in major mergers involving metal-enriched gas considerably after $z = 5$. The Beasley et al. (2002)

model also neglected the effects of the dynamical evolution that shaped the present cluster distribution. In contrast, in our model some old blue clusters are disrupted and some are unable to form at recent times because the protogalaxies are gas-poor. Another difference is that in our model major mergers contribute both red and blue clusters, while in Beasley et al. they contribute only red clusters. We also find that globular clusters form significantly earlier than the bulk of field stars and therefore the two cannot be linked by a constant formation efficiency at all times (see Fig. 4.10).

We have compared the metallicity distribution of globular clusters to the mass-weighted metallicity distributions of other stellar populations as predicted by our scaling relations given in Section 2. We find that galaxy field stars overall have a single-peaked distribution with a mean of $[\mathrm{Fe/H}] \approx 0$, a metal poor tail, and no stars with $[\mathrm{Fe/H}] > 0.4$. This is consistent with our current understanding of the metallicity of stars in the Galactic disk. The stars in surviving satellites, which correspond to Milky Way dwarf galaxies, also appear to have a single-peaked distribution with a mean metallicity $[\mathrm{Fe/H}] \approx -1$. Only the globular cluster system display a bimodal metallicity distribution.

We derived some simple scaling relations for the overall efficiency of globular cluster formation. We adopted the cluster formation rate in gas-rich, high-redshift merger events (eq. 4.10) that scales with the host system mass as $M_{GC} \sim 10^{-4} M_g / f_b \sim 10^{-4} M_h$. We have later learned of a similar empirical relation for all types of massive galaxies, derived independently by Spitler & Forbes (2009) and Georgiev et al. (2010). The outcome of the model is a prediction that the fraction of galaxy stellar mass locked in star clusters, M_{GC}/M_*, is of the order $10-20\%$ at $z > 3$ and then declines steadily with time to about 0.1% at present. The specific frequency parameter follows a similar decline with time and reaches $N/(M_*/10^9\,\mathrm{M}_\odot) \sim 1$ at the present. These efficiencies are in agreement with the compilations of McLaughlin (1999), Rhode et al. (2005), and Peng et al. (2008). We also find that the globular cluster system overall is significantly more metal-poor than the galactic spheroid, which is populated by stars from the disrupted satellites.

Our scenario can be applied to other galactic environments, such as those of elliptical galaxies which contain much larger samples of globular clusters. For example, Peng et al. (2008) showed that the fraction of red clusters increases from 10% to 50% with increasing

luminosity of elliptical galaxies in the Virgo cluster. In our model, globular cluster formation is entirely merger-driven. We showed that the Galactic sample may have arisen from early super-gas-rich low-mass mergers and later metal-rich high-mass mergers. Compared to the Galaxy, giant ellipticals are expected to experience more high-mass mergers which would contribute more prominently to the globular cluster system. As Figure 4.7 shows, such mergers would produce comparable numbers of red and blue clusters simultaneously. Thus the fraction of red clusters should increase with galaxy mass, reaching $\sim 50\%$ for giant ellipticals. This trend, observed by Peng et al. (2008), may be a natural outcome of the hierarchical formation.

At the other end of the galactic spectrum, dwarf galaxies likely lacked metal-rich mergers and produced only metal-poor blue clusters. In particular, dE and dSph type dwarfs which are now deprived of cold gas are not expected to contain any young and metal-rich clusters. Some dIrr galaxies, such as the LMC, still possess considerable amounts of cold gas and may produce younger clusters, although they are still likely to have subsolar metallicity. The variety of globular cluster ages observed in the LMC indicates that it may have had bursts of star formation throughout its cosmic evolution.

Our study places interesting constraints on galaxy formation models. Within the framework of our model, acceptable mass and metallicity distributions result only from a certain range of the parameters. In particular, the minimum ratio of masses of merging protogalaxies strongly correlates with the cluster formation rate. If the clusters form very efficiently only a few massive mergers are needed; if the clusters form inefficiently many mergers are needed, which requires a lower merger threshold. However, mass ratios of less than 0.2 are disfavored in the model (see Fig. 4.13). Formation of massive clusters in very gas-rich systems without detected mergers (our `case-2` scenario) improves the final mass function but is not required for reproducing the metallicity distribution. Thus, globular cluster formation solely in major mergers is consistent with the available observations. Finally, our results rest on the derived prescription for the cold gas fraction as a function of halo mass and cosmic time. This prescription (Fig. 4.1) can be tested by future observations of high-redshift galaxies with JWST and by detailed hydrodynamic simulations.

CHAPTER 5

Conclusions

Here, I present a summary of all topics discussed in the thesis, and provide a roadmap for future work that can further illuminate each subject.

5.1 Chapter 2 Summary

In this chapter, I revisited the formation and evolution of the first galaxies using new hydrodynamic cosmological simulations with the ART code, and attempted to quantify the effect that Pop III stars had on the evolution of these galaxies. Our simulations featured separate prescriptions for the formation of Pop III and Pop II stars. The latter featured recently developed model for H_2 formation and dissociation as presented in Gnedin & Kravtsov (2011), and a star formation recipe that is based on molecular rather than atomic gas. For the former, we developed and implemented a recipe for the formation of metal-free stars in galaxy-scale simulations that resolve primordial clouds with sufficiently high density. We based our recipe on the results of prior zoom-in simulations that resolved the protostellar collapse in pre-galactic objects. We focused our investigation on the possibility that a large fraction of Pop III stars were massive, and that many exploded in pair-instability supernovae. Our recipe prescribed very high output for ionizing radiation from Pop III stars, and thermal injections for the supernovae that generated shocks and outflows.

Sufficiently massive galaxies which hosted Pop III stars did not retain dynamical signatures of their thermal and radiative feedback for more than 10^8 yr after the lives of the stars ended in pair-instability supernovae, even when we consider the maximum reasonable efficiency of the feedback. Though metals ejected by the supernovae traveled well beyond the virial radius of the host galaxy, they typically begin to fall back quickly, and

do not enrich a large fraction of the intergalactic medium. Galaxies with total mass in excess of $3 \times 10^6\,M_\odot$ re-accrete most of their baryons and transition to metal-enriched Pop II star formation. This accretion process is driven by a combination of the expanding virial radius, increasing gravitational potential, and rapid inflows of the early stages of galaxy formation. Galaxies less massive than this limit were more vulnerable to the feedback, especially those in underdense regions of the universe. We found that roughly 20% of all pair-instability supernovae may result in ejections that are permanently incorporated into the diffuse intergalactic medium.

5.1.1 Future Work

As discussed in the introduction of this thesis, studying the first stars and galaxies with ab initio cosmological simulations is inherently appealing due to the relative simplicity of the initial conditions. This simplicity was exploited by researchers at the turn of the century to make bold predictions regarding the first stars, differentiating them significantly from modern-day counterparts by the top-heavy initial mass function. Our study showed that even with a top-heavy initial mass function, it is doubtful that Pop III stars made much of a long-lasting dynamical impact on their host galaxies. We can still hope to observe the impact of the first stars is by the actual detections of their explosions, which may be possible with JWST (Wise & Abel, 2005; Hummel et al., 2012; Pan et al., 2012a; Whalen et al., 2012). Alternatively, we can gain some clues by observing the chemical signatures of the stars which formed from their ejecta (Bromm & Loeb, 2003; Frebel et al., 2007).

Under detailed scrutiny, the formation of the first stars even in completely metal-free gas was shown to be considerably more complicated than initially believed. Indeed, turbulence (Wise et al., 2008; Greif et al., 2008; Prieto et al., 2012; Latif et al., 2013), magnetic fields (Turk et al., 2012), and radiative feedback that could terminate accretion (Hosokawa et al., 2011; Stacy et al., 2012) could potentially push the IMF to lower characteristic masses. Perhaps most importantly, simulations that continued to follow the protostellar cores for long periods of time find that fragmentation can naturally occur from classic disk instabilities (Stacy et al., 2010; Clark et al., 2011; Greif et al., 2011, 2012). It must be said that no simulation has actually followed the accretion process to its termination,

and the protostellar cores may continue to grow to be more massive. Future simulations should focus on following these processes for longer periods of time and culminating into a statistical initial mass function.

From the perspective of galaxy formation simulations like the ones which constitute my thesis, the way forward most likely involves scaling down our estimate for the characteristic mass of Pop III stars in future works. At the very least, predictions should be made for the chemical signatures and supernova rates for Pop III stars of varying masses. We have already begun to look at this with the "low-mass Pop III IMF" run discussed in sections of Chapters 2 and 3. In future works, this can be made more realistic by tracking the details of ordinary stellar explosions in the minihalo environments. Studies such as Ritter et al. (2012) have already taken the initiative on this frontier. Once the sample of low-metallicity stars becomes well characterized, and JWST is able to probe to sufficiently high redshifts, our models will finally be constrained and ab initio simulations of galaxy formation can accurately connect to later epochs.

5.2 Chapter 3 summary

In this chapter, we investigated the transition from primordial Pop III star formation to normal Pop II star formation in the first galaxies using the cosmological hydrodynamic simulations first presented in Chapter 2. Continuing the simulations after the pair-instability supernova explosions revealed that while these explosions seeded their host galaxies with metals, the enrichment of the intergalactic medium was very inefficient, even with our assumption of a top-heavy initial mass function. This means that Pop III star formation could potentially continue until $z \approx 6$ in different unenriched regions of the universe, before being ultimately shut off by cosmic reionization.

Given this pattern, each galaxy is responsible for its own chemical enrichment. A single pair instability supernova is all that is needed to seed Pop II star formation, and Pop II stars typically overtake Pop III stars in 20-200 Myr, based on the amount of stellar feedback and metal production. The exception to this rule happens for minihalos in underdense regions with very slow accretion rates.

5.2.1 Future work

As discussed in the summary of Chapter 2, the conclusions of this chapter are also liable to change when considering alternative forms for the mass function of Population III stars. If the chemical feedback of the first stars is somewhat more modest, one supernova will no longer be enough to facilitate a complete transition to Pop II star formation within a galaxy. This puts greater emphasis on the need to understand star formation in low, but non-zero metallicity environments. We postulate that the inclusion of dust physics into fully radiative hydrodynamic simulations, combined with realistic dust yields from early supernovae will allow for the most sophisticated simulations of the process yet. The critical metallicity for the transition to Pop II star formation is subject to vary depending on dust content of the gas. Recent studies (e.g., Aykutalp & Spaans 2011) indicate that the interplay between UV radiation field and metallicity must be considered when determining whether gas can cool enough to fragment efficiently in star-forming regions. Even in the high-density regime of protostellar collapse, dust cooling makes a difference (Omukai et al., 2005).

Another way to make these simulations more realistic is by considering the feedback effects of radiation pressure from starlight. This is currently being explored in modern-day star-formation (Hopkins et al., 2011; Agertz et al., 2013) as well as mostly primordial environments (Wise et al., 2012b). The strength of this feedback may also depend significantly on dust physics, as dust provides an additional source of optical depth to drive winds. Generally speaking, ab initio simulations of galaxy formation during unresolved epochs can only be fully trusted once the community can agree on a single comprehensive, realistic recipe to model present-day stellar feedback.

The payoff for performing these more realistic high-redshift galaxy simulations is immeasurable. These galaxies are attributed to being the yet-undetected faint sources that contributed to reionization (Robertson et al., 2013), the building blocks of present-day massive galaxies, and in some cases, they can even survive to become remnant dwarf galaxies (Ricotti & Gnedin, 2005). Ultimately, a self-consistent theory of galaxy formation will require that ab-initio simulations of the first galaxies can be connected to the present day. The accuracy of our modeling should become increasingly constrained when new observa-

tional facilities such as JWST, ALMA, LOFAR, and the next generation of ground-based optical telescopes become fully realized.

5.3 Chapter 4 Summary

In this chapter, I presented a model which prescribes the formation of globular clusters semi-analytically using galaxy assembly history from cosmological simulations coupled with observed scaling relations for the amount and metallicity of cold gas available for star formation. We assumed that massive star clusters form only during mergers of massive gas-rich galaxies and tune the model parameters to reproduce the observed distribution in the Milky Way. A wide, but not entire, range of model realizations produced metallicity distributions consistent with the data. The resulting metallicity distribution of the clusters was bimodal, with the peaks occurring at the same metallicities as in the Milky Way's. The distribution of cluster masses was also consistent with the Milky Way's once we applied a realistic treatment of dynamical mass disruption.

Many of the clusters were formed in now-disrupted satellite galaxies. Others were formed in surviving satellites, with some potentially accounting for the massive star cluster population associated with local dwarfs like the Large Magellanic Cloud. A third population that was primarily metal-rich formed in the most massive progenitor, making them analogous to the bulge population of the Milky Way. We found that early mergers of smaller hosts create exclusively blue clusters, whereas subsequent mergers of more massive galaxies create both red and blue clusters. Thus bimodality arises naturally as the result of a small number of late massive merger events. This conclusion is not significantly affected by the large uncertainties in our knowledge of the stellar mass and cold gas mass in high-redshift galaxies. The fraction of galactic stellar mass locked in globular clusters declines from over 10% at $z > 3$ to 0.1% at present.

5.3.1 Future work

One very notable challenge remains for the hierarchical galaxy formation scenario to explain the properties of globular clusters - dark matter. Our study, as well as other works that have modeled and simulated the formation of globular clusters (Beasley et al., 2002;

Bromm & Clarke, 2002; Boley et al., 2009; Griffen et al., 2010), generally prescribe globular clusters to form in environments where the dark matter density should also be high. However, to date, studies have generally found that there is no dynamical evidence for the presence of dark matter within globular clusters (Conroy et al., 2011), even taking into account the gradual dynamical evaporation of dark matter particles (Baumgardt & Mieske, 2008). If the evidence continues to be inclined this way, simulators and theorists will need to provide explanations for how gas physics can cause the clusters to form off-center in halos, thereby averting significant dynamical contributions from the dark matter. Fortunately, studies are already beginning to show that this is a plausible phenomenon (Powell et al., 2013). Our group's future simulations that will include detailed treatment of dust physics can shed further light on this, as we may be able to discover regions where the star-formation efficiency ought to be high enough where massive star clusters may form shielded from the negative effects of internal and external feedback.

Another recently discovered phenomenon in globular clusters is the apparent presence of multiple stellar populations, or at the very least, a variety of enrichment patterns within each cluster (Conroy & Spergel, 2011). Future work should be able to self-consistently account for this phenomenon, allowing for multiple bursts and/or self-enrichment by re-accretion of outflows.

The models presented in our study were specifically created for Milky Way-like galactic environments, and meant to reproduce the properties of only the Milky Way's globular cluster system. Future semi-analytical efforts will have to broaden the focus of modeling to also self-consistently reproduce the varying properties of globular clusters in other galaxies, from massive ellipticals to dwarf irregulars. Age, color, brightness, and metallicity measurements for large samples of extragalactic globular clusters will significantly increase the database available for constraining future modeling attempts (Chies-Santos et al., 2011; Cho et al., 2012; Pota et al., 2013). Once better constraints become available for the gas and stellar content of high-redshift galaxies, the models may need to be re-tooled to match these constraints.

One final possible future direction of research that will synthesize the work presented in all three chapters of this thesis is a study of the potential contribution of early globular

clusters to cosmic reionization. This possibility has been discussed in literature (Ricotti, 2002; Griffen et al., 2010, 2013; Katz & Ricotti, 2013), as globular clusters' ages are not in conflict with the possibility of forming prior to reionization. Indeed, in order to reproduce the current mass function of globular clusters, the total ratio of star formation that occurs in massive clusters approaches unity at high redshift, as discussed in Chapter 3 (Muratov & Gnedin, 2010). If massive clusters do form efficiently in low-mass galaxies, especially if they can form off-center, we expect that the escape fraction of ionizing radiation to be high, easing observational constraints on reionization. If this is indeed a physical possibility, this process may be resolved in my next generation of simulations by our group. Otherwise, the simulations should resolve the formation of globular clusters after the epoch of reionization, whether it be in protogalactic mergers (Kravtsov & Gnedin, 2005) or in star-bursting Lyman-alpha emitters (Elmegreen et al., 2012). Either conclusion would be a ground-breaking theoretical study in both galaxy formation and globular clusters.

BIBLIOGRAPHY

Abel, T., Anninos, P., Norman, M. L., & Zhang, Y. 1998, ApJ, 508, 518

Abel, T., Bryan, G. L., & Norman, M. L. 2000, ApJ, 540, 39

Abel, T., Bryan, G. L., & Norman, M. L. 2002, Science, 295, 93

Abel, T., Wise, J. H., & Bryan, G. L. 2007, ApJ, 659, L87

Ahn, K., Iliev, I. T., Shapiro, P. R., et al. 2012, ApJ, 756, L16

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

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Physical Review, 73, 803

Alvarez, M. A., Bromm, V., & Shapiro, P. R. 2006, ApJ, 639, 621

Alvarez, M. A., Finlator, K., & Trenti, M. 2012, ApJ, 759, L38

Ambartsumian, V. A. 1938, Uch. Zap. Leningrad G. U., 22, 19 [English transl. in IAU Symp. 113, Dynamics of Star Clusters, ed. J. Goodman & P. Hut (Dordrecht: Reidel), 521]

Ashman, K. M. & Zepf, S. E. 1992, ApJ, 384, 50

Ashman, K. M., Bird, C. M., & Zepf, S. E. 1994, AJ, 108, 2348

Ashman, K. M., & Zepf, S. E. 2001, AJ, 122, 1888

Aykutalp, A., & Spaans, M. 2011, ApJ, 737, 63

Barkana, R., & Loeb, A. 2001, Physics Reports, 349, 125

Barkat, Z., Rakavy, G., & Sack, N. 1967, Physical Review Letters, 18, 379

Bauer, D. J. 2007, Multivariate Behavioral Research, 42, 757

Baumgardt, H. 2001, MNRAS, 325, 1323

Baumgardt, H. & Makino, J. 2003, MNRAS, 340, 227

Baumgardt, H., & Mieske, S. 2008, MNRAS, 391, 942

Beasley, M. A., Baugh, C. M., Forbes, D. A., Sharples, R. M., & Frenk, C. S. 2002, MNRAS, 333, 383

Bell, E. F., Zheng, X. Z., Papovich, C., Borch, A., Wolf, C., & Meisenheimer, K. 2007, ApJ, 663, 834

Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2006, ApJ, 642, L137

Belokurov, V., Zucker, D. B., Evans, N. W., et al. 2007, ApJ, 654, 897

Benson, A. J. 2010, Physics Reports, 495, 33

Binney, J. & Tremaine, S. 2008, Galactic Dynamics (Princeton: Princeton University Press)

Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517

Boley, A. C., Lake, G., Read, J., & Teyssier, R. 2009, ApJ, 706, L192

Bond, H. E. 1981, ApJ, 248, 606

Bond, J. R., & Efstathiou, G. 1984, ApJ, 285, L45

Bond, J. R., Arnett, W. D., & Carr, B. J. 1984, ApJ, 280, 825

Borch, A., Meisenheimer, K., Bell, E. F., Rix, H., Wolf, C., Dye, S., Kleinheinrich, M., Kovacs, Z., & Wisotzki, L. 2006, A&A, 453, 869

Brodie, J. P. & Strader, J. 2006, ARA&A, 44, 193

Bromm, V., Coppi, P. S., & Larson, R. B. 1999, ApJ, 527, L5

Bromm, V., Ferrara, A., Coppi, P. S., & Larson, R. B. 2001, MNRAS, 328, 969

Bromm, V., Kudritzki, R. P., & Loeb, A. 2001, ApJ, 552, 464

Bromm, V., & Clarke, C. J. 2002, ApJ, 566, L1

Bromm, V., Coppi, P. S., & Larson, R. B. 2002, ApJ, 564, 23

Bromm, V., & Loeb, A. 2003, Nature, 425, 812

Bromm, V., Yoshida, N., & Hernquist, L. 2003, ApJ, 596, L135

Bromm, V., & Larson, R. B. 2004, ARA&A, 42, 79

Bromm, V., & Yoshida, N. 2011, ARA&A, 49, 373

Bromm, V. 2013, arXiv:1305.5178

Brooks, A. M., Governato, F., Booth, C. M., Willman, B., Gardner, J. P., Wadsley, J., Stinson, G., & Quinn, T. 2007, ApJ, 655, L17

Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000

Burbidge, E. M., Burbidge, G. R., Fowler, W. A., & Hoyle, F. 1957, Reviews of Modern Physics, 29, 547

Carr, B. J., Bond, J. R., & Arnett, W. D. 1984, ApJ, 277, 445

Ceverino, D., & Klypin, A. 2009, ApJ, 695, 292

Chernoff, D. F. & Weinberg, M. D. 1990, ApJ, 351, 121

Chies-Santos, A. L., Larsen, S. S., Kuntschner, H., et al. 2011, A&A, 525, A20

Cho, J., Sharples, R. M., Blakeslee, J. P., et al. 2012, MNRAS, 422, 3591

Ciardi, B., Ferrara, A., & Abel, T. 2000, ApJ, 533, 594

Clark, P. C., Glover, S. C. O., Klessen, R. S., & Bromm, V. 2011, ApJ, 727, 110

Clarke, C. J., & Bromm, V. 2003, MNRAS, 343, 1224

Conroy, C. & Gunn, J. E. 2009, ArXiv:0911.3151

Conroy, C. & Wechsler, R. H. 2009, ApJ, 696, 620

Conroy, C., & Spergel, D. N. 2011, ApJ, 726, 36

Conroy, C., Loeb, A., & Spergel, D. N. 2011, ApJ, 741, 72

Cote, P., Marzke, R. O., & West, M. J. 1998, ApJ, 501, 554

Couchman, H. M. P., & Rees, M. J. 1986, MNRAS, 221, 53

Crain, R. A., Eke, V. R., Frenk, C. S., Jenkins, A., McCarthy, I. G., Navarro, J. F., & Pearce, F. R. 2007, MNRAS, 377, 41

Dahlen, T., Mobasher, B., Dickinson, M., Ferguson, H. C., Giavalisco, M., Kretchmer, C., & Ravindranath, S. 2007, ApJ, 654, 172

Davé, R., Finlator, K., & Oppenheimer, B. D. 2007, EAS Publications Series, Volume 24, 183-189; astro-ph/0608537

De Angeli, F., Piotto, G., Cassisi, S., Busso, G., Recio-Blanco, A., Salaris, M., Aparicio, A., & Rosenberg, A. 2005, AJ, 130, 116

Dekel, A., & Silk, J. 1986, ApJ, 303, 39

Dotter, A., Sarajedini, A., Anderson, J., Aparicio, A., Bedin, L. R., Chaboyer, B., Majewski, S., Marín-Franch, A., Milone, A., Paust, N., Piotto, G., Reid, I. N., Rosenberg, A., & Siegel, M. 2010, ApJ, 708, 698

Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, ApJ, 136, 748

Elmegreen, B. G., Malhotra, S., & Rhoads, J. 2012, ApJ, 757, 9

Einstein, A. 1916, Annalen der Physik, 354, 769

Erb, D. K., Shapley, A. E., Pettini, M., Steidel, C. C., Reddy, N. A., & Adelberger, K. L. 2006, ApJ, 644, 813

Evrard, A. E. 1988, MNRAS, 235, 911

Fall, S. M. & Zhang, Q. 2001, ApJ, 561, 751

Forbes, D. A., Brodie, J. P., & Grillmair, C. J. 1997, AJ, 113, 1652

Forbes, D. A., & Bridges, T. 2010, MNRAS, 404, 1203

Frebel, A., Johnson, J. L., & Bromm, V. 2007, MNRAS, 380, L40

Frebel, A. 2011, arXiv:1108.4692

Frebel, A., & Bromm, V. 2012, ApJ, 759, 115

Friedmann, A. 1922, Zeitschrift fur Physik, 10, 377

Friedmann, A. 1999, General Relativity and Gravitation, 31, 1991

Fryer, C. L., Woosley, S. E., & Heger, A. 2001, ApJ, 550, 372

Fusi Pecci, F., Bellazzini, M., Buzzoni, A., De Simone, E., Federici, L., & Galleti, S. 2005, AJ, 130, 554

Gamow, G. 1948, Nature, 162, 680

Gebhardt, K. & Kissler-Patig, M. 1999, AJ, 118, 1526

Georgiev, I. Y., Puzia, T. H., Goudfrooij, P., & Hilker, M. 2010, MNRAS, 406, 1967

Gieles, M. & Baumgardt, H. 2008, MNRAS, 389, L28

Gieles, M. 2010, IAU Symposium, 266, 69

Gnedin, N. Y., & Ostriker, J. P. 1997, ApJ, 486, 581

Gnedin, O. Y., & Ostriker, J. P. 1997, ApJ, 474, 223

Gnedin, O. Y., Lee, H. M., & Ostriker, J. P. 1999, ApJ, 522, 935

Gnedin, N. Y. 2000, ApJ, 542, 535

Gnedin, N. Y., & Abel, T. 2001, New Astronomy, 6, 437

Gnedin, O. Y., Lahav, O., & Rees, M. J. 2001, arXiv:astro-ph/0108034

Gnedin, N. Y., & Kravtsov, A. V. 2006, ApJ, 645, 1054

Gnedin, N. Y., Kravtsov, A. V., & Chen, H.-W. 2008, ApJ, 672, 765

Gnedin, N. Y., Tassis, K., & Kravtsov, A. V. 2009, ApJ, 697, 55

Gnedin, N. Y., & Kravtsov, A. V. 2010, ApJ, 714, 287

Gnedin, N. Y., & Kravtsov, A. V. 2011, ApJ, 728, 88

Gnedin, N. Y., Kravtsov, A. V., & Rudd, D. H. 2011, ApJS, 194, 46

Greif, T. H., Johnson, J. L., Klessen, R. S., & Bromm, V. 2008, MNRAS, 387, 1021

Greif, T. H., Glover, S. C. O., Bromm, V., & Klessen, R. S. 2010, ApJ, 716, 510

Greif, T. H., Springel, V., White, S. D. M., et al. 2011, ApJ, 737, 75

Greif, T. H., Bromm, V., Clark, P. C., et al. 2012, MNRAS, 424, 399

Greif, T. H., Springel, V., & Bromm, V. 2013, arXiv:1305.0823

Griffen, B. F., Drinkwater, M. J., Thomas, P. A., Helly, J. C., & Pimbblet, K. A. 2010, MNRAS, 405, 375

Griffen, B. F., Drinkwater, M. J., Iliev, I. T., Thomas, P. A., & Mellema, G. 2013, MNRAS, 431, 3087

Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633

Guth, A. H., & Pi, S.-Y. 1982, Physical Review Letters, 49, 1110

Haiman, Z., Thoul, A. A., & Loeb, A. 1996, ApJ, 464, 523

Haiman, Z. 2011, Nature, 472, 47

Harris, W. E. & Pudritz, R. E. 1994, ApJ, 429, 177

Harris, W. E. 1996, AJ, 112, 1487

—. 2010, Phil. Trans. Royal Soc. A, in press; arXiv:0911.0798

Hartigan, J. A. & Hartigan, P. M. 1985, The Annals of Statistics, 13, 70

Heger, A., & Woosley, S. E. 2002, ApJ, 567, 532

Heger, A., Fryer, C. L., Woosley, S. E., Langer, N., & Hartmann, D. H. 2003, ApJ, 591, 288

Heger, A., & Woosley, S. E. 2010, ApJ, 724, 341

Hénon, M. 1961, Annales d'Astrophysique, 24, 369

Hoeft, M., Yepes, G., Gottlöber, S., & Springel, V. 2006, MNRAS, 371, 401

Holtzman, J. A., Faber, S. M., Shaya, E. J., Lauer, T. R., Groth, J., Hunter, D. A., Baum, W. A., Ewald, S. P., Hester, J. J., Light, R. M., Lynds, C. R., O'Neil, E. J., & Westphal, J. A. 1992, AJ, 103, 691

Hopkins, P. F., Quataert, E., & Murray, N. 2011, MNRAS, 417, 950

Hosokawa, T., Omukai, K., Yoshida, N., & Yorke, H. W. 2011, Science, 334, 1250

Hubble, E. P. 1926, ApJ, 64, 321

Hummel, J. A., Pawlik, A. H., Milosavljević, M., & Bromm, V. 2012, ApJ, 755, 72

Hummels, C. B., & Bryan, G. L. 2012, ApJ, 749, 140

Hurley, J. R., Pols, O. R., & Tout, C. A. 2000, MNRAS, 315, 543

Hurley, J. R., Shara, M. M., Richer, H. B., King, I. R., Saul Davis, D., Kalirai, J. S., Hansen, B. M. S., Dotter, A., Anderson, J., Fahlman, G. G., & Rich, R. M. 2008, AJ, 135, 2129

Jasche, J., Ciardi, B., & Enßlin, T. A. 2007, MNRAS, 380, 417

Jeans, J. H. 1902, Royal Society of London Philosophical Transactions Series A, 199, 1

Jeon, M., Pawlik, A. H., Greif, T. H., et al. 2012, ApJ, 754, 34

Johnson, J. L., & Bromm, V. 2006, MNRAS, 366, 247

Johnson, J. L., Greif, T. H., & Bromm, V. 2007, ApJ, 665, 85

Johnson, J. L., Dalla, V. C., & Khochfar, S. 2013, MNRAS, 428, 1857

Karlsson, T., Bromm, V., & Bland-Hawthorn, J. 2013, Reviews of Modern Physics, 85, 809

Katz, H., & Ricotti, M. 2013, MNRAS, 1358

Kirby, E. N., Simon, J. D., Geha, M., Guhathakurta, P., & Frebel, A. 2008, ApJ, 685, L43

Kissler-Patig, M., Jordán, A., & Bastian, N. 2006, A&A, 448, 1031

Kitayama, T., Yoshida, N., Susa, H., & Umemura, M. 2004, ApJ, 613, 631

Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18

Kravtsov, A. V., Klypin, A. A., & Khokhlov, A. M. 1997, ApJS, 111, 73

Kravtsov, A. V. 1999, Ph.D. Thesis, New Mexico State University

Kravtsov, A. V. 2003, ApJ, 590, L1

Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, ApJ, 609, 482

Kravtsov, A. V. & Gnedin, O. Y. 2005, ApJ, 623, 650

Kroupa, P. 2001, MNRAS, 322, 231

Kruijssen, J. M. D. 2008, A&A, 486, L21

Krumholz, M. R., Dekel, A., & McKee, C. F. 2012, ApJ, 745, 69

Lamers, H. J. G. L. M., Gieles, M., & Portegies Zwart, S. F. 2005, A&A, 429, 173

Latif, M. A., Schleicher, D. R. G., Schmidt, W., & Niemeyer, J. 2013, arXiv:1306.2745

Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, ApJS, 123, 3

Liddle, A. R. 2009, Annual Review of Nuclear and Particle Science, 59, 95

Lo, Y., Mendell, N. R., & Rubin, D. B. 2001, Biometrika, 88, 767

Lo, Y. 2008, Statistics and Computing, 18, 233

Loeb, A., & Barkana, R. 2001, ARA&A, 39, 19

Machacek, M. E., Bryan, G. L., & Abel, T. 2001, ApJ, 548, 509

Maio, U., Ciardi, B., Dolag, K., Tornatore, L., & Khochfar, S. 2010, MNRAS, 407, 1003

Maio, U., Khochfar, S., Johnson, J. L., & Ciardi, B. 2011, MNRAS, 414, 1145

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388

Marín-Franch, A., Aparicio, A., Piotto, G., Rosenberg, A., Chaboyer, B., Sarajedini, A., Siegel, M., Anderson, J., Bedin, L. R., Dotter, A., Hempel, M., King, I., Majewski, S., Milone, A. P., Paust, N., & Reid, I. N. 2009, ApJ, 694, 1498

Mateo, M. L. 1998, ARA&A, 36, 435

McGaugh, S. S. 2005, ApJ, 632, 859

McLachlan, G. J. 1987, Journal of the Royal Statistical Society. Series C (Applied Statistics), 36, 318

McLaughlin, D. E. 1999, AJ, 117, 2398

McKee, C. F., & Ostriker, E. C. 2007, ARA&A, 45, 565

Mesinger, A., Bryan, G. L., & Haiman, Z. 2009, MNRAS, 399, 1650

Miller, G. E., & Scalo, J. M. 1979, ApJS, 41, 513

Miralda-Escudé, J. 2003, Science, 300, 1904

Moore, B., Diemand, J., Madau, P., Zemp, M., & Stadel, J. 2006, MNRAS, 368, 563

Muratov, A. L., & Gnedin, O. Y. 2010, ApJ, 718, 1266

Muratov, A. L., Gnedin, O. Y., Gnedin, N. Y., & Zemp, M. 2013, arXiv:1212.0909

Muratov, A. L., Gnedin, O. Y., Gnedin, N. Y., & Zemp, M. 2013, arXiv:1212.1157

Murray, N. & Rahman, M. 2010, ApJ, 709, 424

Muthén, B. 2003, Psychological Methods, 8, 369

Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, ApJ, 699, L178

Nakamura, F., & Umemura, M. 2002, ApJ, 569, 549

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

O'Connell, R. W., Gallagher, J. S., Hunter, D. A., & Colley, W. N. 1995, ApJ, 446, L1

O'Shea, B. W., & Norman, M. L. 2007, ApJ, 654, 66

O'Shea, B. W., & Norman, M. L. 2008, ApJ, 673, 14

Oh, S. P., & Haiman, Z. 2002, ApJ, 569, 558

Okamoto, T., Gao, L., & Theuns, T. 2008, MNRAS, 390, 920

Omukai, K., Tsuribe, T., Schneider, R., & Ferrara, A. 2005, ApJ, 626, 627

Orban, C., Gnedin, O. Y., Weisz, D. R., Skillman, E. D., Dolphin, A. E., & Holtzman, J. A. 2008, ApJ, 686, 1030

Ostriker, J. P., & Gnedin, N. Y. 1996, ApJ, 472, L63

Pan, T., Kasen, D., & Loeb, A. 2012, MNRAS, 422, 2701

Pan, T., Loeb, A., & Kasen, D. 2012, MNRAS, 423, 2203

Pawlik, A. H., Milosavljević, M., & Bromm, V. 2011, ApJ, 731, 54

Pawlik, A. H., Milosavljević, M., & Bromm, V. 2013, ApJ, 767, 59

Peebles, P. J. E., & Dicke, R. H. 1968, ApJ, 154, 891

Peng, E. W., Jordán, A., Côté, P., Takamiya, M., West, M. J., Blakeslee, J. P., Chen, C.-W., Ferrarese, L., Mei, S., Tonry, J. L., & West, A. A. 2008, ApJ, 681, 197

Penzias, A. A., & Wilson, R. W. 1965, ApJ, 142, 419

Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431

Pota, V., Forbes, D. A., Romanowsky, A. J., et al. 2013, MNRAS, 428, 389

Powell, L. C., Bournaud, F., Chapon, D., & Teyssier, R. 2013, arXiv:1306.2641

Press, W. H., & Schechter, P. 1974, ApJ, 187, 425

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, Numerical recipes. The art of scientific computing, 3rd ed. (Cambridge: University Press)

Prieto, J. L. & Gnedin, O. Y. 2008, ApJ, 689, 919

Prieto, J., Jimenez, R., & Martí, J. 2012, MNRAS, 419, 3092

Puzia, T. H., Perrett, K. M., & Bridges, T. J. 2005, A&A, 434, 909

Rey, S., Rich, R. M., Sohn, S. T., Yoon, S., Chung, C., Yi, S. K., Lee, Y., Rhee, J., Bianchi, L., Madore, B. F., Lee, K., Barlow, T. A., Forster, K., Friedman, P. G., Martin, D. C., Morrissey, P., Neff, S. G., Schiminovich, D., Seibert, M., Small, T., Wyder, T. K., Donas, J., Heckman, T. M., Milliard, B., Szalay, A. S., & Welsh, B. Y. 2007, ApJS, 173, 643

Rhode, K. L., Zepf, S. E., & Santos, M. R. 2005, ApJ, 630, L21

Ricotti, M. 2002, MNRAS, 336, L33

Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2002, ApJ, 575, 33

Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2002, ApJ, 575, 49

Ricotti, M., & Ostriker, J. P. 2004, MNRAS, 350, 539

Ricotti, M., & Gnedin, N. Y. 2005, ApJ, 629, 259

Ricotti, M., Gnedin, N. Y., & Shull, J. M. 2008, ApJ, 685, 21

Ritter, J. S., Safranek-Shrader, C., Gnat, O., Milosavljević, M., & Bromm, V. 2012, ApJ, 761, 56

Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71

Rubin, V. C., Ford, W. K. J., & . Thonnard, N. 1980, ApJ, 238, 471

Rudd, D. H., Zentner, A. R., & Kravtsov, A. V. 2008, ApJ, 672, 19

Safranek-Shrader, C., Agarwal, M., Federrath, C., et al. 2012, MNRAS, 426, 1159

Savaglio, S., Glazebrook, K., Le Borgne, D., Juneau, S., Abraham, R. G., Chen, H., Crampton, D., McCarthy, P. J., Carlberg, R. G., Marzke, R. O., Roth, K., Jørgensen, I., & Murowinski, R. 2005, ApJ, 635, 260

Saslaw, W. C., & Zipoy, D. 1967, Nature, 216, 976

Scannapieco, E., Schneider, R., & Ferrara, A. 2003, ApJ, 589, 35

Scannapieco, E., Madau, P., Woosley, S., Heger, A., & Ferrara, A. 2005, ApJ, 633, 1031

Schaerer, D. 2002, A&A, 382, 28

Schneider, R., Salvaterra, R., Ferrara, A., & Ciardi, B. 2006, MNRAS, 369, 825

Shapley, H. 1918, ApJ, 48, 154

Sirko, E. 2005, ApJ, 634, 728

Slater, C. T., & Bell, E. F. 2013, arXiv:1306.1829

Smith, B. D., Turk, M. J., Sigurdsson, S., O'Shea, B. W., & Norman, M. L. 2009, ApJ, 691, 441

Smith, R. J., Iocco, F., Glover, S. C. O., et al. 2012, ApJ, 761, 154

Spitler, L. R. & Forbes, D. A. 2009, MNRAS, 392, L1

Spitzer, L. J. 1940, MNRAS, 100, 396

Spitzer, L. J. & Chevalier, R. A. 1973, ApJ, 183, 565

Spitzer, L. 1987, Dynamical Evolution of Globular Clusters (Princeton: Princeton University Press)

Stacy, A., & Bromm, V. 2007, MNRAS, 382, 229

Stacy, A., Greif, T. H., & Bromm, V. 2010, MNRAS, 403, 45

Stacy, A., Greif, T. H., & Bromm, V. 2012a, MNRAS, 422, 290

Stacy, A., Greif, T. H., Klessen, R. S., Bromm, V., & Loeb, A. 2013, MNRAS, 431, 1470

Stewart, K. R., Bullock, J. S., Wechsler, R. H., & Maller, A. H. 2009, ApJ, 702, 307

Strader, J., Brodie, J. P., Cenarro, A. J., Beasley, M. A., & Forbes, D. A. 2005, AJ, 130, 1315

Strader, J., Smith, G. H., Larsen, S., Brodie, J. P., & Huchra, J. P. 2009, AJ, 138, 547

Tassis, K., Kravtsov, A. V., & Gnedin, N. Y. 2008, ApJ, 672, 888

Tegmark, M., Silk, J., Rees, M. J., et al. 1997, ApJ, 474, 1

Tinker, J., Kravtsov, A. V., Klypin, A., et al. 2008, ApJ, 688, 709

Tornatore, L., Ferrara, A., & Schneider, R. 2007, MNRAS, 382, 945

Tremonti, C. A., Heckman, T. M., Kauffmann, G., Brinchmann, J., Charlot, S., White, S. D. M., Seibert, M., Peng, E. W., Schlegel, D. J., Uomoto, A., Fukugita, M., & Brinkmann, J. 2004, ApJ, 613, 898

Trenti, M., Heggie, D. C., & Hut, P. 2007, MNRAS, 374, 344

Trenti, M., & Stiavelli, M. 2009, ApJ, 694, 879

Trenti, M., Stiavelli, M., & Michael Shull, J. 2009, ApJ, 700, 1672

Tumlinson, J., & Shull, J. M. 2000, ApJ, 528, L65

Turk, M. J., Abel, T., & O'Shea, B. 2009, Science, 325, 601

Turk, M. J., Oishi, J. S., Abel, T., & Bryan, G. L. 2012, ApJ, 745, 154

Uhlig, M., Pfrommer, C., Sharma, M., et al. 2012, MNRAS, 423, 2374

van den Bergh, S. 2001, ApJ, 559, L113

Volonteri, M., & Bellovary, J. 2012, Reports on Progress in Physics, 75, 124901

Waters, C. Z., Zepf, S. E., Lauer, T. R., & Baltz, E. A. 2009, ApJ, 693, 463

Wehner, E. M. H., Harris, W. E., Whitmore, B. C., Rothberg, B., & Woodley, K. A. 2008, ApJ, 681, 1233

Whalen, D., Abel, T., & Norman, M. L. 2004, ApJ, 610, 14

Whalen, D., van Veelen, B., O'Shea, B. W., & Norman, M. L. 2008, ApJ, 682, 49

Whalen, D. J., Even, W., Frey, L. H., et al. 2012, arXiv:1211.4979

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341

Whitmore, B. C., Zhang, Q., Leitherer, C., Fall, S. M., Schweizer, F., & Miller, B. W. 1999, AJ, 118, 1551

Wise, J. H., & Abel, T. 2005, ApJ, 629, 615

Wise, J. H., & Abel, T. 2007, ApJ, 671, 1559

Wise, J. H., & Abel, T. 2008, ApJ, 684, 1

Wise, J. H., & Abel, T. 2008, ApJ, 685, 40

Wise, J. H., Turk, M. J., & Abel, T. 2008, ApJ, 682, 745 t

Wise, J. H., & Cen, R. 2009, ApJ, 693, 984

Wise, J. H. 2012, arXiv:1201.4820

Wise, J. H., Turk, M. J., Norman, M. L., & Abel, T. 2012, ApJ, 745, 50

Wise, J. H., Abel, T., Turk, M. J., Norman, M. L., & Smith, B. D. 2012, MNRAS, 427, 311

Wolfe, J. H. 1971, in Technical Bulletin STB 72-2 (San Diego: U.S. Naval Personnel and Training Research Laboratory)

Woo, J., Courteau, S., & Dekel, A. 2008, MNRAS, 390, 1453

Woosley, S. E., & Heger, A. 2012, ApJ, 752, 32

Woosley, S. E., & Weaver, T. A. 1986, ARA&A, 24, 205

Yoshida, N., Abel, T., Hernquist, L., & Sugiyama, N. 2003, ApJ, 592, 645

Yoshida, N., Bromm, V., & Hernquist, L. 2004, ApJ, 605, 579

Yoshida, N., Omukai, K., Hernquist, L., & Abel, T. 2006, ApJ, 652, 6

Yoshida, N., Omukai, K., & Hernquist, L. 2007, ApJ, 667, L117

Yoshida, N., Omukai, K., & Hernquist, L. 2008, Science, 321, 669

Zackrisson, E., Rydberg, C.-E., Schaerer, D., Östlin, G., & Tuli, M. 2011, ApJ, 740, 13

Zackrisson, E., Zitrin, A., Trenti, M., et al. 2012, MNRAS, 427, 2212

Zemp, M., Gnedin, O. Y., Gnedin, N. Y., & Kravtsov, A. V. 2012, ApJ, 748, 54

Zepf, S. E. & Ashman, K. M. 1993, MNRAS, 264, 611

Zepf, S. E., Ashman, K. M., English, J., Freeman, K. C., & Sharples, R. M. 1999, AJ, 118, 752

Zinn, R. 1985, ApJ, 293, 424

Zwicky, F. 1933, Helvetica Physica Acta, 6, 110