Title

Formation of the First Stars Under the Influence of Streaming

Author

Lake, Frederick William

Publication Date

2024

UNIVERSITY OF CALIFORNIA
Los Angeles

Formation of the First Stars Under the Influence of Streaming

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Astronomy and Astrophysics

by

William Lake

2024

ABSTRACT OF THE DISSERTATION

Formation of the First Stars Under the Influence of Streaming

by

William Lake
Doctor of Philosophy in Astronomy and Astrophysics
University of California, Los Angeles, 2024
Professor Smadar Naoz, Chair

JWST has begun to reveal the nature of the small-scale structures of our Universe at unprecedented distances, probing an era of galaxy formation with remarkably different properties. This has offered the opportunity for novel tests of our models of cosmology and galaxy formation. In this thesis, I focused on a defining characteristic of ΛCDM structure formation, the streaming velocity, which has a variety of impacts on small-scale structure. Specifically, I developed models of and observational predictions for a new class of small-scale structures called SIGOs derived from this mechanism, with the aim of enabling future detections of these objects that would place constraints on alternative cosmologies. I also showed that–in contrast to previous expectations–the streaming velocity enhances star formation on a per-galaxy basis in the dwarf regime at high redshifts. Taken together, I showed that the streaming velocity has significant influence on high-redshift small-scale structure formation, and may serve as an important signal of our standard cosmology.

The dissertation of William Lake is approved.

Alice E. Shapley

Steven R. Furlanetto

Tommaso Treu

Smadar Naoz, Committee Chair

University of California, Los Angeles

2024

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGMENTS

First, I want to thank my advisor, Smadar Naoz, for all the help and support over the course of my PhD. I've grown in leaps and bounds as a scientist in the past 6 years and it wouldn't have been possible without you. I couldn't have asked for a better thesis advisor.

I also want to thank the many research collaborators who made this thesis possible. Specifically, I want to thank Claire Williams, Blakesley Burkhart, Federico Marinacci, Mark Vogelsberger, Mike Grudić, Naoki Yoshida, Gen Chiaki, Avi Chen, Yeou Chiou, Kyle Kremer, and Ned Wright, all of whom have given me valuable research advice and guidance over the course of these projects.

I also want to thank the many friends and family members who helped me through this. Firstly, of course, my mom, uncle, sister, and brother, and my whole extended family. In addition, I want to thank the graduate students, especially Briley Lewis, Luke Finnerty, Ben Calvin, and everyone in the Smadar group for all their support and occasionally commiseration. I also want to thank my other friends near and far–particularly Massimo Pascale, Ryan Carlson, Madelyn Stroder, Tal Minear, Tom Gorordo, Carly Robison, Francesco Macagno, Jake Ketchum, and Queso. I'm so grateful for all of you.

I also want to thank my cat, Artemis, for all her moral support. Last but not least, it would be impossible for me to convey my full appreciation for my wife, Amy Zhou-Lake, for all her support through these years. For understanding, for her advice and help, for her cooking, and for a million other things. I am incredibly lucky to have you.

I acknowledge the results presented in this thesis are based upon published works with additional co-authors, and are reprinted here with permission when required. In particular, Chapter 2 is a version of Lake et al. (2021). I thank co-authors Blakesley Burkhart, Federico Marinacci, Mark Vogelsberger, and Kyle Kremer for useful discussions, and co-authors

Smadar Naoz and Yeou S. Chiou for help and guidance with running the numerical simulations. Chapter 3 is a version of Lake et al. (2023a). I thank co-authors Smadar Naoz, Blakesley Burkhart, Federico Marinacci, Mark Vogelsberger, Gen Chiaki, Yeou S. Chiou, Naoki Yoshida, Yurina Nakazato, and Claire E. Williams for useful discussions. Chapter 4 is a version of Lake et al. (2023b). I thank co-authors Smadar Naoz, Federico Marinacci, Blakesley Burkhart, Mark Vogelsberger, Claire E. Williams, Yeou S. Chiou, Yurina Nakazato, and Naoki Yoshida for useful discussions, and Gen Chiaki for creating the numerical simulation package we used for the simulations. Chapter 5 is a version of Lake et al. (2024b). I thank co-authors Smadar Naoz, Federico Marinacci, Blakesley Burkhart, Mark Vogelsberger, Naoki Yoshida, Gen Chiaki, Avi Chen, and Yeou S. Chiou for useful discussions, and Claire E. Williams for providing some of the simulation data used. Chapter 6 is a version of Lake et al. (2024a). I thank co-authors Smadar Naoz, Naoki Yoshida, Claire E. Williams, Blakesley Burkhart, Federico Marinacci, Mark Vogelsberger, and Avi Chen for useful discussions, and Mike Grudić for designing, aiding with, and updating the STARFORGE simulation package used.

This work was made possible through the support of the UCLA Graduate Division through the Dissertation Year Fellowship, as well as the NASA ATP grants 19- ATP19-0020 and 3-ATP23-0149 and the XSEDE and ACCESS AST180056 allocation, as well as the Simons Foundation Center for Computational Astrophysics and the UCLA cluster Hoffman2.

VITA

2017-2019	**B.S., Astronomy and Physics**, University of Arizona, Tucson, AZ
2019–2024	**Teaching Assistant**, Department of Physics and Astronomy, UCLA, Los Angeles, CA
2019–present	**Graduate Student Researcher**, Department of Physics and Astronomy, UCLA, Los Angeles, CA
2021	**M.S., Astronomy and Astrophysics**, UCLA, Los Angeles, CA

PUBLICATIONS

The Supersonic Project: SIGOs, a Proposed Progenitor to Globular Clusters, and their Connections to Gravitational Wave Anisotropies

Lake, W., Naoz, S., Chiou, Y. S., et al. 2021, ApJ, 922, 86

The Supersonic Project: The Early Evolutionary Path of SIGOs

Lake, W., Naoz, S., Burkhart, B., et al. 2023a, ApJ, 943

The Supersonic Project: Star Formation in Early Star Clusters without Dark Matter

Lake, W., Naoz, S., Marinacci, F., et al. 2023b, ApJL, 956

The Supersonic Project: Early Star Formation with the Streaming Velocity

Lake, W., Williams, C. E., Naoz, S., et al. 2024a, ApJ, 973

The Stellar Initial Mass Function of Early Dark Matter-free Gas Objects

Lake, W., Grudić, M.Y., Naoz, S., et al. 2024b, arXiv e-prints, arXiv:2410.02868

H2 Cooling and Gravitational Collapse of Supersonically Induced Gas Objects

Nakazato, Y., Chiaki, G., Yoshida, N., ..., Lake, W., et al. 2022, ApJL, 927, L12

The Supersonic Project: The Eccentricity and Rotational Support of SIGOs and DM GHOSts

Williams, C. E., Naoz, S., Lake, W., et al. 2023, ApJ, 945, 6

The Supersonic Project: Lighting Up the Faint End of the JWST UV Luminosity Function

Williams, C. E., Lake, W., Naoz, S., et al. 2024, ApJL, 960, L16

CHAPTER 1

Introduction

Before the era of Recombination, when dark matter decoupled from the photon field, large-scale overdensities pulled dark matter towards dense regions. This produced dark matter overdensities that were about $10^5\times$ larger than baryon overdensities at Recombination, when baryons decoupled from the photon field (Naoz & Barkana, 2005b). The growth of baryon overdensities and the formation of galaxies was then dominated by these earlier, larger dark matter overdensities. This pre-Recombination period of growth had effects on scales larger than individual halos, as the early collapse of dark matter created relative velocities (known as streaming velocities) between dark matter and baryons. As the baryons cooled after decoupling from the photon field, these streaming motions, with an RMS value of $\mathrm{v_{RMS}} \sim 30$ km s^{-1} at Recombination, became highly supersonic (Tseliakhovich & Hirata, 2010). These streaming velocities were coherent on few-comoving-Mpc scales. Although these motions are formally second order in perturbation theory, the combination of the coherence of these streaming velocities and their large magnitudes significantly influenced the development of small-scale structures in the early Universe and thus impact both the Universe we see today and the early Universe we are beginning to probe with JWST (Tseliakhovich & Hirata, 2010; Visbal et al., 2012; Naoz & Narayan, 2013, 2014; Fialkov & Barkana, 2014).

On small scales ($\lesssim 10^9$ M$_\odot$), as gas collapses into early dark matter overdensities, the effect of these streaming velocities is to advect the gas past its parent halo, offsetting the center of mass of a newly formed gas overdensity from its parent halo (Naoz & Narayan, 2014). This has a few important effects: firstly, at high redshifts (prior to reionization), this substantially reduces the fraction of gas in small halos, an effect that fades with mass and redshift (e.g., Naoz & Narayan, 2014; Popa et al., 2016; Kulkarni et al., 2021; Williams et al., 2023). In addition, this can create dwarf galaxies with low gas fractions and unusual mass distributions, known as Dark Matter + Gas Halos Offset by Streaming, or DM GHOSts, as I will discuss momentarily. Furthermore, streaming velocities can generate even more extreme objects, when the center of mass of the gas overdensity is so far offset from its parent halo through advection that the resulting gas object is entirely outside of its parent halo, at which time it is known as a Supersonically Induced Gas Object, or SIGO (Naoz & Narayan, 2014). These objects form in the mass range of 10^4- few $\times 10^6$ $M_\odot$ (more massive gas overdensities tend to form closer in to their parent halos, and their parent halos are larger), and are generally very deficient in dark matter relative to the overall Universe because, unlike all other structures in that mass range, they form outside of halos. It is important to note that this is a distinctive prediction of ΛCDM structure formation, and thus the detection of a SIGO would constrain non-ΛCDM cosmologies in new ways. Therefore, it is important to understand what a SIGO would look like today, to potentially enable observational confirmation of these objects.

One suggested possibility for the descendant of a SIGO is that it could form a globular cluster. We know that globular clusters (GCs) are comprised of up to $\sim 10^6$ $M_\odot$ of very dense, very old, metal-poor stars–some of the oldest stars in the Universe–and that GCs are deficient in dark matter (Heggie & Hut, 1996; Bradford et al., 2011; Conroy et al., 2011; Ibata et al., 2013). The oldest GCs, which are the most metal-poor, also have distinct properties from younger clusters (Bastian & Lardo, 2018). However, while there are theories for how GCs may have formed, there is no consensus as to the primary formation mechanism of these clusters (Gunn, 1980; Peebles, 1984; Ashman & Zepf, 1992; Harris & Pudritz, 1994; Kravtsov & Gnedin, 2005; Mashchenko & Sills, 2005a; Saitoh et al., 2006; Gray & Scannapieco, 2011; Bekki & Yong, 2012; Kruijssen, 2015; Mandelker et al., 2018).

It has been suggested that GCs may form as the most extreme end of normal star formation, in the dense, hot gas of metal-poor halos (e.g., Elmegreen & Efremov, 1997; Kravtsov & Gnedin, 2005; Shapiro et al., 2010; Grudić et al., 2022), supported by observations of similar processes in high-density gas in the Antennae system (Whitmore & Schweizer, 1995; Whitmore et al., 1999). This would naturally explain why GCs are quite old, as these conditions were more prevalent in the early Universe, with more dynamic events like mergers forming high-density gas, and less metals. However, there are open questions in this model, such as whether it predicts an age distribution for clusters that matches observations, and why the GC luminosity function is similar across a range of environments.

Alternatively, GCs could form inside DM halos, which are then stripped (Peebles, 1984). This naturally lends itself to connections between a GC's properties and the properties of its host galaxy. However, this theory predicts extended dark matter halos around the evolved GCs (remnants of their original tidally stripped halo), which shield the GC from further tidal processes. Therefore, it is difficult to reconcile this theory with observations of stellar tidal tails in today's clusters (e.g., Grillmair et al., 1995; Moore, 1996; Odenkirchen et al., 2003; Mashchenko & Sills, 2005a).

In this thesis, I show that SIGOs can also produce globular-cluster-like objects in regions with substantial streaming. The SIGOs form in low-metallicity regions inside relatively dense accretion flows in the early Universe, and molecular hydrogen cooling allows the gas to reach the densities needed for rapid bursts of star formation. Because our Local Group may have formed in a region of the Universe with significant streaming (and in general, streaming is associated with dense regions), SIGOs are also expected to have been quite abundant in the parts of the Universe where we observe globular clusters (Uysal & Hartwig, 2022). SIGOs' ages and low metallicities lend themselves to comparisions to globular clusters, and I will also detail several other observables that both further this comparison and provide avenues for testing the ubiquity of this mechanism. In this thesis, I'll also connect SIGOs' properties to present-day giant molecular clouds (GMCs), showing that SIGOs have similar properties and thus some modern-day theory relating to cluster formation in molecular clouds can be used to understand SIGOs.

Further, SIGOs serve as a natural testbed for Population III star formation. As SIGOs form from fairly pristine gas[1], Pop III stars are an expected consequence of their evolution. SIGOs form stars at the peak of Pop III star formation, between $z = 10$ and $z = 20$. As I'll overview, SIGOs have masses comparable to the minihalos expected to make up the bulk of Pop III formation, between 10^5 and 10^6 $M_\odot$. The most important difference is that SIGOs are made predominantly of gas, and therefore may reach higher gas surface densities at lower temperatures than classical halos.

I will present new methods for determining the properties of Pop III star clusters in SIGOs, including their star formation rates and stellar masses. I will show that SIGOs are expected to form compact stellar clusters even in the absence of metals, and probe the detectability of star clusters derived from SIGOs with the James Webb Space Telescope (JWST). Present-day globular clusters are found inside galaxies, so I also aim to understand the evolution of SIGOs within their environment, in order to explain why we don't see isolated globular clusters derived from SIGOs.

This simulation-based work also naturally lends itself to studying star formation in minihalos in regions affected by streaming. As mentioned above, the streaming velocity creates dwarf galaxies with gas offset from their dark matter centers of mass at high z (DM GHOSts). This has significant effects on the evolution of these galaxies, as they tend to form shallower gas density profiles and may exhibit higher population-level luminosities (Williams et al., 2023, 2024a). In the fourth project of my thesis, I studied these DM GHOSts in depth, to understand the surface densities and star formation rates in their star-forming regions (their Kennicutt-Schmidt relations/KSRs; Schmidt, 1959) as well as their star formation rate den-

---

[1] Though some metal enrichment from nearby star-forming halos may be possible

sities. Regions with substantial streaming tend to form stars in objects with higher masses, with low-mass structures delayed in their collapse to form stars (e.g., Hegde & Furlanetto, 2023; Williams et al., 2024a), so this study permitted me to understand whether enhanced accretion flows onto high-mass halos led to higher individual luminosities for galaxies in regions of significant streaming. Furthermore, I compared the strength of this effect to the strength of streaming's suppression of minihalo formation, to understand the impact of streaming on observed quantities such as star formation rate densities. With these observables in hand, I suggest a strategy for using JWST to constrain the value of the streaming velocity in a patch of the sky, permitting an observational study of streaming.

My thesis is organized as follows: in Chapter 2 I detail a new semi-analytic model for the abundance of SIGOs formed simply through adiabatic cooling, to understand their abundance variations on the sky and connect them to potential observables. In Chapter 3 I further this work by adding molecular hydrogen cooling, in order to understand the evolution of SIGOs as molecular clouds in the context of modern theory on GMCs. Chapter 4 takes this a step further by adding explicit star formation and presenting the first simulation to explicitly show the presence and abundance of stars within a star-forming SIGO. Chapter 5 generalizes this, showing the impacts of the streaming velocity that gives rise to SIGOs on star formation in both DM GHOSts and SIGOs and suggesting key observables for understanding streaming. Chapter 6 presents the first sub-solar-resolution study of star formation in SIGOs, as well as the first study of the effects of mechanical feedback in SIGOs, showing the stellar IMF and overall star cluster derived from these objects. Finally, I summarize my results in Chapter 7.

CHAPTER 2

SIGOs, a Proposed Progenitor to Globular Clusters, and their Connections to Gravitational Wave Anisotropies

2.1 Introduction

Globular clusters (GCs) are very old ($\sim$ 13 Gyr, e.g., Trenti et al., 2015) structures with masses between $\sim 10^5 - 10^6 M_\odot$ (e.g., Elmegreen & Efremov, 1997; Fall & Zhang, 2001; McLaughlin & Fall, 2008; Elmegreen, 2010). Their high stellar densities and low metallicities make them a promising nurturing ground for gravitational wave sources via few body dynamics (e.g., Portegies Zwart & McMillan, 2000; Wen, 2003; O'Leary et al., 2006; Rodriguez et al., 2015, 2016b; Chatterjee et al., 2017; Kremer et al., 2020; Rodriguez et al., 2021).

Significantly, observations suggest that GCs contain little to no dark matter (e.g., Heggie & Hut, 1996; Bradford et al., 2011; Conroy et al., 2011; Ibata et al., 2013). These observations pose a challenge to the formation of these objects in the context of hierarchical structure formation. Accordingly, different GC formation scenarios exist in the literature. One popular mechanism is that GCs formed as a byproduct of active star formation in galaxy discs (e.g., Elmegreen, 2010; Shapiro et al., 2010; Kruijssen, 2015), for example as a result of strong

---

The contents of this chapter appeared in Lake et al. (2021)

shocks when gas is compressed during galaxy mergers, as first proposed by Gunn (1980). The discovery of many massive young star clusters in the interacting Antennae system (e.g., Whitmore & Schweizer, 1995; Whitmore et al., 1999) supports this idea. Furthermore, this scenario has also been incorporated into cosmological hierarchical structure formation models (e.g. Harris & Pudritz, 1994; Ashman & Zepf, 1992; Kravtsov & Gnedin, 2005; Muratov & Gnedin, 2010). However, this paradigm is challenged by observations of nuclear star clusters that resemble GCs (e.g., in total mass and core/half-light radii), which imply that some GC-like structures may form inside dark matter (DM) halos and thus may have a DM halo origin (see for example Böker et al., 2004; Walcher et al., 2005, 2006; Brown et al., 2014).

Another popular theory is that GCs initially formed inside dark matter halos (as suggested by Peebles, 1984), but that these halos were later stripped by the tidal field of their host galaxies, leaving the central parts deficient of dark matter (e.g. Bromm & Clarke, 2002; Mashchenko & Sills, 2005b; Saitoh et al., 2006; Bekki & Yong, 2012). However, some GCs are observed to have stellar tidal tails, which is difficult to explain in the context of this scenario. If the objects have extended dark matter halos, the halos should have shielded them from forming tidal tails (e.g., Grillmair et al., 1995; Moore, 1996; Odenkirchen et al., 2003; Mashchenko & Sills, 2005b).

Recently, Naoz & Narayan (2014) proposed a formation pathway for GCs that relies on the relative motion between baryons and DM at the time of recombination, known as the stream velocity (Tseliakhovich & Hirata, 2010; Tseliakhovich et al., 2011). In the standard model of structure formation, due to the baryon-photon coupling, dark matter began to collapse to form overdensities far more efficiently than baryons. By the time of recombination, when baryons decoupled from photons, baryon overdensities were about 5 orders of magnitude smaller than dark matter overdensities (e.g. Naoz & Barkana, 2005b). This meant that the existing dark matter overdensities dominated the dynamics of baryon overdensity formation. Tseliakhovich & Hirata (2010) showed that in addition to the difference in amplitude between baryonic and DM overdensities, there was a significant difference in their velocities in the period following recombination. As the baryons cooled, the typical relative velocity between dark matter and baryons (about 30 km s^{-1}) became supersonic. They also showed that this relative velocity was coherent on scales of $\sim 2-3$ Mpc, allowing it to be modelled as a stream velocity on these scales.

This stream velocity suppresses formation of the earliest baryonic structures, such as minihalos, and therefore has an effect on early star formation and on the temperature of the early Universe. This has been explored in a variety of studies, having such diverse impacts as creating temperature-induced fluctuations in the cosmological 21 cm line (e.g., Dalal et al., 2010; Visbal et al., 2012; McQuinn & O'Leary, 2012; Cain et al., 2020), enhancing primordial black hole formation (e.g., Tanaka et al., 2013; Tanaka & Li, 2014; Latif et al., 2014; Hirano et al., 2017b; Schauer et al., 2017), and even creating primordial magnetic fields through temperature fluctuation-induced vorticity (Naoz & Narayan, 2013). Studies also show that this stream velocity has major impacts on the number densities of halos (Asaba et al., 2016; Tanaka et al., 2013; Tanaka & Li, 2014; Bovy & Dvorkin, 2013; O'Leary & McQuinn, 2012; Naoz et al., 2012; Fialkov et al., 2012; Tseliakhovich & Hirata, 2010; Maio et al., 2011), as well as the overall gas fraction in halos (Asaba et al., 2016; Richardson et al., 2013; Maio et al., 2011; O'Leary & McQuinn, 2012; Naoz et al., 2013; Greif et al., 2011; Fialkov et al., 2012; Naoz et al., 2012; Tseliakhovich et al., 2011; Dalal et al., 2010) and the size of halos able to retain gas at each redshift (Naoz et al., 2013). In addition, the stream velocity impacts the gas density and temperature profiles (Richardson et al., 2013; O'Leary & McQuinn, 2012; Fialkov et al., 2012; Greif et al., 2011; Maio et al., 2011; Liu & Wang, 2011; Druschke et al., 2020), and the halo mass threshold at which star formation occurs (Bovy & Dvorkin, 2013; O'Leary & McQuinn, 2012; Fialkov et al., 2012; Greif et al., 2011; Maio et al., 2011; Liu & Wang, 2011; Schauer et al., 2019).

The aforementioned proposal by Naoz & Narayan (2014) suggested that this stream velocity effect could lead to a possible formation mechanism for globular clusters. They found that a stream velocity of sufficient magnitude between a dark matter and baryonic overdensity could create a spatial offset between the collapsing baryonic overdensity and its parent dark matter halo. In certain instances, this effect is large enough to cause the baryonic overdensity to collapse outside the parent halo's virial radius, allowing it to be separated from the parent halo's gravitational influence entirely. This would create a baryonic clump depleted of dark matter in a similar mass range to present-day globular clusters. In addition, such a baryon clump would likely have a low metallicity attributable to its early formation, possibly consistent with that of the low-metallicity population of GCs.

These objects, known as Supersonically Induced Gas Objects (SIGOs), have since been found in follow-up simulations (Popa et al., 2016; Chiou et al., 2018, 2019, 2021). However, there are still many outstanding questions about these objects. Notably, their large-scale abundance distribution has not been studied yet. This large-scale abundance is expected to be correlated with the magnitude of the stream velocity (Popa et al., 2016). SIGO abundances determine their possible global effect on reionization as well as the distribution of the very first star clusters and possibly GCs.

The question of the connection between SIGO abundances and GC abundances has particular relevance given the recent detections of gravitational wave (GW) emission from merging stellar-mass black hole (BH) binaries by LIGO-Virgo that have expanded our ability to sense the Universe (e.g., Abbott et al., 2016, 2017b). It remains challenging to explain the formation channels of these sources, but recent studies have emphasized the significant contribution of dynamical formation channels in dense stellar environments to the overall population of GW signals (e.g., Portegies Zwart & McMillan, 2000; Wen, 2003; O'Leary et al., 2006, 2009, 2016; Kocsis & Levin, 2012; Antonini et al., 2015; Rodriguez et al., 2016a; Chatterjee et al., 2017; Stone & van Velzen, 2016; Hoang et al., 2018; Stephan et al., 2019; Kremer et al., 2020; Wang et al., 2021). Accordingly, GCs have been suggested as a primary source of black-hole binary (BBH) mergers (e.g., Rodriguez et al., 2021). Should this be the case, and should SIGOs indeed be connected to GCs, then SIGO abundances should be connected to the abundance of BBH mergers.

In this paper, we study the large-scale abundances of SIGOs using a combination of analytical and numerical methods. This is a challenging task due the following reasons:

- The stream velocity is constant only on scales of a few Mpc (e.g., Tseliakhovich & Hirata, 2010). Thus, the implementation of the initial conditions in numerical simulations can be done self-consistently only on small box simulations (e.g., Naoz et al., 2012, 2013; McQuinn & O'Leary, 2012; O'Leary & McQuinn, 2012; Stacy et al., 2011; Schauer et al., 2019; Chiou et al., 2018, 2019, 2021). In these small box simulations, the stream velocity is implemented as a uniform boost along one axis.

- Even at the event of successfully implementing initial conditions that allow for the stream velocity to change coherently over large scales (>> few Mpc), the simulation will still need to resolve objects at the order of 10^4 $M_\odot$ with at least around 100 particles, requiring unrealistic numerical resources.

We therefore take a combined approach, utilizing analytical and numerical tools. We use a series of small-box `AREPO` runs (side length 2 Mpc) with varying stream velocity magnitudes and compare them to analytical calculations. Using simulation results to derive an abundance normalization factor, we create a fully analytic model of the spatial variation of SIGO abundances. If SIGOs are indeed linked to GCs, they can host gravitational wave sources, which allows us to hypothesize a spatial variation in GC and GW abundances related to that of SIGOs.

For this work, we have assumed a ΛCDM cosmology with $\Omega_\Lambda = 0.73$, $\Omega_{\rm M} = 0.27$, $\Omega_{\rm B} = 0.044$, $\sigma_8 = 1.7$, and $h = 0.71$.

This paper is organized as follows: we first provide an overview of our simulations in Section 2.2.1. We then discuss our analytic model in Section 2.2.2. We provide a comparison between the simulation and model results in Section 2.3, as well as connecting our model results to the real-world abundance of GCs. We consider the implications of these results to gravitational wave abundances in Section 2.4. We discuss our model results in Section 2.5. Finally, we show how we normalized our analytic model to simulations in Appendix A and provide an analytic approximation to our model in Appendix B.

2.2 Methods

We use a combination of analytical and numerical methods described below to analyze the large-scale SIGOs number density.

2.2.1 Simulations

We present three simulations with the moving-mesh code `AREPO` (Springel, 2010) in a 2 Mpc box[1] with 512^3 DM particles of mass $M_{\rm DM} = 1.9 \times 10^3$ $\mathrm{M}_\odot$ and 512^3 Voronoi mesh cells with $M_{\rm b} = 360$ $\mathrm{M}_\odot$, evolved from $z = 200$ to $z = 20$. These runs had stream velocities of $v_{\rm bc} = 1\sigma_{\rm vbc}$, $2\sigma_{\rm vbc}$, and $3\sigma_{\rm vbc}$ where $\sigma_{\rm vbc}$ is the rms value of the stream velocity–the relative velocity of the gas component with respect to the dark matter component. $\sigma_{\rm vbc} = 5.9$ km sec^{-1} at $z = 200$. We note that these runs do not include radiative cooling. Cooling does not significantly change the physical properties of SIGOs, and only moderately affects the classical objects (i.e., DM halos with gas), as shown in Chiou et al. (2021).

The initial conditions for our cosmological simulations were generated using transfer functions calculated using a modified `CMBFAST` code (Seljak & Zaldarriaga, 1996) that takes into account the first-order correction of scale-dependent temperature fluctuations (Naoz &

---

[1] Note that the simulated abundances of SIGOs at the relevant masses were shown to converge for small (few Mpc) boxes (e.g., Popa et al., 2016).

Barkana, 2005b). These transfer functions also include second-order corrections to the equations presented in Tseliakhovich & Hirata (2010) that describe the evolution of the stream velocity. There are two transfer functions, one for the baryons and one for the dark matter, as it was pointed out that the gas fraction evolution strongly depends on the baryons' initial conditions (e.g., Naoz et al., 2009, 2011, 2013; Park et al., 2020). The stream velocity was implemented in the initial conditions as a uniform boost to the gas in the x-direction, as in Popa et al. (2016). Initial conditions were generated at $z = 200$.

For this paper, we use the object classifications described in Chiou et al. (2018). The first step in our identification of SIGOs is to identify dark matter-primary objects (dark matter halos) using a Friends-of-Friends (FOF) algorithm with a linking length that is 20% of the mean particle separation on the DM component of the simulation output[2], about 780 comoving pc. This algorithm identifies the location of the DM halos in the simulation box. It also calculates the virial radius for each halo, assuming sphericity for simplicity (although DM halos show distinct triaxiality e.g., Sheth et al., 2001; Lithwick & Dalal, 2011; Vogelsberger & White, 2011; Schneider et al., 2012; Vogelsberger et al., 2020). Next, we find gas-primary objects using the same FOF algorithm run only on the gas component of the simulation output. We require that gas primary objects contain at least 32 particles to be considered a SIGO (Chiou et al., 2021). Because these objects tend to be more attenuated, each gas-primary object is fit to an ellipsoid, by identifying an ellipsoidal surface that encloses every particle in the gas object (Popa et al., 2016). We then tighten these ellipsoids by shrinking their axes by 5% until either 20% of their particles have been removed, or until the ratio of the axes lengths of the tightened ellipsoid to that of the original ellipsoid is greater than the ratio of the number of gas cells contained in each, as in Popa et al. (2016). Because many of these gas-primary objects are actually just the gas component of the previously mentioned DM halos, SIGOs are then defined as gas-primary objects which have a gas fraction above

---

[2] This linking length was shown to give converging values of object abundances by Naoz et al. (2011)

40%[3], and are outside the virial radius of the nearest dark matter halo.

2.2.2 Analytic Model

Our analytic model, in contrast to our simulations, ran on a large-scale box (~ 1365 Mpc on a side), composed of grid cells that were 3 Mpc on a side. Within each grid cell, as in the simulations, the relevant scales are small enough that $v_{\rm bc}$ is approximately constant. We assigned a value of $v_{\rm bc}$ for each cell using an algorithm for generating Maxwell distributed random fields given a power spectrum of their spatial fluctuations (Brown, 2013) which we calculated using a modified version of `CMBFAST` (Seljak & Zaldarriaga, 1996), that includes the spatial perturbations of the baryon sound speeds, as outlined in Naoz & Barkana (2005b).

Stream velocities follow a Maxwell distribution with scale parameter $\sigma = \sigma_{\rm vbc}/\sqrt{3}$, and a known power spectrum given by the output of `CMBFAST` described above. Using the spectral distortion method outlined in Brown (2013), we generated a Maxwell-distributed random field of velocities in a grid of $456 \times 456 \times 456$ cells, with each cell being 3 Mpc on a side (small enough such that the stream velocity within each cell is coherent). This grid was generated with the computed power spectrum through the following recursive steps:

1. We generated a Gaussian random field using the computed power spectrum of stream velocity fluctuations $P_{\rm vbc}$ as the input power spectrum $P_{\rm I}$.
2. We then transformed this Gaussian random field to a Maxwell distributed random field using a quantile transform, and calculated the output power spectrum $P_{\rm F}$ of that field.
3. If this power spectrum output is consistent with the target output, we accepted this Maxwell-distributed field as our velocity grid. Otherwise, we set our input power spectrum

---

[3] Note that 40% here represents a somewhat arbitrary compromise between doubling the cosmic baryon fraction and the estimated baryon content of GCs, 50% or more.

$$P_{\mathrm{I,new}} = \frac{P_{\mathrm{F}}(k)}{P_{\mathrm{I}}(k)} \times P_{\mathrm{vbc}}(k) \ , \tag{2.1}$$

and returned to the first step using this new input power spectrum.

This yields a grid of cells with constant stream velocity to be used in the density evolution equations that follow.

For completeness, we provide the full set of differential equations of the perturbation theory. We solve the differential equations for the dimensionless overdensities of both the dark matter, δ_{dm}, and the baryons, δ_{b} in the presence of the relative velocity between dark matter and baryons in small regions within which the velocity is coherent (few Mpc, e.g. Tseliakhovich & Hirata, 2010; Tseliakhovich et al., 2011; Naoz et al., 2013). These can be expressed by the following set of coupled equations:

$$\ddot{\delta}_{\mathrm{dm}} + 2H\dot{\delta}_{\mathrm{dm}} - f_{\mathrm{dm}}\frac{2i}{a}\mathbf{v}_{\mathrm{bc}}\cdot\mathbf{k}\dot{\delta}_{\mathrm{dm}} = \frac{3}{2}H_0^2\frac{\Omega_{\mathrm{m}}}{a^3}(f_{\mathrm{b}}\delta_{\mathrm{b}} + f_{\mathrm{dm}}\delta_{\mathrm{dm}}) + (\frac{\mathbf{v}_{\mathrm{bc}}\cdot\mathbf{k}}{a})^2\delta_{\mathrm{dm}} \ , \tag{2.2}$$

$$\ddot{\delta}_{\mathrm{b}} + 2H\dot{\delta}_{\mathrm{b}} = \frac{3}{2}H_0^2\frac{\Omega_{\mathrm{m}}}{a^3}(f_{\mathrm{b}}\delta_{\mathrm{b}} + f_{\mathrm{dm}}\delta_{\mathrm{dm}}) - \frac{k^2}{a^2}\frac{k_{\mathrm{b}}\bar{T}}{\mu}(\delta_{\mathrm{b}} + \delta_{\mathrm{T}}) \ , \tag{2.3}$$

where $\mathbf{k}$ is the comoving wavenumber vector, a is the scale factor, μ is the mean molecular weight, δ_{T_γ} is the photon temperature fluctuations, $\bar{T}$ and f_{b} and f_{c} are the baryon and DM fractions, respectively.

We also include the baryons' temperature fluctuations δ_{T} which include scale-dependent temperature time evolution (according to Naoz & Barkana, 2005b) in our calculations. These evolve according to

$$\frac{\delta_{\mathrm{T}}}{dt} = \frac{2}{3}\frac{\delta_{\mathrm{b}}}{dt} + \frac{x_{\mathrm{e}}(t)}{t_\gamma}a^{-4}\left\{\delta_\gamma\left(\frac{\bar{T}_\gamma}{\bar{T}} - 1\right) + \frac{\bar{T}_\gamma}{\bar{T}}\left(\delta_{T_\gamma} - \delta_{\mathrm{T}}\right)\right\} \ .$$

As was discussed in Naoz et al. (2012) and Naoz & Narayan (2014), the stream velocity intro-

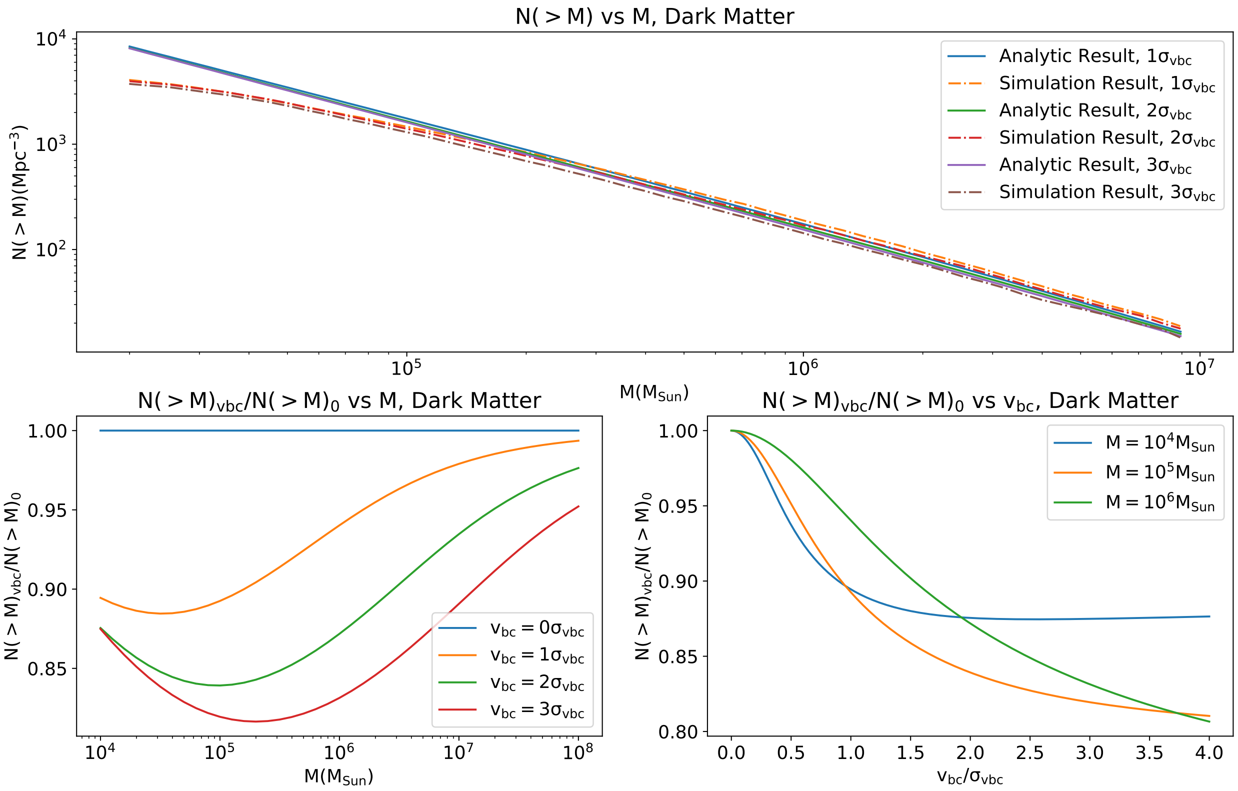


Figure 2.1: **The effect of the stream velocity on the "classical" dark matter halos.** In the top panel, we consider the number abundance $N(>M)_{\rm Halo,vbc}$ as a function of the DM halo mass, and compare the analytical calculation (solid lines) and simulation results (dot-dashed lines). In the bottom panels, we consider the analytical reduction fraction of the number of halos due to the stream velocity (i.e., $N(>M)_{\rm Halo,vbc}/N(>M)_{Halo,0}$, where subscript "0" indicates $v_{\rm bc}=0$). We present this fraction as a function of DM halo mass, for different stream velocity values (bottom left panel) and as a function of the stream velocity for different halo masses (bottom right panel). Results are for $z=20$.

duces a phase shift between baryon and DM overdensities. This phase shift creates a spatial separation between baryonic overdensities and their parent DM overdensities. Because increasing stream velocities create increasing phase shifts (which in turn create increasing spatial separations between the overdensities), sufficiently high stream velocities can cause baryonic clumps to collapse outside of the virial radii of their parent DM overdensities. This allows them to survive as independent, DM-depleted objects. Simulations suggest that these objects could potentially evolve into present-day globular clusters (Chiou et al., 2019).

We adopt the generalized Press-Schechter formalism (Press & Schechter, 1974) to allow for non-spherical halos. This model, based on Gaussian random fields and including linear growth, allows us to calculate abundances of objects at different masses. The formalism depends on two functions $\sigma(M, z)$ and δ_c. In this case, $\sigma^2(M, z)$ is the variance, calculated from the power spectrum, as a function of halo mass at a given redshift, and δ_c is the critical collapse overdensity[4]. The comoving number density of halos of mass M at redshift z in this model is given by

$$\frac{dn}{dM} = \frac{\rho_0}{M} f_{ST} \left| \frac{dS}{dM} \right| , \tag{2.4}$$

where we have used the Sheth et al. (2001) mass function that both fits simulations, and includes non-spherical effects on the collapse. The function f_{ST} is the fraction of mass in halos of mass M:

$$\begin{aligned} f_{ST}(\delta_c, S) &= A' \frac{\nu}{S} \sqrt{\frac{a'}{2\pi}} \left[1 + \frac{1}{(a'\nu^2)^{q'}} \right] \exp\left(\frac{-a'\nu^2}{2} \right) \\ \text{and} \quad \nu &= \frac{\delta_c}{\sigma} = \frac{\delta_c}{\sqrt{S}} \ . \end{aligned} \tag{2.5}$$

We use best-fit parameters $A' = 0.75$ and $q' = 0.3$ (Sheth & Tormen, 2002). By evolving the power spectrum analytically, we can use this model to effectively predict halo abundances, as shown in Section 2.3.

It not straightforward to extrapolate the DM halo Press-Schechter formalism to SIGOs because these objects are non-spherical. Significantly, unlike a DM overdensity that grows due to its own gravity, SIGOs by themselves do not have enough material to grow independently (Peebles, 1969), and instead are still coupled to the DM potential wells (Naoz & Narayan, 2014). In the presence of the stream velocity the gas does not accumulate over the DM overdensities (e.g., Naoz et al., 2012; Popa et al., 2016) and some of it results in the forma-

---

[4] We note that in general δ_c is a function of the redshift (e.g. Naoz et al., 2006; Naoz & Barkana, 2007) because the baryons have smoother initial conditions. However, since we normalize our abundances according to the simulations, we neglect the redshift contribution.

tion of SIGOs. Thus, we postulate that the SIGOs overdensity may be related to the DM underdensity, and could obey a simple relation such as:

$$\left.\frac{dN(v_{\rm bc})}{dM}\right|_{\rm SIGO} \propto \left(\left.\frac{dN(v_{\rm bc}=0)}{dM}\right|_{\rm Halo} - \left.\frac{dN(v_{\rm bc})}{dM}\right|_{\rm Halo}\right) \ , \tag{2.6}$$

where the proportionality here is aimed to emphasize that the non-linear effects result in a normalization factor (see Appendix A for details). In other words, some of the gas that does not fall onto the DM potential wells does not become SIGOs. Motivated by simulations, we find a simple normalization power law in $\mathrm{v}_{\rm bc}$ (equation [A.1]).

By evolving the power spectrum analytically, we can determine an analytical SIGO abundance, and thereby determine properties of their distribution on large scales (i.e., the sky). We can then compare this predicted spatial variation of SIGO abundances to observations of globular cluster abundances to test the hypothesis that these objects are their dominant formation mechanism.

2.3 Comparison between analytical and numerical calculations

2.3.1 Dark Matter Halos

In Figure 2.1 we show the agreement between the analytic model based on the generalized Press-Schechter formalism and simulation results for DM halos, for various values of the stream velocity effect at $z = 20$. As depicted, the simulations and the analytical calculations are consistent for $M \approx 10^5 - 10^7 M_\odot$, the region in which the simulation results are expected to be less sensitive to resolution effects (requiring a minimum of 100 DM particles per halo). Thus, because of the limited resolution of the simulation, we observe fewer small-mass halos in the simulation than our analytic model would predict, as expected.

The bottom panels of Figure 2.1, present only analytical calculations. We show the fraction of DM number density $N(>M)_{\rm Halo}$ with the stream velocity compared to the number density without the stream velocity. In other words:

$$f_{\rm DM} = \frac{N(>M)_{\rm Halo,vbc}}{N(>M)_{\rm Halo,0}} \ , \tag{2.7}$$

where the subscript "0" means $v_{\rm bc} = 0$ and the cumulative comoving number density of DM haloes $N(>M)$ is given by

$$N(>M)_{\rm Halo} = \int_{\rm M}^{\infty} \frac{dN_{\rm Halo}}{dM} dM \ . \tag{2.8}$$

The bottom left panel of Figure 2.1 shows Equation (2.7) as a function of the DM halo mass for different stream velocity effects. The bottom right panel shows Equation (2.7) as a function of $v_{\rm bc}$, for different DM halo masses. As depicted, the higher stream velocities reduce the abundance of dark matter halos, particularly those of mass $\sim 10^5 - 10^6 M_\odot$, with the total reduction in abundance with stream velocities on the order of $\sigma_{\rm vbc}$ being on the order of tens of percent.

In Figure 2.2 we depict the DM halo fluctuations due to the stream velocity effects on a large scale. In particular, we show

$$\delta_{\rm Halo} = \frac{N(>M)_{\rm Halo,vbc} - \langle N(>M)_{\rm Halo}\rangle}{\langle N(>M)_{\rm Halo}\rangle} \ , \tag{2.9}$$

with $N(>M)_{\rm Halo}$ as defined in Equation (2.8), and using the velocity field generated with the method described in Section 2.2.2.

2.3.2 SIGOs

Using the tightly fitted ellipsoid method to find the SIGOs (see Section 2.2) we find $8, 75$ and 188 SIGOs for the $1, 2$ and $3\sigma_{\rm vbc}$ simulation runs. From this simulation data, we construct $N(>M)_{\rm SIGO,sim}$. As discussed in Section 2.2.2, we also calculate the abundance of SIGOs as

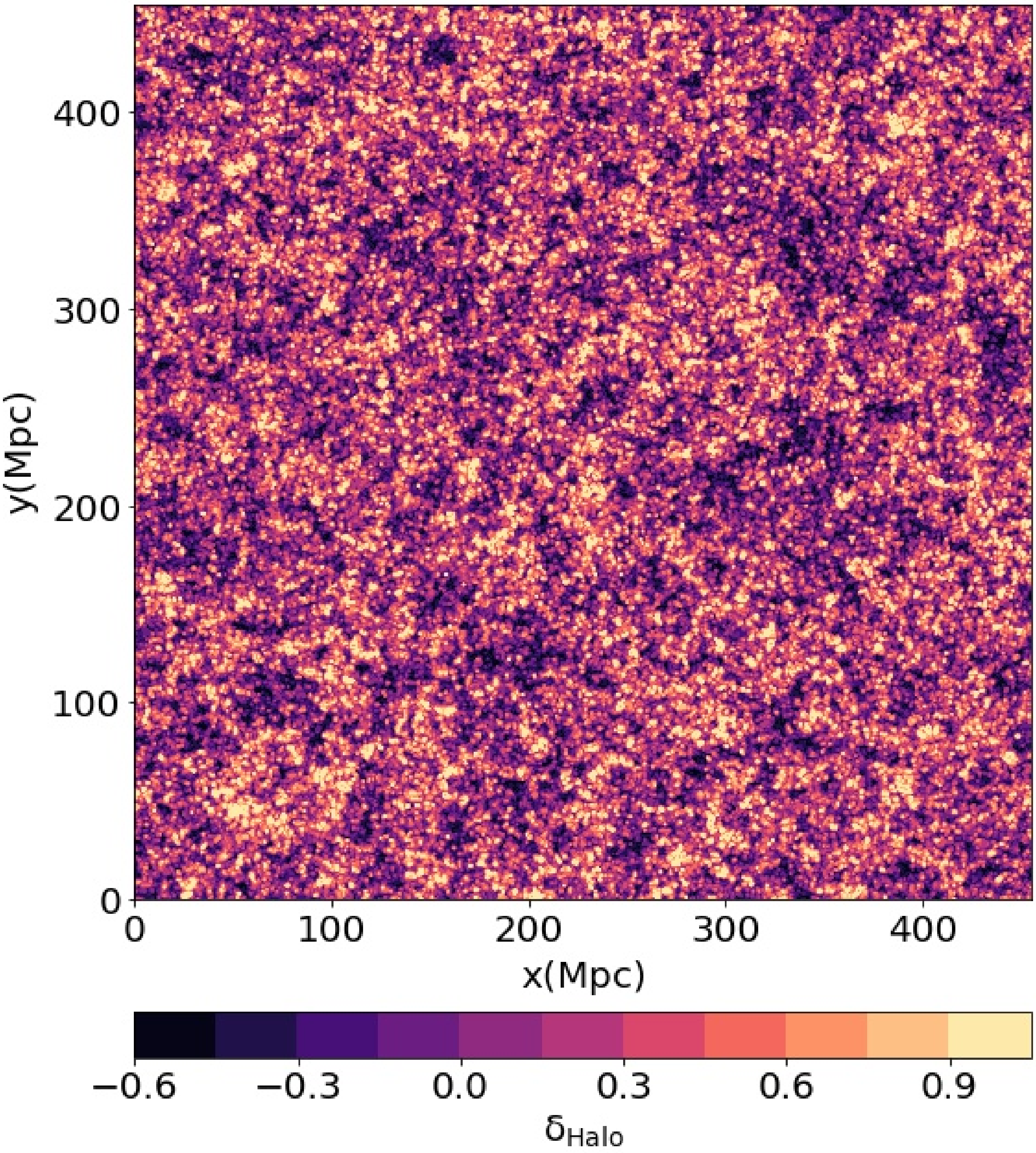


Figure 2.2: Map of DM halo density contrast at $z = 20$, showing a 456 Mpc × 456 Mpc × 1 Mpc box (a subset of our model box, to enhance the visibility of structures).

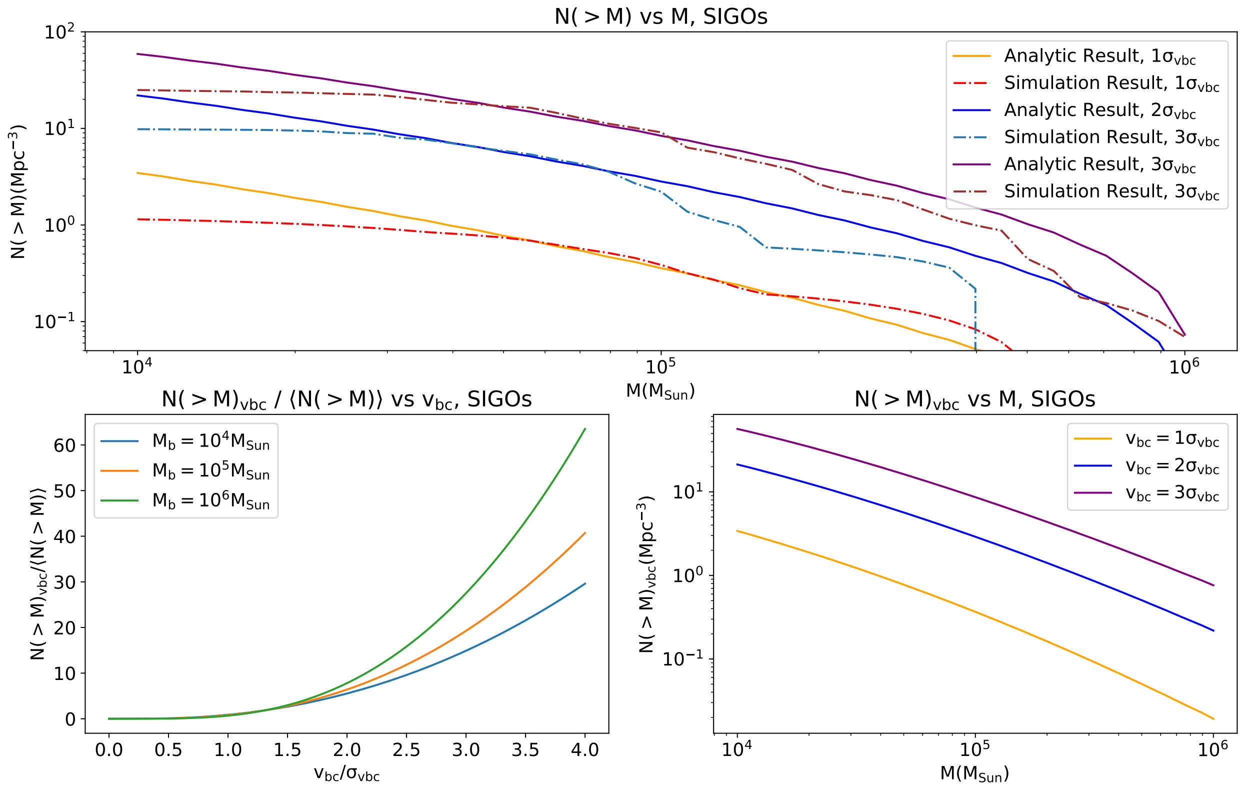


Figure 2.3: **The effect of the stream velocity on SIGO abundances.** In the top panel, we consider the number abundance $N(>M)_{\mathrm{SIGO,vbc}}$ as a function of SIGO mass, and compare the analytical calculation (solid lines) and simulation results (dot-dashed lines). See text for an explanation of how SIGO abundances are estimated in the analytical calculations and how they are defined in the simulations, and see Equation (2.13) for a definition of $\langle N(>M)_{\mathrm{SIGO}}\rangle$. In the bottom left panel, we consider the number density of SIGOs as a function of the stream velocity magnitude for different SIGO masses, based on analytical calculations. In the bottom right panel, we depict the number density of SIGOs as a function of mass for different stream velocities, based on analytical calculations, without integrating from a maximum mass. Results are for $z = 20$.

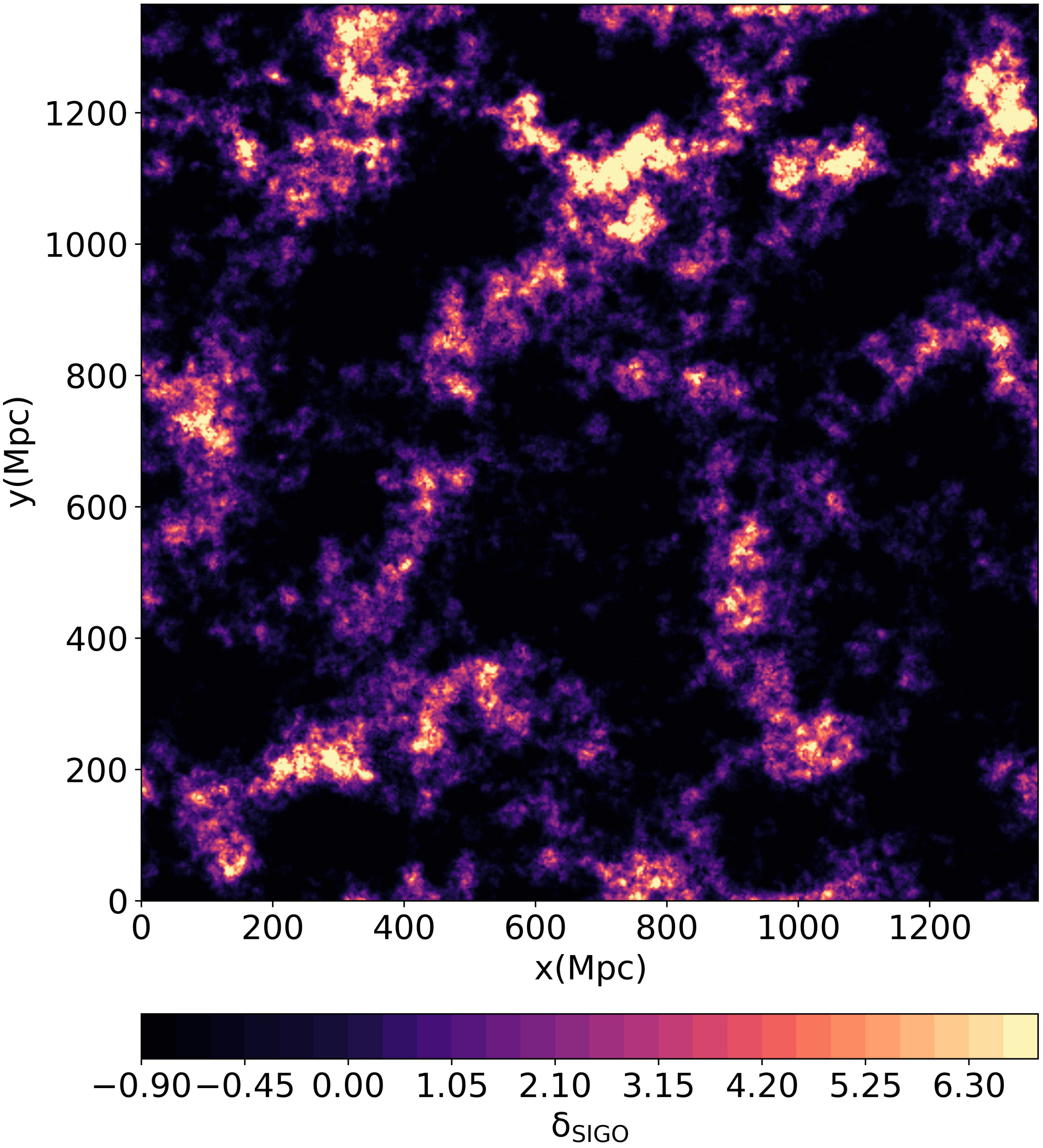


Figure 2.4: Map of SIGO density contrast at z=20. Specifically, we plot Eq. (2.14), for a minimum mass of $10^5\ M_\odot$. Here the color scale was capped at $\delta_{\rm SIGO} = 7.0$ to enhance visibility of smaller fluctuations; the true maximum in this plot is $\delta_{\rm SIGO} \approx 30$(which represents a rare fluctuation).

a function $\sigma_{\rm vbc}$ analytically using Equation (2.6). We normalize the analytical results to the simulation results as described in appendix A. The comparison yields a simple functional form for SIGO abundances as function of mass and stream velocity, i.e.,

$$\left.\frac{dN(u_{\rm vbc})}{dM}\right|_{\rm SIGO} \approx \left(\frac{u_{\rm vbc}}{20.76}\right)^{2.43} \times \left(\left.\frac{dN(u_{\rm vbc}=0)}{dM}\right|_{\rm Halo} - \left.\frac{dN(u_{\rm vbc})}{dM}\right|_{\rm Halo}\right) \ , \tag{2.10}$$

with $u_{\rm vbc} = v_{\rm bc}/\sigma_{\rm vbc}$. At the time of recombination, for example, $\sigma_{\rm vbc} = 30$ km sec^{-1}, which corresponds to $u_{\rm vbc} = 1$.

Figure 2.3, top panel, shows the comparison between the SIGOs model and simulation results (as in Figure 2.1). In the analytical model, we have integrated the number density dN/dM (see Equation [2.10]) only from a defined cutoff mass ($\mathrm{M}_{\rm cutoff} = M_{\rm SIGO,max} = 1.1 \times 10^6 M_\odot$) for the purpose of comparing to simulations. This cutoff mass is motivated by Naoz & Narayan (2014), who found that SIGOs have an upper mass limit of around a few $\times 10^6$ $\mathrm{M}_\odot$, above which they are incapable of escaping their parent DM halo. Because of this, we select $\mathrm{M}_{\rm cutoff}$ as the mass of the largest SIGO observed in any of our simulations. In particular, the cumulative number density can be expressed as:

$$N(>M_{\rm min}) = \int_{\mathrm{M}_{\rm min}}^{M_{\rm SIGO,max}} \left(\frac{dN}{dM}\right)_{\rm SIGO} dM \ , \tag{2.11}$$

where dN/dM is defined in Eq. (2.10). We show the analytical calculations based on this equation in Figure 2.3, solid lines, in all panels. From top to bottom, we consider $1, 2$, and 3 $\sigma_{\rm vbc}$ effects, and compare these effects to the cumulative SIGO abundance estimated from our simulation boxes, shown as dashed lines. Note the consistency between the analytical and simulation SIGOs abundances at the range of few$\times 10^4$-few$\times 10^5$ $\mathrm{M}_\odot$.

At small masses ($M \approx 10^4$ $\mathrm{M}_\odot$), the limited resolution of the simulation yields lower abundances. At large masses, Poisson fluctuations increase the uncertainty of our simulation results, again creating an apparent disagreement between our simulation results and ana-

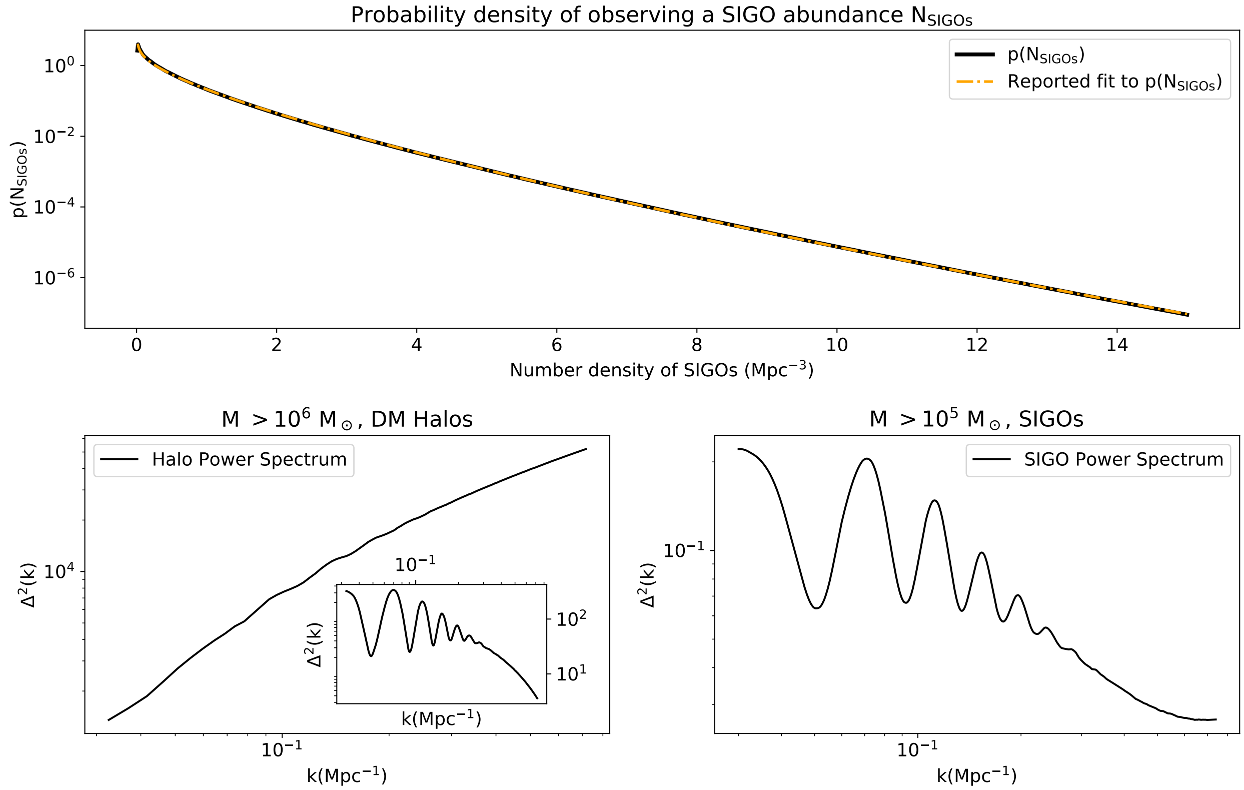


Figure 2.5: **The distribution of SIGOs in the Universe.** In the top panel, we show the probability density $p(N_{\rm SIGO})$ of observing a given abundance of SIGOs with M $> 10^5$M$_\odot$ in a randomly selected region of the Universe with an unknown (but constant) stream velocity. We also show a fit to this probability density function, given by Equation (2.16). In the bottom left panel, we present the power spectrum of the distribution of M$_{\rm Halo} > 10^6$ M$_\odot$ halos in our analytic model. Displayed is $\Delta^2(k) \equiv k^3/(2\pi^2)P(k)$, a non-dimensional quantity describing the variance in N(>M)$_{\rm Halo}$ per $\ln k$. Inset in the bottom left panel, we also show the portion of the same power spectrum that is solely the result of the stream velocity, and not an effect of large scale density fluctuations. In the bottom right panel, we show the power spectrum of the abundances of SIGOs with M $> 10^5$M$_\odot$ in our analytic model, also given as $\Delta^2(k)$. Note the different y-scale in the two bottom panels.

lytic model[5]. Given these caveats and the agreement over the relevant range we are confident that the analytical model provides a reasonable approach to estimate SIGO abundances.

Therefore, the bottom panels of Figure 2.3 show analytical calculations only, analogous to those of Figure 2.1. The bottom right panel shows our analytical results for comoving SIGO number densities as a function of mass at given values of the stream velocity. Here, unlike in the top panel, we integrate Equation (2.11) to infinity, rather than to the maximum mass we previously used to match simulations. We also find a simple relation between the SIGOs' cumulative number density as a function of mass $M_{\rm SIGO}$ and $\sigma_{\rm vbc}$ that fits the analytical model.

$$\log_{10}\left(\frac{N(>M_{\rm SIGO})}{\rm Mpc^3}\right) = a(v_{\rm bc})\left\{\log_{10}\left(\frac{M_{\rm SIGO}}{\rm M_\odot}\right)\right\}^{b(v_{\rm bc})} + c(v_{\rm bc}) \ .$$

Fits for the prefactors $a(v_{\rm bc}), b(v_{\rm bc})$ and $c(v_{\rm bc})$ are given in Appendix B.

In the bottom left panel, we show the ratio of SIGO number densities at given stream velocities and masses to the mean SIGO number density in the Universe at their mass. In other words, we show:

$$f_{\rm SIGO} = \frac{N(>M)_{\rm SIGO,vbc}}{\langle N(>M)_{\rm SIGO}\rangle} \ , \tag{2.12}$$

with

$$\langle N(>M)\rangle = \int_0^\infty p(v_{\rm bc}) N(>M)_{\rm SIGO,vbc} dv_{\rm bc} \ , \tag{2.13}$$

where $p(v_{\rm bc})$ is the likelihood of a given stream velocity given by a Maxwell distribution, and $\langle N(>M)_{\rm SIGO}\rangle$ is the mean number of SIGOs per comoving Mpc^3 at or above mass M in the Universe. Note, we could not use the number density of SIGOs with no stream velocity as the denominator for this ratio as in Figure 2.1, because SIGOs are only found when the stream velocity effect is present. As shown in Figure 2.3, left panel, large values of the stream velocity ($\sim 3.5\sigma_{\rm vbc}$) result in an enhancement in SIGO abundances on the order of $20-30\times$

---

[5] For example, the largest such fluctuation,for M$\sim$few$\times 10^5$M$_\odot$, is about a 3σ deviation in the $2\sigma_{\rm vbc}$ data. This represents a 1% fluctuation, which is consistent with the largest fluctuations due to Poisson statistics we would expect in our data set.

the mean abundance of SIGOs in the Universe. However, we note that only a small fraction of the Universe by volume (about 5×10^{-8}) have stream velocities above $3.5\sigma_{\rm vbc}$. We can also see that larger stream velocity values result in larger maximum SIGO masses.

Using our analytical approach, we can now examine the large-scale abundance of SIGOs. In particular, in Figure 2.4 we depict the fluctuations in SIGO abundances resulting from the variations in stream velocity on large scales. Analogous to the DM case, we show

$$\delta_{\rm SIGO} = \frac{N(>M)_{\rm vbc} - \langle N(>M)\rangle}{\langle N(>M)\rangle} \ , \tag{2.14}$$

with $N(>M)$ as defined in Equation (2.11). The overdensities in the plot were generated using the method for generating appropriately distributed density fields outlined in Section 2.2.2.

We also present a power spectrum of the fluctuations in SIGO abundances (i.e., fluctuations in $N[M_{\rm SIGO} > 10^5 M_\odot]$) on large scales, the bottom right panel of Figure 2.5. The power spectrum of these number densities is calculated as

$$\left\langle N_{\mathbf{k},\rm SIGO} N^*_{\mathbf{k},\rm SIGO} \right\rangle = (2\pi)^3 P(k) \delta^{(3)}(\mathbf{k} - \mathbf{k}') \ . \tag{2.15}$$

The plot shows the variance of these number density fluctuations per $\ln k$: $\Delta^2(k) \equiv k^3/(2\pi^2) \times P(k)$. Also shown is the power spectrum of M$> 10^6$M$_\odot$ DM halo abundances, for comparison, both with the effects of large-scale density fluctuations (main figure) and with only the effects of the stream velocity (inset figure). Note the similarities between the inset figure and the power spectrum of SIGO abundances, which are primarily set by velocity fluctuations. Note that the coherence scale of the number density is set by the range of scales over which $\Delta^2(k)$ is nonzero. In this case, as with the stream velocity that gives rise to this effect, that scale is approximately $k > 0.5$ Mpc^{-1}, implying that number densities are coherent on scales of a few comoving Mpc.

In the top panel of Figure 2.5, we show the probability density of observing an abundance $N_{\rm SIGO}(M > 10^5\ {\rm M}_\odot)$ of SIGOs per Mpc^3 within a region of the Universe with constant (but unknown) stream velocity. This can be expressed mathematically as $p(N_{\rm SIGO}) = p(v_{\rm bc}) \times dv_{\rm bc}/d(N_{\rm SIGO})$, where $p(v_{\rm bc})$ is given by a Maxwell distribution with scale parameter $\sigma_{\rm vbc}/\sqrt{3}$. To facilitate quick calculations for future semi-analytical studies, we provide a fit (also showed in the figure) for this probability density function:

$$\begin{aligned} \log_{10}[p(N_{\rm SIGO})] \approx &-0.933 \times \exp[1.734 \times \log_{10}(N_{\rm SIGO})] \\ &+ 0.261 \times \exp[-0.507 \times \log_{10}(N_{\rm SIGO})], \end{aligned} \tag{2.16}$$

calculated using a chi-squared fit with a Trust Region algorithm, where $N_{\rm SIGO}$ is given in units per Mpc^3 comoving. Note that this fit diverges as $\rm N_{SIGO} \rightarrow 0$, and is valid for $\rm N_{SIGO} \gtrsim 0.001$ Mpc^{-3}.

Using this equation, we see a fairly high likelihood that a given (small) region of space will contain virtually no SIGOs ($\sim 29\%$ chance that a given region will contain fewer than 0.1 SIGOs per Mpc^3), but a long tail–there is about a 13% chance that a given region of space will contain more than 1 SIGO per Mpc^3.

If indeed SIGOs are the progenitors of globular clusters, we would expect the distribution of globular clusters on the sky, as well as their overall abundance, to be similar. Using our model, we can predict the average abundance of SIGOs in the local Universe at early redshifts. Note that here we have used $\sigma_8 = 1.7$, in order to enhance the abundance of SIGOs in our simulations. However, we may still use this figure to compare to the real Universe for several reasons. The density power spectrum scales as $P(k) \propto a^2 \approx z^{-2}$, so doubling the normalization of the power spectrum as we have is similar (at the redshifts and scales we are discussing) to increasing the redshift discussed by a factor of $\sqrt{2}$ (Park et al., 2020). In other words, we can expect comparable SIGO abundances in the real Universe at $z \sim 14$, which is consistent with our analytical calculations. In addition, clusters in the Universe

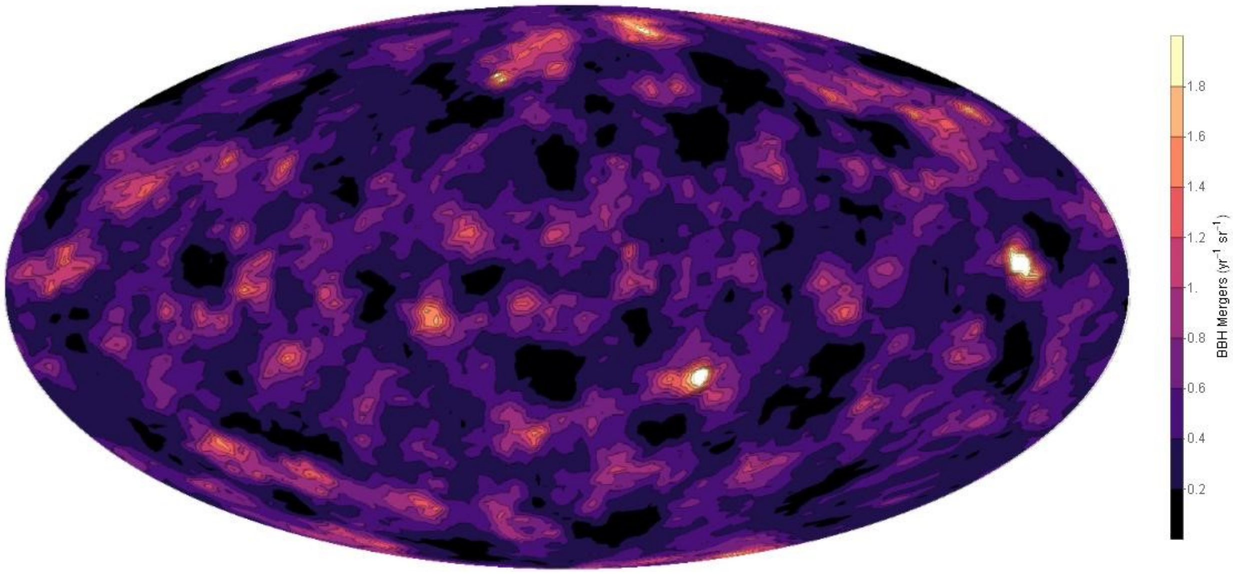


Figure 2.6: Sky-map of integrated BBH merger abundances in globular clusters, to a distance of 675 Mpc. Numbers are given in mergers per steradian per year, assuming a merger rate of 10^{-8} per year per cluster.

tend to form on high sigma peaks of the large-scale density field (e.g., Kaiser, 1984; Sheth & Tormen, 1999; Barkana & Loeb, 2004; Topping et al., 2018). Because of our proximity to the Virgo cluster, it is probable that we are at such a high sigma peak, which will enhance the concentration of structures of all masses relative to this work, which assumed a density consistent with the average matter density of the Universe.

With that in mind, our model predicts an average SIGO number density of ~ 0.5 Mpc^{-3} above $M > 10^5$ M$_\odot$, which may be extrapolated to recent times, as was done in Chiou et al. (2019), yielding a possibly similar number density. This result is consistent to order of magnitude with the observed local density of globular clusters, estimates of which range from 0.72 Mpc^{-3} (Rodriguez et al., 2015) to a few × Mpc^{-3} (e.g., Portegies Zwart & McMillan, 2000; Harris et al., 2013). Since SIGOs are early structures, they are likely to have low metallicities. Rodriguez et al. (2015) estimates the local density of low metallicity GCs as 0.46 Mpc^{-3}, which is also consistent with our estimate for the abundance of SIGOs. It is worthwhile, however, to be cautious with these comparisons, as not all of these early SIGOs necessarily evolve into present-day globular clusters (Naoz & Narayan, 2014; Popa et al., 2016; Chiou et al., 2021), and as additional SIGOs are expected to form after the redshifts considered, increasing the overall number of SIGOs formed (Naoz & Narayan, 2014; Popa et al., 2016; Chiou et al., 2018, 2019). An additional consideration is the possibility that SIGOs will fragment in their collapse from their initial size, which is on the order of tens to hundreds of pc (e.g., Chiou et al., 2021), to the size of a globular cluster (~ 1 pc). A typical SIGO, modelled as a puffy disc, has a Toomre stability criterion above unity (Toomre, 1964), and therefore will not fragment (excepting some central overdense regions, which could potentially collapse further to create high density star forming regions) in its collapse into GCs. In addition, in connecting the mass of SIGOs to the mass of globular clusters, we have implicitly assumed a high star formation efficiency. However, cluster formation and star formation are not 100% efficient (e.g., Baumgardt & Kroupa, 2007; Krumholz et al., 2019; Li et al., 2019; Grudić et al., 2021b), and while preliminary studies by Chiou et al. (2021) do support a relatively high star formation efficiency in SIGOs, star formation in SIGOs is still not fully understood. Nonetheless, these comparisons highlight an additional possible connection between these SIGOs and (particularly low-metallicity) globular clusters.

2.4 Implications to Gravitational Wave Anisotropies

For a standard initial stellar mass function, thousands of stellar-mass BHs likely form in a typical GC, many of which are likely initially retained (Belczynski et al., 2006; Willems et al., 2005; Wong et al., 2012). On a timescale of $\lesssim 1\,\mathrm{Gyr}$, these BHs sink to their host clusters' centers through dynamical friction, where they dynamically interact with other BHs to form BH binaries (e.g., Sigurdsson & Hernquist, 1993). It has been suggested that this process may be one of the leading sources for binary BH mergers [6] (Portegies Zwart & McMillan,

---

[6] Although other processes are suggested to be comparable, from isolating binaries (e.g., de Mink & Mandel, 2016; Belczynski et al., 2016; Marchant et al., 2016; Breivik et al., 2019, 2020) to dynamical evolution at nuclear star clusters at the center of galaxies (e.g., Hoang et al., 2018; Stephan et al., 2019; Wang et al., 2021).

2000; O'Leary et al., 2006; Rodriguez et al., 2015, 2016b; Chatterjee et al., 2017; Rodriguez et al., 2021).

If indeed most BBH mergers form in GCs, and if SIGOs are indeed the main progenitor of GCs, there should be an anisotropy in GW signals from binary BHs derived from and comparable to the anisotropies in SIGO abundances due to spatial variations in stream velocity. Using our analytic model, we can estimate the expected observed anisotropy.

In a typical GC similar to the ones observed in the Milky Way at present (mass of roughly a few$\times 10^5\,M_\odot$, core radii of roughly 1 pc, metallicity of roughly 10% solar; e.g., Harris, 1996), recent models predict roughly 100 total binary BH mergers over a roughly 10 Gyr cluster lifetime (e.g., Rodriguez et al., 2018; Fragione & Kocsis, 2018; Kremer et al., 2020; Antonini & Gieles, 2020). Assuming that these mergers are roughly uniformly distributed in time at zero-th order, this implies a BBH merger rate of roughly $10^{-8}\,\mathrm{yr}^{-1}$ per GC. By combining this order-of-magnitude rate estimate with the expected spatial distribution of GCs linked to SIGOs, we build a sense of the potential anistropy in BBH mergers from GCs.

As a proof-of-concept, in Figure 2.6 we show a sky-map of a line-of-sight integrated GW merger abundance, from GW BBH mergers in GCs, up to a distance of 675 Mpc (determined by our analytical box, see Section 2.2.2). At this distance redshift effects on the GW are negligible. As shown, we predict the integrated rate of BBH mergers may vary by as much as an order-of-magnitude over scales of $\sim 10\,\mathrm{deg}$. To a distance of 675 Mpc, we also estimate an approximate rate of BBH mergers of 0.5 mergers sr $^{-1}$ yr^{-1}, with a standard deviation of approximately 0.3 mergers sr^{-1} yr^{-1} across the sky. Payne et al. (2020) found that ten black hole mergers from the first LIGO/Virgo catalog are consistent with an isotropic distribution, over large scales ($>$ 100 Mpc). Further, recent endeavors by Abbott et al. (2021b) to search for anisotropic stochastic GW backgrounds did not find anisotropies on the three directions in the sky. Thus, both analyses imply that larger scales than the ones predicted here do not exhibit anisotropies, yielding a possible clear GW-sky signatures of SIGOs.

As the catalog of binary BH mergers continues to grow through both current and ongoing LIGO/Virgo/KAGRA detections (e.g., Abbott et al., 2021a) and detections by proposed third-generation detectors such as the Einstein Telescope (e.g., Punturo et al., 2010) and Cosmic Explorer (e.g., Abbott et al., 2017a), this anisotropy may potentially be observable and if observed, would further constrain the connection between GCs and SIGOs.

We stress, however, that Figure 2.6 and this discussion represent an ideal case where we assume that *all* SIGOs are directly linked to globular clusters and that their distribution (as well as the distribution of BH mergers) does not vary with redshift. While SIGOs may be linked to globular clusters (e.g., Naoz & Narayan, 2014; Chiou et al., 2019), it is unlikely that all SIGOs become globular clusters (e.g., Popa et al., 2016; Chiou et al., 2021), and we have yet to test the redshift evolution of SIGOs over large ranges. Nevertheless, this result suggests that if indeed SIGOs are GCs' progenitors, they may imprint an anisotropic sky distribution on the GW emission signal[7] .

2.5 Discussion

Supersonically induced gas objects (SIGOs) containing little to no DM are expected to exist in the early Universe (before reionization) with masses of $\lesssim$ few $\times\ 10^6\ \mathrm{M}_\odot$ (Naoz & Narayan, 2014; Popa et al., 2016; Chiou et al., 2018, 2019, 2021). They are the result of a decoupling between the DM and baryon fluids at the time of recombination because of the relative velocity between them (a.k.a. stream velocity Tseliakhovich & Hirata, 2010). This stream velocity is coherent only on small scales ($\sim$ few Mpc), which means that numerical simulations that track this effect can do so, over those scales. This small coherent scale poses a challenge when exploring large-scale SIGO abundances using numerical simulations.

---

[7] Note that a Kroupa (2001) initial mass function is often invoked; however, it is still unknown whether initial mass functions in the early Universe will follow a Kroupa profile. As was mentioned in Rodriguez et al. (2015), a 1σ variation in the slope of the high-mass end of the IMF can cause significant variation in the abundance of BBHs.

In addition, the relatively small mass of SIGOs ($\sim$few$\times 10^5 - 10^6$ M$_\odot$) also poses a mass-resolution challenge for numerical simulations.

Here, we combined small scale numerical simulations with analytical perturbation theory, and explored the large-scale distribution of SIGOs. Using normalizations obtained from high-resolution, small scale, numerical simulation results, we connect the decrease in dark matter halo formation at large stream velocities to SIGO abundances. We demonstrate that perturbation theory can be used to adequately model dark matter halo abundances, by comparing the results of perturbation theory to simulation results (Figure 2.1). We additionally show a comparison between our model results and simulations of SIGO formation and abundance (Figure 2.3), demonstrating that our model is useful for predicting SIGO abundances at typical stream velocities.

Our major results are as follows:

(i) **Halo abundance**: We show that increasing stream velocity decreases the number density of dark matter halos at $M <$ few $\times 10^7 M_\odot$ in both our analytic model and simulations, consistent with previous studies (e.g., Tseliakhovich & Hirata, 2010; Naoz et al., 2012; Popa et al., 2016) and with each other (Figure 2.1). Using this model and a power spectrum of the distribution of stream velocities on the sky, we present simulated maps of the number density of dark matter halos at a given mass (Figure 2.2). For a comparison more directly related to our Universe, we also show a power spectrum of dark matter halo abundances across the sky (Figure 2.5).

(ii) **SIGO abundance**: We use our analytic model for DM halo abundances to estimate SIGO abundances, giving the first fully analytic model for SIGO number densities. We posit that the decline in halos due to the stream velocity can be linked directly to the increase in SIGOs, and use this relation to analytically model SIGO number densities, using the abundance of SIGOs in simulations at various stream velocities to normalize our results (Figure

2.3). The strong agreement between our analytical model and the small-box simulation results, as depicted in the top panel of Figure 2.3, motivates us to use our analytical model to calculate the SIGOs abundance on large scales.

(iii) **Anisotropy in the distribution of SIGOs**: Because prior simulations relied on constant stream velocities, they could not be extended to large scales. Our analytic model does not have this limitation, and can be therefore be used to measure the large-scale variations of SIGO quantities on the sky. We use this to simulate maps of SIGO distributions (Figure 2.2), and to create a power spectrum of SIGO abundances (Figure 2.5). Our model predicts an average of 0.5 SIGOs of mass $M > 10^5 M_\odot$ per comoving Mpc^3 at $z \sim 14$, with a standard deviation $\sigma_{\rm SIGO} = 0.6$ Mpc^{-3}.

We also use the probability of a given stream velocity (given by a Maxwell distribution), along with the relation between stream velocity and SIGO number densities, to compute a probability density function for varying SIGO abundances (Equation (2.16) and the top panel of Figure 2.5).

(iv) **Connection to high redshift observations:** Simulations such as those in Chiou et al. (2019) suggest that SIGOs occupy a distinctive region in luminosity-size parameter space that may be distinguishable in future JWST observations (specifically, SIGOs are predicted to be dimmer than classical objects of the same radius). Follow-up studies may therefore soon be able to place observational constraints on the abundance of SIGOs as well as on their variation across the sky, contributing additional physical insight to these results. Should JWST indeed be able to observe them, we would expect their large-scale abundances to vary based on a power spectrum in agreement with that presented in Figure 2.5. The distribution of observed SIGOs should be qualitatively similar to the map presented in Figure 2.4.

(v) **Connections to globular clusters:**

If SIGOs are a progenitor of globular clusters, we can connect our conclusions about the variation in SIGO abundances on the sky to GC abundances. We find a mean SIGO number density of ~ 0.5 Mpc^{-3} at $z \sim 14$, which is consistent to order of magnitude with the observed local density of globular clusters, estimates of which range from 0.72 Mpc^{-3} (Rodriguez et al., 2015) to a few × Mpc^{-3} (e.g., Portegies Zwart & McMillan, 2000; Harris et al., 2013). More notably, this analytic SIGO number density is almost equal to the observed abundance of low metallicity GCs, 0.46 Mpc^{-3} (Rodriguez et al., 2015). As SIGOs are early structures with correspondingly low (expected) metallicities, this could be another indicator of a connection between SIGOs and (particularly low-metallicity) GCs.

(vi) **Anisotropy in the distribution of Gravitational Wave emission due to BBH merger events:** As discussed in Section 2.4, if we assume that SIGOs are indeed connected to GCs, we can connect the abundance of SIGOs to the abundance of GCs and therefore possibly to the abundance of BBH merger events. As a result, we suggest an anisotropy in the distribution of BBH merger events derived from the variation in SIGO abundances on the sky. As shown in Figure 2.6, this anisotropy could cause the integrated abundance of BBH mergers to vary by as much as an order of magnitude over scales of approximately 10 deg, to a distance of 675 Mpc, in an idealized case (though we caution that that variation would decrease on longer sightlines). Future observations of BBH merger events from LIGO, Virgo, KAGRA, and others may be able to observe this anisotropy, and could therefore further constrain the relation between SIGOs and GCs.

To summarize, we presented the larger scale abundance of supersonically induced gas objects (SIGOs), using a combination of analytical and simulation approaches. We thus predict variation of these high density gas objects in the early Universe ($z \sim 14$) with an average of ~ 0.5 Mpc^{-3}, possibly observable by JWST. The average number density of SIGOs is consistent with the local number density of globular clusters, further supporting the proposal that these SIGOs are the progenitors of globular clusters (e.g., Naoz & Narayan, 2014; Chiou et al., 2019). Finally, since globular clusters are natural birthplaces of black hole binary mergers, we propose that SIGOs may leave a distinct anisotropic signature on the gravitational wave signal on the sky.

CHAPTER 3

The Early Evolutionary Path of SIGOs

3.1 Introduction

The standard model of structure formation, invoking dark energy and dark matter (DM) that dominate the Universe's energy density budget, has had great success in explaining a wide variety of observations. The ΛCDM model successfully explains the anisotropies in the Universe's radiation field and the large-scale distribution of galaxies, as well as more generally explaining properties of the Universe's structure on scales larger than 100 Mpc (Springel, 2005; Vogelsberger et al., 2014a,b; Schaye et al., 2015). This model predicts that the Universe's structure formed hierarchically, with very early primordial baryon overdensities at $z \lessapprox 30$ collapsing to form larger objects, which eventually form galaxies and other structures.

These early baryon overdensities evolved in turn from interactions with existing DM overdensities following recombination. Because baryon objects had been inhibited from collapse prior to recombination by the photon field, DM overdensities at the time of recombination were about 10^5 times larger than baryon overdensities (e.g., Naoz & Barkana, 2005b). Because of this imbalance, the formation and growth of baryon overdensities at this time was

---

The contents of this chapter appeared in Lake et al. (2023a) driven by these large existing DM overdensities.

However, this process was complicated by the fact that there was a significant relative velocity between baryons and DM following recombination (Tseliakhovich & Hirata, 2010). This relative velocity had an rms value at recombination of 30 km s^{-1} (which was about 5 times the average speed of sound in the Universe at this time) and was coherent on few-Mpc scales. Because of its coherence on these scales, it is also known as the stream velocity.

Because the stream velocity was so highly supersonic, it induced significant spatial offsets between baryon overdensities and their parent DM overdensities, forming collapsed baryon objects outside the virial radii of their parent DM halos (Naoz & Narayan, 2014). These collapsed baryon objects, known as Supersonically Induced Gas Objects or SIGOs, have been proposed as a progenitor of some modern-day globular clusters (GCs), and have been shown to have similar properties to GCs (e.g., Chiou et al., 2018, 2019, 2021; Lake et al., 2021; Nakazato et al., 2022). However, SIGOs remain poorly understood. Early simulations of SIGOs, aimed at confirming their existence and putting basic constraints on their properties (e.g., Popa et al., 2016; Chiou et al., 2018, 2019, 2021; Lake et al., 2021), included only adiabatic and sometimes atomic hydrogen cooling. However, molecular hydrogen cooling has important effects on the structure of SIGOs (Nakazato et al., 2022), and on the abundance of gas objects in general in the early Universe (e.g. Glover, 2013; Schauer et al., 2021; Nakazato et al., 2022; Conaboy et al., 2022). For example, SIGOs in molecular cooling simulations tend to be far more filamentary than those in atomic cooling simulations (Nakazato et al., 2022), potentially for reasons analogous to the Zel'dovich approximation, which states that collapse will occur along successive axes, starting with the shortest (Zel'dovich, 1970). H_2 cooling also lowers the temperature in these SIGOs to $\sim$ 200 K, which lowers the Jeans mass to about 10^3 $M_{\odot}$, potentially leading to star formation in SIGOs (Nakazato et al., 2022, Lake et al. in prep).

In order to understand and contextualize the properties of SIGOs with molecular hydrogen cooling, it is useful to have an analogous class of objects to compare to, which share many of the properties of SIGOs. A natural class of objects to compare to here are giant molecular clouds (GMCs). Like SIGOs, GMCs form in environments shaped by supersonic turbulence (e.g., Larson, 1981; Padoan & Nordlund, 1999; Mac Low & Klessen, 2004; Krumholz & McKee, 2005; McKee & Ostriker, 2007; Krumholz & Tan, 2007; Bergin & Tafalla, 2007; Padoan & Nordlund, 2011; Burkhart & Lazarian, 2012; Semenov et al., 2016; Mocz et al., 2017; Burkhart, 2018, 2021; Appel et al., 2022). The energy from this turbulence cascades from the driving scale of the turbulence $\mathrm{L}_{\mathrm{Drive}}$ down to the sonic scale λ_s which marks the boundary between supersonic and subsonic turbulence:

$$\lambda_s = \frac{\mathrm{L}_{\mathrm{Drive}}}{\mathcal{M}^2} \ , \tag{3.1}$$

where $\mathcal{M}$ is the mach number of the turbulent flow on the driving scale, and where we have assumed Larson's law (Larson, 1981). This energy cascade is important in both cases, because it creates high density peaks in the turbulent medium on scales comparable to the sonic scale (e.g., Krumholz & McKee, 2005), which is key to understanding both objects' structure. Ultimately, this process leads to a critical density for star formation that can be expressed in the form

$$\rho_{\mathrm{crit,GMC}} = \frac{\pi c_s^2 \mathcal{M}^4}{G \mathrm{L}_{\mathrm{Drive}}^2} \ , \tag{3.2}$$

where c_s is the speed of sound and G is the gravitational constant.

GMCs are also interesting as a point of comparison for SIGOs because they may be a major formation channel for globular clusters at early times (e.g., $z \sim 6$ Harris & Pudritz, 1994; Elmegreen & Efremov, 1997; Kravtsov & Gnedin, 2005; Shapiro et al., 2010; Grudić et al., 2022). GMCs form with a wide range of masses and densities, and it has been shown that the high-density, high-mass end of this formation channel may be capable of efficiently forming globular clusters. This formation process is supported by observations of the merging Antennae system and its population of massive young clusters (e.g., Whitmore & Schweizer, 1995; Whitmore et al., 1999). Therefore, a better understanding of the comparison between SIGOs and GMCs may also be important to understanding the link between SIGOs and globular clusters.

Motivated by these factors, in the present paper we present a suite of simulations using the cosmological simulation code `AREPO`, including primordial chemistry and accounting for the effects of molecular hydrogen cooling. We constrain SIGOs' structural and kinematic properties, including the effects of the stream velocity and supersonic turbulence on the size and density of SIGOs, and compare these properties to those of GMCs to better contextualize them. As the formation of SIGOs is now well-understood, we aim to characterize the next step of SIGOs' lives by providing physical intuitions for the processes through which SIGOs evolve. To this end, we provide an analysis of the various timescales important for early SIGO evolution. This includes the cooling time, as well as the timescale on which they collapse gravitationally. It also includes the free-fall time on which SIGOs fall into nearby DM halos. Taken together, these characteristic timescales provide a better understanding of the potential for star formation in SIGOs. We also discuss two other mechanisms relevant to SIGO evolution: growth and two-body relaxation.

This paper is organized as follows: in Section 3.2 we discuss the setups of the simulations used, as well as modifications made to definitions of SIGOs used in prior works. In Section 3.3 we expand upon the importance of molecular hydrogen cooling to SIGOs' properties. In Section 3.4, we discuss the similarities and differences between SIGOs and GMCs, and revisit previous density-driven analyses of SIGO evolution. In Section 3.5, we overview the evolution of SIGOs from a timescale perspective, primarily discussing the potential for SIGOs to collapse through the lens of molecular hydrogen cooling. Lastly, in Section 3.6 we overview future avenues of exploration towards understanding star formation in SIGOs and summarize our results.

For this work, we have assumed a ΛCDM cosmology with $\Omega_{\Lambda} = 0.73$, $\Omega_{\rm M} = 0.27$, $\Omega_{\rm B} = 0.044$,

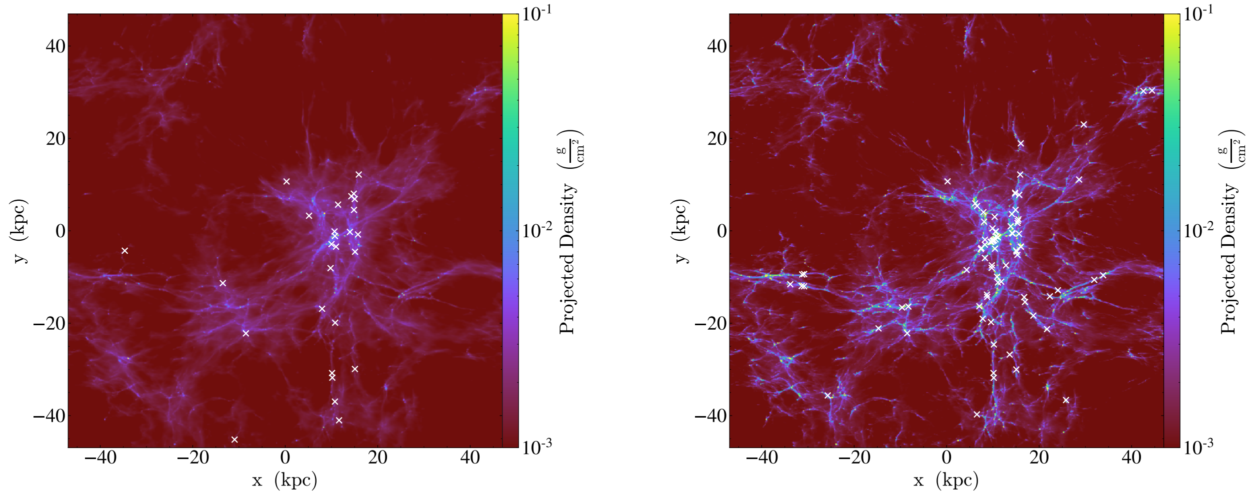
$\sigma_8 = 1.7$, and $h = 0.71$.

3.2 Numerical Setup and Object Classification

Using the cosmological and hydrodynamic simulation code `AREPO`, we present a suite of 4 simulations demonstrating the effect of molecular hydrogen cooling on the properties of SIGOs. The main parameters of these simulations are listed in Table 3.1. Here, we indicate Runs with molecular hydrogen cooling using "H2", and indicate runs without molecular hydrogen cooling (simply using adiabatic and atomic hydrogen cooling) with "H". Runs without the stream velocity are labelled with "0v", whereas Runs with a 2σ stream velocity (a 2 $v_{\rm rms}$ relative velocity between baryons and dark matter) are labelled as "2v". The stream velocity in the latter Runs is implemented as a uniform boost to baryon velocities in the x direction in the initial conditions–at the initial redshift of $z = 200$, this is a boost of 11.8 km s^{-1}.

Our initial conditions were generated using transfer functions calculated using a modified version of `CMBFAST` (Seljak & Zaldarriaga, 1996) taking into account the first-order correction of scale dependent temperature fluctuations (e.g., Naoz & Barkana, 2005b; Naoz & Narayan, 2013) and second-order corrections to the equations presented in Tseliakhovich & Hirata (2010), describing the evolution of the stream velocity. We use two transfer functions, one each for the baryons and DM, as the evolution of the gas fraction strongly depends on the initial conditions of the baryons (e.g., Naoz et al., 2009, 2011, 2012, 2013; Park et al., 2020). The glass file for baryons uses positions shifted by a random vector, rather than tracing the dark matter density perturbations (Yoshida et al., 2003a).

Our simulations use 512^3 DM particles with a mass of 1.9×10^3 $\mathrm{M}_\odot$ and 512^3 Voronoi mesh cells corresponding to a gas mass of 360 $\mathrm{M}_\odot$, in a box 2 comoving Mpc on a side. This box size aims to study a patch of the Universe with constant stream velocity, as a proof of concept, rather than analysing structure formation in larger regions with variable stream



(a) Atomic Cooling Only

(b) Molecular Hydrogen Cooling

Figure 3.1: **Comparison of the abundance of SIGOs with and without molecular cooling at $z = 20$.** We show the gas density field in our simulation box for simulation 2vH2 (molecular cooling, right panel), compared to simulation 2vH (atomic cooling, left panel). SIGOs are marked with Xs. Note here that SIGOs trace gas and halo abundances on these scales. Note also that molecular hydrogen cooling dramatically increases the abundance of SIGOs.

velocity. Following Chiou et al. (2019, 2021); Lake et al. (2021); Nakazato et al. (2022), we choose $\sigma_8 = 1.7$, representing a rare, overdense region where structure forms early in a large volume, in order to increase the number of gas objects in our simulation, allowing increased statistical power. The simulations begin at $z = 200$, and they run to $z = 20$.

In Runs 0vH2 and 2vH2, using the chemistry and cooling library GRACKLE (Smith et al., 2017; Chiaki & Wise, 2019), we track nonequilibrium chemical reactions and their associated radiative cooling explicitly in the gas. This Run includes H_2 and HD molecular cooling, as well as chemistry for 15 primordial species: e^-, H, H^+, He, He^+, He^{++}, H^-, H_2, H_2^+, D, D^+, HD, HeH^+, D^-, and HD^+. The radiative cooling rate of H_2 follows both rotational and vibrational transitions (Chiaki & Wise, 2019).

For comparison, in Runs 0vH and 2vH, we consider only atomic hydrogen cooling, following the species e^-, H, H^+, He, He^+, and He^{++} in equilibrium with a spatially uniform, redshift-dependent photoionizing background, as described in Vogelsberger et al. (2013).

We use the object classifications laid out in Chiou et al. (2018). We identify DM halos with an FOF (friends-of-friends) algorithm that uses a linking length of 20% of the mean particle separation of the DM component, which is about 780 comoving pc. Assuming sphericity for simplicity, we use this to calculate positions and radii of all DM halos in the simulation output (though note that DM halos at this time resemble triaxial ellipsoids, e.g. Sheth et al., 2001; Lithwick & Dalal, 2011; Vogelsberger & White, 2011; Schneider et al., 2012; Vogelsberger et al., 2020).

We then run the FOF finder on the gas component, allowing us to identify gas-primary objects that contain over 100 gas cells. Especially when molecular hydrogen is included, SIGOs at these redshifts are distinctly filamentary, similar to in the Zel'dovich approximation (Zel'dovich, 1970). We cannot, therefore, assume sphericity when determining these objects' properties, as we did with DM halos. To address this issue, we follow the procedure introduced in Popa et al. (2016). In particular, we fit these gas-primary objects to ellipsoids determined as the smallest ellipsoidal surface that surrounds all of the constituent gas cells. We then reduce the major axis of this ellipsoid by 5% until either the ratio of the axes lengths of the tightened ellipsoid to that of the original ellipsoid is greater than the fraction of gas cells remaining inside the tightened ellipsoid, or until 20% of their particles have been removed. Finally, in order to distinguish SIGOs from other classes of gas objects, we require that the center of mass of the gas objects be located outside the virial radius of all nearby halos, and that the SIGOs have a baryon fraction above 60%.[1] This condition is discussed and justified in Appendix C.

After identifying every SIGO present in each of the 150 snapshots in Run 2vH2 (evenly

---

[1] This condition was introduced as 40% in initial work by Chiou et al. (2018), and modified by Nakazato et al. (2022) to account for false identifications of SIGOs arising from molecular cooling, as discussed in Section 3.3 and justified in Appendix C.

Run	$v_{\rm bc}$	H_2 Cooling
0vH2	0	Yes
0vH	0	No
2vH2	2σ	Yes
2vH	2σ	No

Table 3.1: Simulation Parameters

spaced in redshift from $z = 30$ to $z = 20$), we compare the gas cells present in each gas-primary object at each redshift to track gas objects across snapshots. If two gas objects in different snapshots share at least 1/3 of the gas cells present in the smaller of the two objects, they are assumed to be the same object at different times. In other words, we track the gas cells' IDs and require that at least 1/3 of the gas cells will be shared by the larger gas object. By identifying which of these objects are SIGOs (and at what redshifts they fulfill the conditions needed to be identified as a SIGO), we can trace the history of SIGOs in Run 2vH2.

3.3 The Importance of Molecular Cooling to SIGOs' Properties

Molecular hydrogen cooling plays a vital role in the formation of the first stars (e.g., Saslaw & Zipoy, 1967; Haiman et al., 1996; Abel et al., 1998; Yoshida et al., 2003b; Stacy et al., 2011; Greif et al., 2011; Yoshida et al., 2008; Bromm, 2013; Glover, 2013; Hummel et al., 2016; Nakazato et al., 2022). Previous studies of SIGOs have argued that adiabatic and atomic cooling may be a sufficient approximation to understand SIGOs' formation and morphology (e.g., Popa et al., 2016; Chiou et al., 2018, 2021; Lake et al., 2021). This is because the main factor in the formation of SIGOs is the phase shift of gravitational fluctuations between DM and baryons (Naoz & Narayan, 2014). However, it was recently shown that molecular cooling may be able to increase the efficiency of SIGO formation, as well as playing an important role in their density (Schauer et al., 2021; Nakazato et al., 2022).

Quantifying the population-level differences in SIGOs' properties with molecular cooling compared to atomic cooling can help to illuminate the physical processes important to their later evolution. Figure 3.1 shows a comparison of SIGO abundances with and without molecular hydrogen cooling (Runs 2vH and 2vH2). SIGOs in this figure are marked with white Xs, displayed against a backdrop of the gas density in the full simulation box. Note that length scales here are in physical kpc. With molecular hydrogen cooling, SIGOs are dramatically more efficient at condensing from overdensities in the gas. Thus, we find 85 SIGOs at $z = 20$ in Run 2vH2, compared to 27 in Run 2vH. With molecular cooling, SIGOs also reach higher densities, due to their increased cooling efficiencies. In Run 2vH2, the most dense SIGO at $z = 20$ had an overall gas density of 39.1 cm^{-3}, compared to a maximum gas density in any SIGO of 2.5 cm^{-3} in Run 2vH. Therefore, SIGOs are more abundant and denser by redshift 20 with molecular cooling.

The increase in SIGO abundances through molecular cooling seen in Figure 3.1 can be seen as a function of redshift in Figure 3.2. This figure shows the time evolution of SIGO abundances with molecular cooling, as well as showing SIGO abundances at integer redshifts in Run 2vH for comparison. As one can see, the abundance of SIGOs increases with time, and consistent with Nakazato et al. (2022), abundances are generally enhanced through molecular cooling, through the process outlined above. This process is most important at later redshifts: as one can see in the figure, the abundance of SIGOs in Run 2vH2 begins to significantly diverge from the abundance without molecular cooling after $z \sim 25$ in our simulations.

Molecular cooling affects the structure around SIGOs as well: SIGOs form embedded in gas filaments, and molecular hydrogen cooling permits these filaments to condense much more efficiently than does atomic hydrogen cooling. As mentioned in Section 3.2, following Nakazato et al. (2022), we revise the gas fraction cutoff for identification of SIGOs with our FOF algorithm upwards compared to e.g. Chiou et al. (2021), as we are working with molecular hydrogen cooling. This exclusively reduces the number of SIGOs considered, in order to prioritize positive identifications over concerns about false negatives, allowing us to

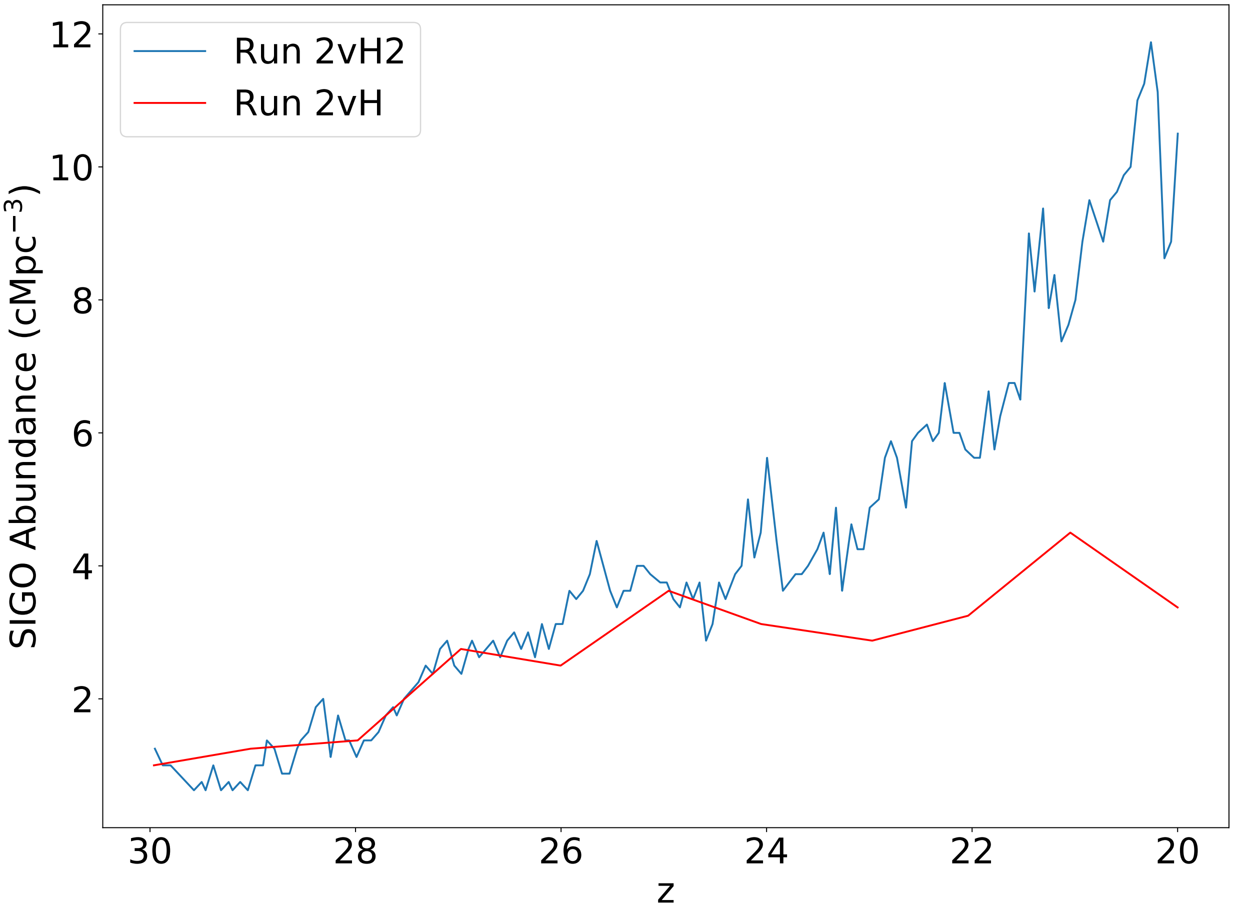
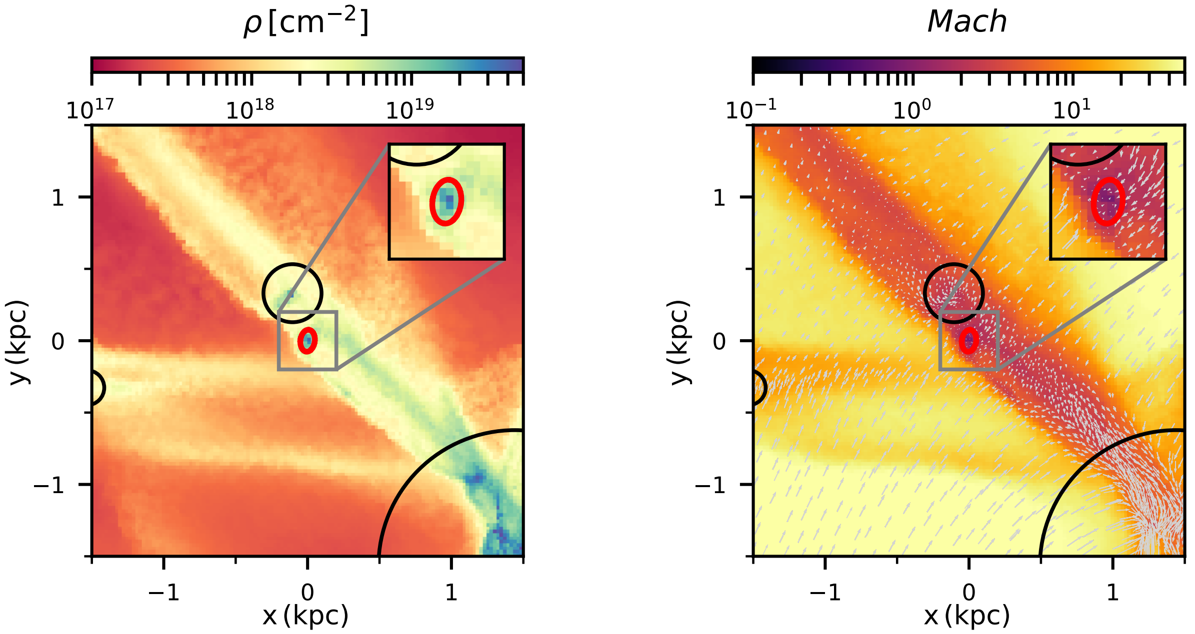


Figure 3.2: **Changes in SIGO abundance with redshift.** Plotted here is the evolution of SIGO abundances with redshift in Run 2vH2. For comparison, SIGO abundances from Run 2vH are plotted in red. Note that the time resolution of Run 2vH2 is higher than that of Run 2vH, for the purposes of this study.



(a) Gas Column Number Density

(b) Mach number

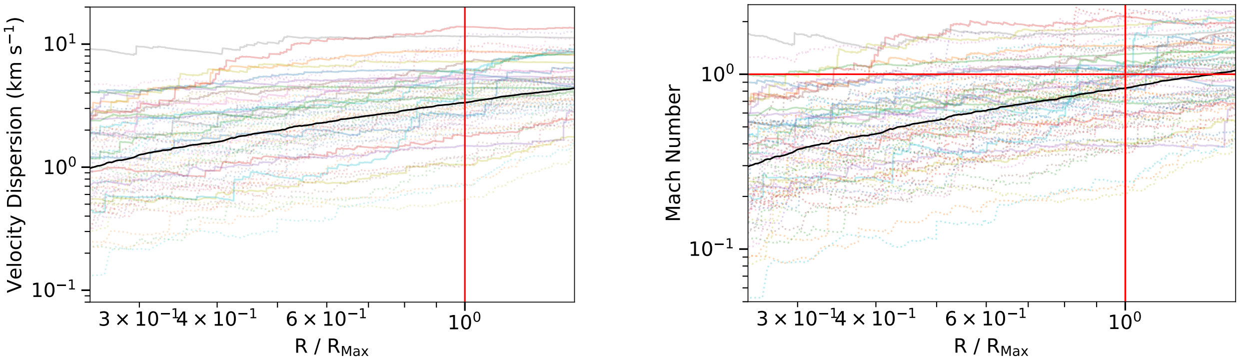
Figure 3.3: **Example of a Young SIGO in a Turbulent Shock:** The velocity field of a typical SIGO (in red) with molecular cooling. This SIGO is associated with a nearby parent DM halo, and there is a larger DM halo at the bottom right of the field. Here, we see that the velocity dispersion inside the SIGO is quite small compared to that outside of it. The sonic scale here is larger than the scale of the SIGO, so turbulent flow plays a small role in the SIGO's potential further collapse. As can be seen the SIGO is embedded in a shock front, which has formed on a scale where the Mach number is about unity. See Figure 3.4.

state with confidence that objects studied are truly SIGOs. This is because gas filaments condense so efficiently that they sometimes appear to structure-finding algorithms (even in Run 0vH2 where no SIGOs should form) as collapsed baryon objects outside of halos. Because SIGOs are embedded in these gas filaments, this process is inextricably linked to their formation: the mass and particularly the density of cooled gas around SIGOs in these filaments increases in Run 2vH2 compared to Run 2vH, potentially allowing SIGOs to accrete external gas and grow further with molecular cooling.

3.4 SIGOs as GMC analogues

As mentioned above (§3.1), at face value, GMCs and SIGOs seem to share several key properties. At the most fundamental level, the impact of molecular hydrogen is critical for understanding the evolution of both classes of objects, which also have similar masses ($\sim 10^6$ $M_\odot$) and scales ($\sim$ 100 pc). Additionally, these classes have similar substructures, with SIGOs often exhibiting local density fluctuations such as cores and filaments in simulations, similar to those in GMCs (Mac Low & Klessen, 2004; Krumholz & McKee, 2005; Mocz & Burkhart, 2018). In addition, the formation of both classes of object is dominated by supersonic turbulence from their environment (Burkhart et al., 2009; Padoan & Nordlund, 2011; Burkhart & Lazarian, 2012; Semenov et al., 2016; Mocz et al., 2017; Burkhart, 2018; Chiou et al., 2019; Burkhart, 2021; Appel et al., 2022). In considering star formation within both classes of objects, it is often impossible to consider only the Jeans mass– instead, one must use a critical density that incorporates the impact of turbulence.

To illustrate the nature of the turbulence around SIGOs, in Figure 3.3, we see the velocity field around one SIGO from Run 2vH2, including that of its shock front. Mach numbers are labelled based on the local temperature and the velocity relative to the center of mass of the SIGO. The SIGO itself here has a maximum radius of 118 pc, which is slightly smaller than the sonic scale. The volume-averaged Mach number within the SIGO is 0.97. One way to think about this SIGO is as a density fluctuation induced by this supersonic turbulence: the strongest density perturbations from the supersonic motions take place at the sonic scale, appearing as gas overdensities whose positions are correlated with those of nearby dark matter overdensities. This creates gas objects (SIGOs) offset from their parent halos, with sizes that are by necessity comparable to the sonic scale. Once this turbulence-induced structure formation occurs, molecular cooling helps to cool the gas within the SIGO to temperatures of order 200 K (Nakazato et al., 2022), which enables further collapse. Collapse of the SIGO reduces its axis length from the sonic scale, so a typical SIGO will



(a) Velocity Dispersion

(b) Mach Dispersion

Figure 3.4: **SIGOs' linewidth-size relation:** Panel (a) shows the linewidth-size relation for SIGOs, computed over a sphere centered on the center of mass of each SIGO, with increasing radius (scaled for each SIGO by the length of the longest axis of that SIGO). The translucent lines represent individual SIGOs, and the black opaque line represents an average of all SIGOs in log space. Dotted lines represent low-mass SIGOs below 10^5 $M_\odot$, which are more affected by 2-body relaxation resulting from our limited resolution (see text). Panel (b) shows the mach number dispersion, showing that most SIGOs are somewhat smaller than the sonic scale. The vertical red line in each panel displays the length of the longest axis of the SIGO, and the horizontal red line in panel b indicates Mach 1.

have a maximum axis length that is comparable to but smaller than the sonic scale, unless it either accretes additional gas from its surroundings (permitting it to grow while accreting mass) or collapses.

Figure 3.4 shows a more general comparison of the sonic scales and the maximum axis lengths of the SIGOs in Run 2vH2. In the top panel, each faded line shows the velocity dispersion within a SIGO as a function of its radius, and in the bottom panel each faded line shows the mach number dispersion as a function of radius. The black line is the average of all SIGOs' dispersions in log space. The spatial distributions of each are similar, because SIGOs are close to isothermal: their sound speeds do not change much on the edges of the SIGOs compared to the centers. Typical temperatures at the edge of a SIGO are of order 10% higher than those in the central regions. Therefore, their mach dispersions roughly follow their velocity dispersions.

A typical SIGO has a volume-averaged velocity dispersion of about Mach 0.8, or about 3 km/s (with fairly high variance) on the scale of its largest axis. Put another way, a typical SIGO has a sonic scale that is about 1.6 times larger than its longest axis. This assumes a driving scale for the turbulence on the order of the maximum axis length of the SIGO, as discussed by Chiou et al. (2019). Because the sonic scale is comparable to the size of a SIGO (and in some cases is smaller than the SIGO), it is important to account for this turbulence when considering the ability of a SIGO to undergo gravitational collapse (see also Hirano et al. (2017a) for similar effect in non-SIGOs). Chiou et al. (2019) accounted for this effect by adapting a critical density (previously used in GMCs by e.g., Krumholz & McKee, 2005), defined by Equation (3.2). This critical density defines the scale on which turbulence-induced fluctuations collapse through Jeans collapse, in effect equating the Jeans scale to the sonic scale. However, the analysis in the paper treated SIGOs as spherical objects with uniform density $\rho_{\rm SIGO}$ within a radius $\rm R_{max}$ defined as the longest principal axis of the SIGO. SIGOs are not spherical (particularly with molecular hydrogen cooling accounted for), and have a typical shortest axis length $\rm R_{min}$ that is of order 10 times smaller than their longest axis in Run 2vH2; therefore, this overestimates SIGOs' masses. An updated treatment for the critical density of collapse for SIGOs, then, equates the mass of the SIGO to the Jeans mass, giving a new definition of the critical density of a SIGO for collapse $\rho_{\rm crit}$:

$$\rho_{\rm crit} \approx \frac{\pi c_s^2 \mathcal{M}^4}{4G{\rm R_{min}}{\rm R_{int}}}, \tag{3.3}$$

where $\rm R_{int}$ is the second-largest principal axis of the SIGO and where we have assumed $\rm L_{drive} \approx 2 \times R_{max}$.

At $z = 20$, no SIGOs in Run 2vH exceed this critical density. This result contrasts with that of Chiou et al. (2019), which found that SIGOs can collapse through only atomic hydrogen cooling, because of the additional factor of the prolateness of the SIGOs considered in this analysis. With atomic cooling alone, the cooling timescale for these SIGOs is long, so it is unlikely that SIGOs could collapse to potentially form stars outside of halos without considering molecular hydrogen cooling. However, when considering molecular hydrogen cooling, looking only at these SIGOs' densities at early redshifts does not paint the whole picture of SIGOs' ability to form stars outside of halos.

While the critical density and the Jeans density are both important to SIGOs, their location outside nearby halos allows their evolution to be characteristically slow. At $z = 20$, about 50 Myr or less after SIGOs begin to form in meaningful quantities, SIGOs tend to be fairly isolated objects. They are outside of the immediate vicinity of the DM halos that birthed them, and have characteristically long fall-back times to their nearest halo. Because of this, even though their cooling times are long (Schauer et al., 2021), SIGOs can have shorter cooling times than fall-back times, allowing them to collapse in spite of their early Jeans stability (though typically later than $z = 20$, because these timescales tend to be $\gtrapprox$ 50 Myr). This may permit the formation of stars. Subsequently, they may be able to accrete onto halos on the fall-back timescale, and potentially survive as identifiable clusters due to their compact nature and existing population of stars. In order for this to happen, they must have cooling and collapse timescales that are shorter than their fall-back timescale to nearby halos, as discussed in Section 3.5. At any given time, then, SIGOs may be destined for collapse through cooling despite not exceeding the critical density for collapse. Therefore, to study the evolution of SIGOs, we argue that a timescale analysis is more appropriate than a density analysis, as follows in Section 3.5.

Another of the primary characteristics of GMC populations is their power law mass spectrum. GMCs in the inner disk of the Milky Way follow a power law mass spectrum given by $\mathrm{dN} \propto \mathrm{M}^{\xi}\mathrm{dM}$ with $\xi \sim -1.5$. In contrast, the mass spectra of GMCs in the outer Milky Way ($\xi \sim -2.1 \pm 0.2$) and M33 ($\xi \sim -2.9 \pm 0.4$) are notably steeper (e.g. Rosolowsky, 2005). Characterizing this mass spectrum is vital to understanding both how the clouds formed, as well as their overall contribution to large-scale star formation. Establishing how the mass spectrum of SIGOs compares to these power laws is useful not only for furthering theoretical work with SIGOs, but also for contextualizing similarities between SIGOs and these different populations of GMCs.

Figure 3.5 shows the mass spectrum of SIGOs in Run 2vH2 at $z = 20$. The blue histogram points show the probability density of finding a given SIGO in the labelled mass bin, with Poisson errors. However, immediately translating this estimate of the mass spectrum to a power law is complicated by the artificial 2-body relaxation present in the simulation (see Section 3.5.4 for more details). At low masses, this mass spectrum is distorted, or truncated, by the limited resolution of our simulation. Therefore, we only use mass points above $\mathrm{M}_{\mathrm{SIGO,min}} = 10^5\ \mathrm{M}_\odot$ (278 gas particles) to construct a best-fit power spectrum. This cutoff reflects SIGOs whose artificial 2-body relaxation timescales are longer than the age of the oldest SIGO in the simulation, which are therefore less affected by relaxation. We introduce an upper bound mass cutoff of $10^6\ \mathrm{M}_\odot$, to avoid over-fitting to the SIGO with a mass of $\sim 10^7\ \mathrm{M}_\odot$, because we do not have enough data points to determine whether or not the spectral index varies over the range $10^6\ \mathrm{M}_\odot \leq \mathrm{M}_{\mathrm{SIGO}} \leq 10^7\ \mathrm{M}_\odot$. The best-fit power law is expected to diverge from the data in this range due to Poisson fluctuations. We calculate a maximum likelihood power law index $\hat{\xi}$ using the method of Clauset et al. (2009) (derived from Muniruzzaman 1957), using our lower-bound mass cutoff of $10^5\ \mathrm{M}_\odot$ as $\mathrm{M}_{\mathrm{SIGO,min}}$ as follows:

$$\hat{\xi} = 1 + n \left[\sum_{i=1}^{n} \ln \frac{\mathrm{M}_{\mathrm{SIGO},i}}{\mathrm{M}_{\mathrm{SIGO,min}}} \right]^{-1}, \tag{3.4}$$

where n is the number of SIGOs with $10^5\ \mathrm{M}_\odot \leq \mathrm{M}_{\mathrm{SIGO}} \leq 10^6\ \mathrm{M}_\odot$, and $\mathrm{M}_{\mathrm{SIGO},i}$ is the mass of each of these SIGOs.

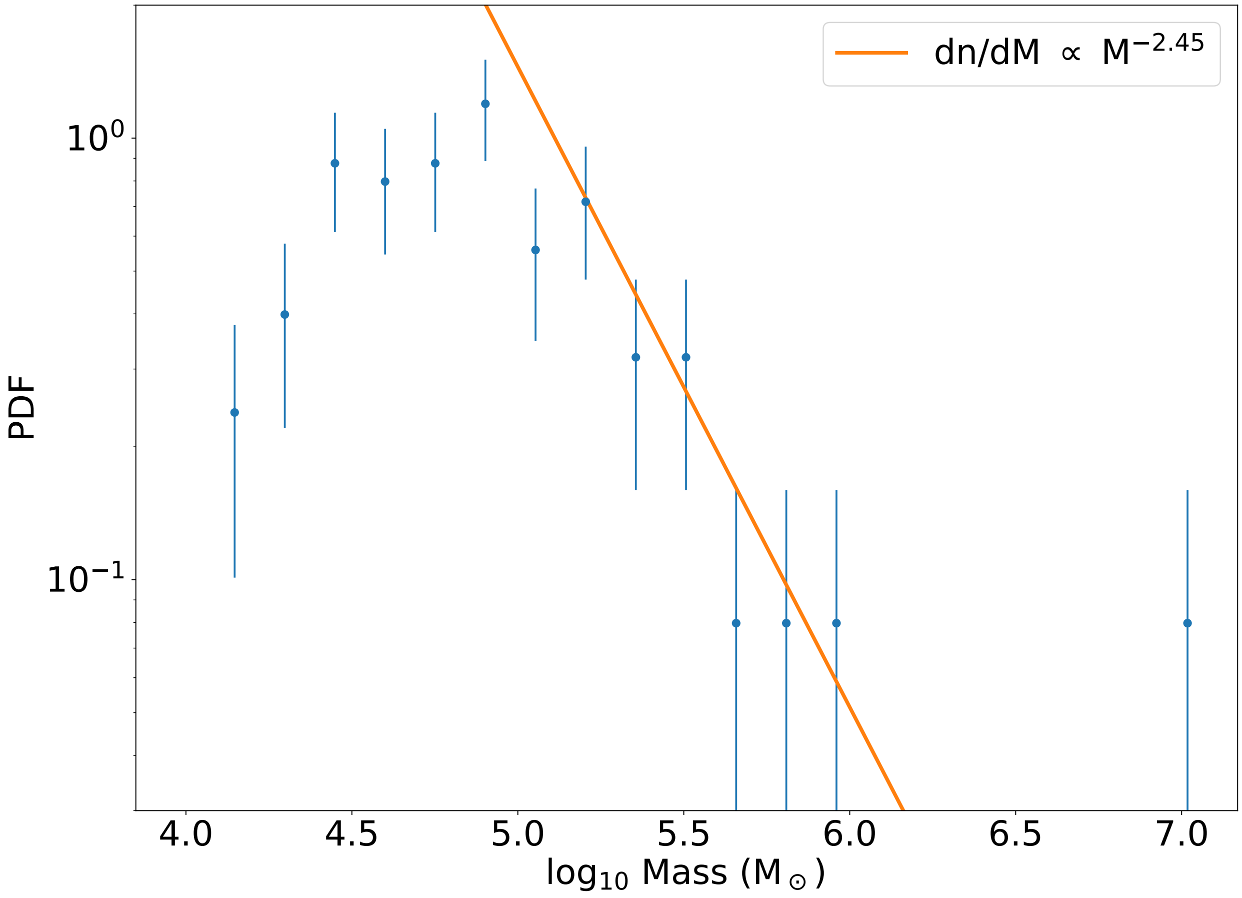


Figure 3.5: **The Mass Spectrum of SIGOs:** Here we show the mass spectrum of SIGOs from Run 2vH2 at $z = 20$ (the blue points and associated Poisson errors, plotted as a probability density in log space), as well as showing (in orange) a best-fit power law mass spectrum for high-mass SIGOs. The mass spectrum of SIGOs in this Run is consistent with a power law index of -2.4 ± 0.3.

The error σ_ξ is then estimated from the width of the maximum likelihood estimate as

$$\sigma_\xi = \frac{\hat{\xi} - 1}{\sqrt{n}} + \mathrm{O}\left(\frac{1}{n}\right). \tag{3.5}$$

With these conditions in place, the orange line in Figure 3.5 shows a maximum likelihood power law for our simulated SIGOs' masses. We find a best-fit mass spectrum index for SIGOs in this Run of $\xi \sim -2.4 \pm 0.3$, which is, interestingly, consistent with the mass

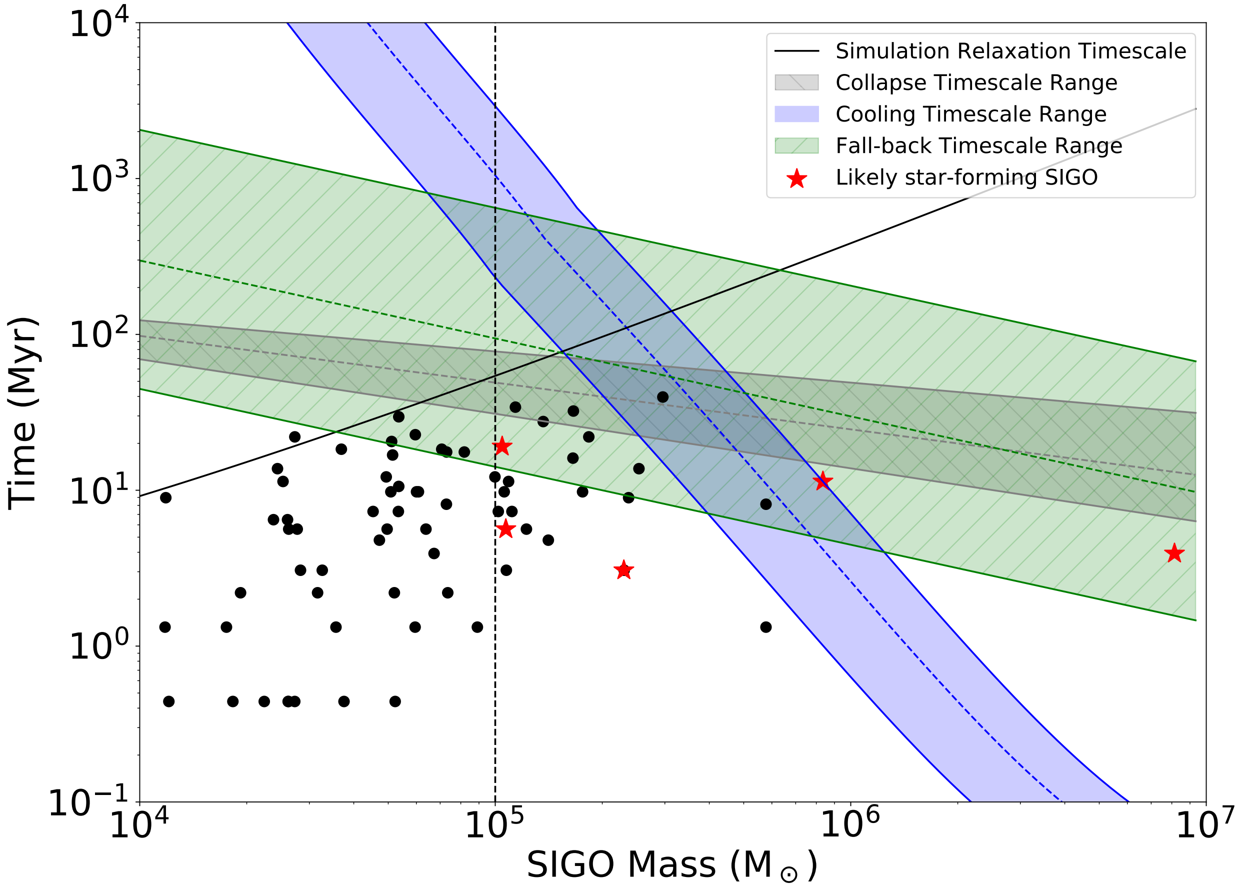


Figure 3.6: **Timescales of SIGOs:** here we show a number of important timescales to the evolution of SIGOs at $z = 20$ in Run 2vH2. Black dots mark the age and baryon mass of SIGOs found at this redshift, which represent 16% of the SIGOs that form in our simulation at all redshifts. Ranges are representative of the properties of SIGOs at $z = 20$ in our simulation (see text for details of assumptions). The green region and lines in these figures show the range of fall-back timescales to the nearest halo, as a function of mass. The green dashed line indicates the same for a SIGO with median properties. The blue region shows the same, but for the range of cooling timescales, and the gray region shows the timescale for gravitational collapse of a SIGO. The black line shows the relaxation timescale of SIGOs in the simulation at the resolution of the simulation (much shorter than in the real Universe). The black dashed line indicates the mass scale above which simulated SIGOs are less affected by relaxation at $z = 20$, with maximum ages shorter than their relaxation timescales. SIGOs with shorter cooling and collapse timescales than their fall-back timescales are marked with red stars as having the potential to form stars outside of a DM halo. As depicted, the main limitation on low mass SIGOs' lifetimes is the numerical evaporation process (i.e., 2-body relaxation, see text). Therefore, in our adopted resolution, high-mass SIGOs can collapse to potentially form stars, while in the Universe we expect that more SIGOs will form stars.

spectrum of GMCs in the outer Milky Way, and is broadly consistent with the range of mass spectra seen in GMC populations, despite their differing formation routes.

3.5 Timescale Analysis of the Evolution of SIGOs

The similarity between SIGOs and GMCs is limited when considering the main challenge for SIGOs: their location outside of a DM halo, and eventual fall back into a DM halo. In considering this challenge, and as a starting point for our timescale analysis, we identify the following evolutionary channels for SIGOs in a simulation (see Figure 3.6 for the relevant processes).

- *Gravitational Collapse*: SIGOs can undergo gravitational collapse in their overdense cores, or on the scale of the SIGO. We expand on this physical process in §3.5.1.
- *Cooling*: As mentioned above, molecular hydrogen cooling lowers the temperature of the gas, reducing the gas pressure and assisting with collapse. See §3.5.2 for more details.
- *Fall-back Into Halos*: SIGOs are inherently found near DM halos (Naoz & Narayan, 2014). Therefore, SIGOs are likely to eventually fall into a DM halo. Accretion is the most likely final state of even potential star-forming SIGOs, and could lead to the formation of accreted clusters. See §3.5.3 for an estimate of this timescale.
- 2-*body Relaxation:* Gas in `AREPO` numerical simulations has an associated mass that depends on the simulation resolution. This gas interacts gravitationally, causing it to undergo artificial two-body relaxation processes. These artificial interactions result in the evaporation of gas structures on a characteristic timescale. We estimate the timescale of this process in §3.5.4.
- *Growth*: SIGOs can grow by accreting gas and dark matter from their environment.

Over time, this changes their mass and gas fraction, and can impact their ability to collapse. We overview this process in §3.5.5

3.5.1 Gravitational Collapse

The timescale for gravitational collapse is important to determining whether a SIGO can achieve the overdensities needed to form stars. In order for this to occur, a SIGO needs to be able to collapse within a shorter time span than the fall-back timescale to nearby DM halos. If this condition is not satisfied, the SIGO will fall into a nearby DM halo before it forms stars, becoming disrupted by tides or external pressure. As a starting point for this timescale comparison, we calculate a free collapse timescale:

$$t_{\rm collapse} \approx \frac{1}{\sqrt{G\rho_{\rm peak}}} \ , \tag{3.6}$$

depicted in Figure 3.6 for a representative SIGO, and in Appendix C, Figure C.1 for a depiction of individual SIGOs. For the collapse timescale for a typical SIGO in Figure 3.6, we assume a representative range of peak densities, based on the range of densities in SIGOs at $z = 20$ in Run 2vH2 (see Appendix C for more details). The upper bound density is taken as

$$\log_{10}(\rho_{\rm max,com}) = 0.7 \times \log_{10}\left(\frac{\rm M_{SIGO}}{\rm M_\odot}\right) - 30.2. \tag{3.7}$$

The lower bound density is taken to be

$$\log_{10}(\rho_{\rm min,com}) = 0.4 \times \log_{10}\left(\frac{\rm M_{SIGO}}{\rm M_\odot}\right) - 29.6, \tag{3.8}$$

where ρ has units of g/cm^3. Here we use peak rather than mean densities to reflect that the SIGO's star formation is likely primarily occurring in its most dense regions: the process of star formation in a SIGO is not 100% efficient.

As seen in Figure 3.6, the collapse timescale has a negative dependence on mass, driven by larger central overdensities in more massive SIGOs. This mass dependence contributes to the enhanced ability of massive SIGOs to form stars.

3.5.2 Cooling

However, this timescale for free gravitational collapse does not paint the whole picture of the collapse of SIGOs. In order for a SIGO to collapse, it also must be able to efficiently cool, allowing its Jeans mass to decline and permitting gravity to overcome gas pressure. There are 2 phases to this cooling in a SIGO that has an initial Jeans mass above its actual mass: in the first phase, the SIGO isochorically cools, lowering its Jeans mass until it drops below the SIGO's mass. Secondly, the SIGO begins to collapse, with the dynamics of its collapse determined by the cooling and collapse timescales. To account for this, we can define a cooling timescale $\mathrm{t}_{\mathrm{cool}}$, given as the time it would take for a gas clump to cool isochorically at a rate $\Lambda(T, n_{\mathrm{H}})$ to its Jeans temperature, or, for SIGOs with supersonic velocity dispersions, to a temperature at which its density exceeds the critical density. We then add the timescale for the second phase of this collapse: the cooling time as defined in Schauer et al. (2021) and derived from Hollenbach & McKee (1979), taken from the temperature at which the SIGO initially collapses.

$$t_{\mathrm{cool}} = \frac{k_{\mathrm{B}} T_{\mathrm{collapse}}}{n_{\mathrm{H}}(\gamma - 1)\Lambda(T_{\mathrm{collapse}})} + t_{\mathrm{iso}}, \tag{3.9}$$

where γ is the adiabatic index and $\mathrm{t}_{\mathrm{iso}}$ is the additional cooling time a SIGO that initially does not exceed its Jeans mass takes to isochorically cool to lower its Jeans mass to $\mathrm{M}_{\mathrm{SIGO}}$

and begin to collapse:

$$\mathrm{t_{iso}} = \begin{cases} 0 & \mathrm{M_{SIGO}{\geq}M_J(T_{init})} \\ \\ \int_{\mathrm{T_{Collapse}}}^{\mathrm{T_{init}}} \frac{k_\mathrm{B}}{n_\mathrm{H}(\gamma-1)\Lambda(T)} dT\ . & \mathrm{M_{SIGO}{<}M_J(T_{init})} \end{cases} \tag{3.10}$$

If the SIGO initially exceeds the Jeans mass, its collapse temperature is taken as its volume-averaged temperature. Otherwise, its collapse temperature (after the SIGO has isochorically cooled), is taken as the Jeans temperature that the SIGO has cooled to:

$$\mathrm{T_J} = \frac{3\mathrm{G}\mu \mathrm{m_H}}{5\mathrm{k_B}} \sqrt[3]{\frac{36\mathrm{M}_{\mathrm{SIGO}}^2 \rho_{\mathrm{SIGO}}}{\pi^2}} \tag{3.11}$$

In SIGOs, Λ(T) is dominated by and approximated as the molecular cooling rate defined in Galli & Palla (1998):

$$\Lambda(\mathrm{T})_{\mathrm{H_2}} = \frac{\Lambda(\mathrm{LTE})}{1 + [n^{\mathrm{cr}}/n(\mathrm{H})]}, \tag{3.12}$$

consistent with the molecular cooling prescription used in GRACKLE. Note here that because the SIGO may accrete gas from its surroundings and gain mass as it cools (raising its Jeans temperature), because the SIGO also contracts as it cools on a similar timescale (Nakazato et al., 2022), and because this does not account for overdense regions within the SIGO which enhance cooling, this cooling timescale assuming an isochoric process is an upper bound on the true cooling time relevant to collapse.

We show the representative timescale based on Equation 3.10 in Figure 3.6 (see also Figure C.1 from Appendix C, for this timescale for individual SIGOs). To depict the typical value of this timescale we take the median initial gas number density of a SIGO, 1.08 $\mathrm{cm^{-3}}$. To represent the maximum cooling time, we take a lower bound gas density of 0.7 $\mathrm{cm^{-3}}$. For the lower bound cooling time, we take $\mathrm{n_H} \sim 2.0\ \mathrm{cm^{-3}}$ (see Figure C.2 from Appendix C).

We assume an initial temperature of 500 K or the Jeans temperature (Eq 3.11), whichever is higher, for this representative line. As depicted in Figure 3.6, most SIGOs at $z = 20$ have not yet cooled substantially.

3.5.3 Fall-back Into Halos

Eventually, the majority of SIGOs will be accreted onto a halo (most often their parent halo), forming globular-cluster-like systems. Like GCs, these are mostly expected to reside in the halos of their host galaxies (e.g., Chaboyer, 1999; Benedict et al., 2002; Chen et al., 2018). Naoz & Narayan (2014) calculated the timescale on which a SIGO free-falls to its parent halo:

$$t_{\rm ff} = 0.27{\rm Gyr}\left(\frac{\Delta r_{\rm phys}}{0.59{\rm kpc}}\right)^{\frac{3}{2}}\left(\frac{\Omega_m}{\Omega_b}\frac{\rm M_b}{10^6{\rm M}_\odot}\right)^{-\frac{1}{2}}, \tag{3.13}$$

where $\Delta r_{\rm phys}$ is the physical separation between the SIGO and the halo, and $\rm M_b$ is the baryon mass of the SIGO. In order for a SIGO to collapse and potentially form stars outside of a DM halo, this timescale must be longer than both the cooling and collapse timescales for that SIGO, permitting the object to cool and reach high central densities prior to being affected by tides or ram pressure stripping near the larger halo.

We show the representative timescale from Eq. (3.13) in Figure 3.6, where we assumed a median physical separation of 0.4 kpc between a SIGO and its nearest halo, with a range of $0.1 - 0.9$ kpc for a characteristic SIGO, representative of the true range of separations in the simulation at $z = 20$. As seen in this figure, some SIGOs "fall" into a DM halo. Figure 3.7 shows the evolutionary path of a single such SIGO (left of the figure with high baryon fractions) that falls back into a dark matter halo (the gas component of which forms the right branch of the figure, with low gas fractions). Notably, this particular SIGO was accreted by a larger nearby DM halo rather than its parent halo, allowing its lifetime outside of a halo to be unusually short for this process, only 15 Myr. We label the final outcome as "GC-like," indicating the potential for this object to form a star cluster at the outskirts of

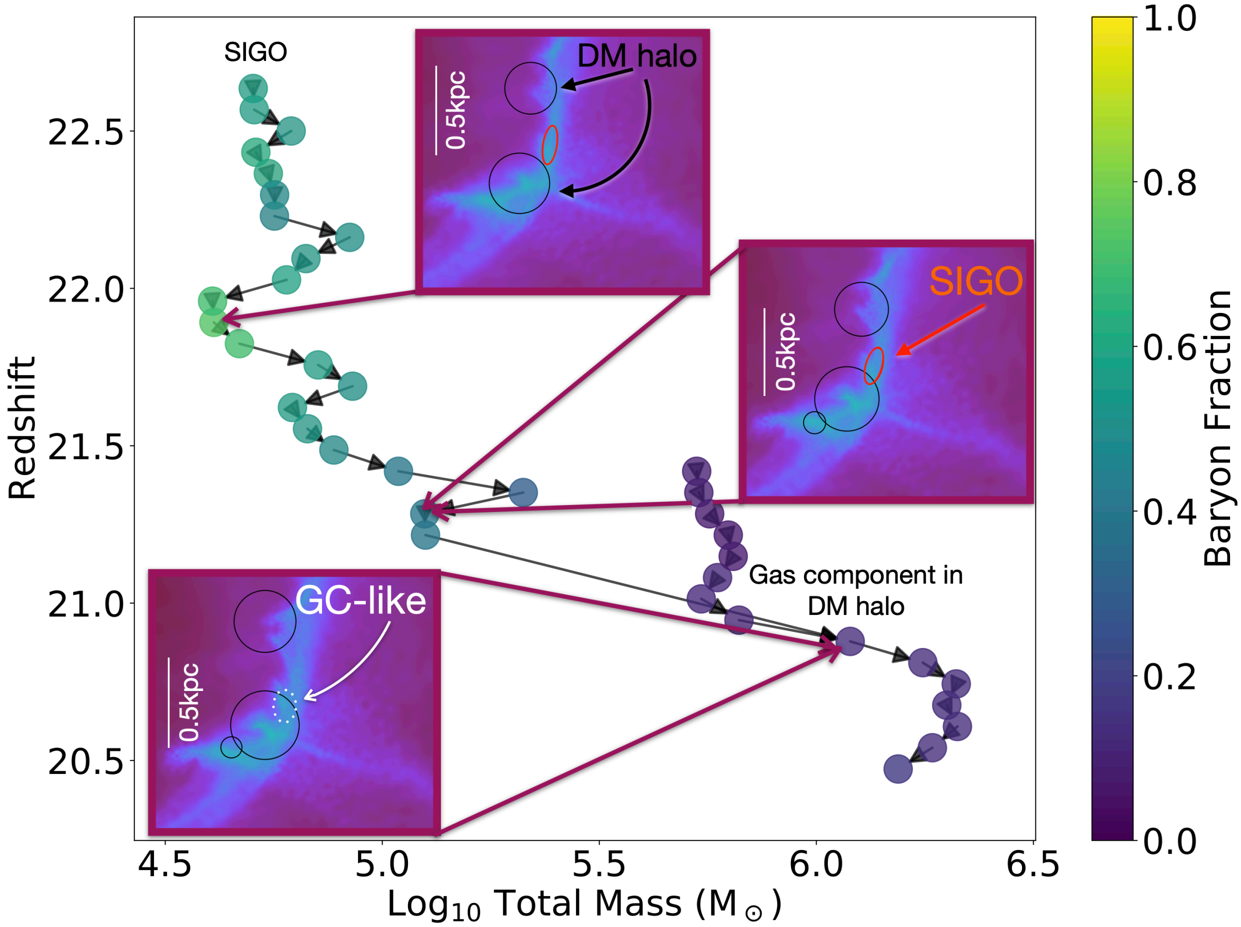


Figure 3.7: **Fall-back of a SIGO (the creation of a GC-like system at high redshift):** here we show the time evolution of a single SIGO in Run 2vH2. This SIGO, the left-most branch of this plot, is accreted by a nearby halo (the gas component of which is shown in purple, the right branch of the plot, with a very low baryon fraction). This accretion produces a more massive merged object, containing the SIGO in its substructure. This SIGO (labelled “GC-like" in the figure, for globular cluster-like candidate), will be subject to future evolution within the halo, and is identifiable as a distinct component, with the potential to continue to evolve into a star cluster. For a movie of this evolution see here.

the halo, similar to a present-day GC.

3.5.4 2-body Relaxation

In a simulation with limited resolution, each Voronoi cell has an associated gas mass. Therefore, gravitational interactions between these cells can lead to changing the object's energy by an order of itself. This processes, called 2-body relaxation, has been shown to be a possible limiting factor on simulations of SIGOs (Naoz & Narayan, 2014; Popa et al., 2016). This process essentially "evaporates" the object as a function of time. The 2-body relaxation timescale for N particles can be written as

$$t_{\rm rlx} \sim \frac{0.1\mathrm{N}}{\ln(\mathrm{N})} \times \frac{\mathrm{r}}{\sigma}, \tag{3.14}$$

(Binney & Tremaine, 2008) where σ is the 1D velocity dispersion within the SIGO and r is taken to be the minimum principle axis length (ergo r/σ is $\mathrm{t}_{\rm cross}$, such that this relaxation time describes the timescale on which the SIGO significantly changes in size). SIGOs can have as few as ~ 100 gas particles in these simulations, which combined with their small sizes and fairly high velocity dispersions leads to artificial relaxation times on the order of tens of Myr. This underscores the need for caution when following the evolution of SIGOs to low redshifts in these cosmological simulations: at our resolution, low-mass SIGOs with around 10^4 $\mathrm{M}_\odot$ of material are already beginning to be destroyed by artificial relaxation at redshift 20.

We depict the timescale from Eq. (3.14) in Figure 3.6 (see also Figure C.1), as a solid black line. As seen in the Figure, the main effect shortening the lifetime of our SIGOs is the relaxation timescale generated by our limited resolution. Lower-mass SIGOs left of the black dashed line at 10^5 $\mathrm{M}_\odot$ may be destroyed by this relaxation before $z = 20$. Thus, we expect SIGOs in the Universe to have increased longevity not captured by our simulation.

3.5.5 Growth

While constraining the full spectrum of growth timescales is challenging at present due to limitations of our spatial and time resolution, it is important to also note that SIGOs can grow through gas and dark matter accretion. As can be seen in Figure 3.3, SIGOs form embedded in gas streams–typical gas densities surrounding the forming SIGOs can be as high as $\sim 20\%$ that of the SIGO itself. In addition, the material around the SIGO may be itself enriched in baryons (though its gas fraction will not necessarily be as high as that of the SIGO). This gas, as well as some dark matter, may be accreted from a SIGO's environment, increasing its mass over time and allowing its gas fraction to change.

Figure 3.8 shows an example of this growth for a single SIGO in Run 2vH2. This SIGO formed with a mass of about 6×10^4 $\mathrm{M}_\odot$ and a gas fraction near 60%. The SIGO was able to accrete gas from its surroundings with a similar baryon fraction to the SIGO itself, growing by a factor of nearly 10 while maintaining a roughly constant gas fraction. This has the potential to impact the SIGO's future evolution, as the more massive SIGO at $z = 20$ is much more likely to quickly cool and collapse to form stars through the process outlined above in §3.5.2.

It is important to note, however, that this process of growth does not have to maintain a SIGO's gas fraction. Over time, accretion of surrounding material can cause the gas fraction in SIGOs to drop below the 60% threshold set for SIGO identification, even though the SIGOs still exist (and are commonly still enriched in gas). In Run 2vH2, about 50% of formed SIGOs at all redshifts were found to eventually drop below a gas fraction of 60% through this process before $z = 20$, while generally maintaining a gas fraction above 40% and maintaining or increasing their density. In future studies modelling star formation in SIGOs to lower redshifts, it will be important to follow these evolved SIGOs, which have higher masses than newborn SIGOs and may be promising grounds for star formation in SIGOs at lower redshifts.

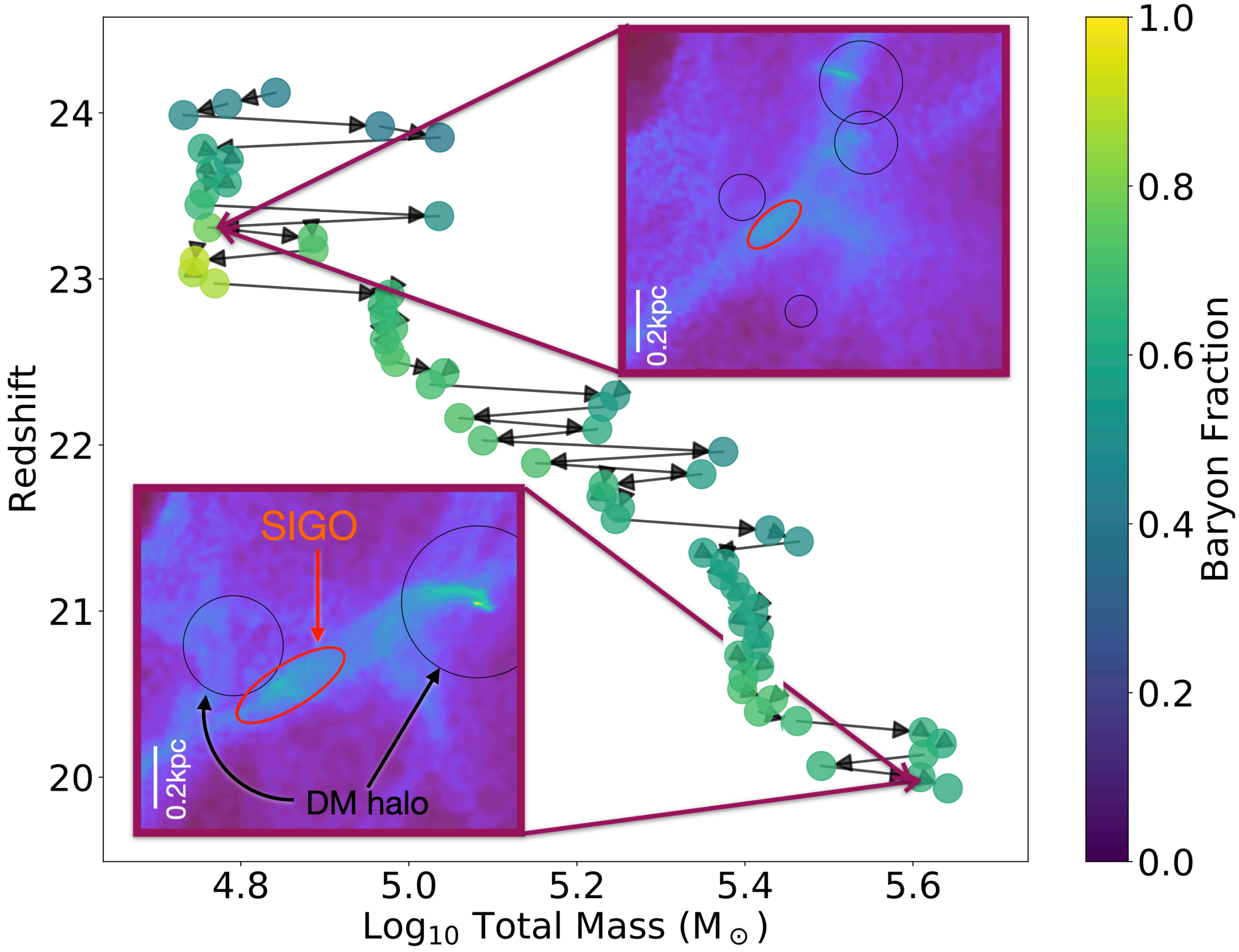


Figure 3.8: **Growth of a SIGO:** here we show the time evolution of a single SIGO in Run 2vH2. This SIGO increases its mass by nearly a factor of 10 between its formation and the final redshift of our simulation, accreting gas from its surroundings with a similar baryon fraction to the SIGO itself and thereby maintaining its baryon fraction over time. For a movie of this evolution see here.

3.5.6 Timescale Comparison

Figure 3.6 shows these characteristic timescales for the evolutionary channels for SIGOs mentioned above. In grey, one can see the characteristic timescale for gravitational collapse of SIGOs identified at $z = 20$ and at least one other snapshot[2] of Run 2vH2. In green, we show the range of fall-back timescales to the nearest halo in the simulation. In blue, we show the range of cooling times for the SIGOs to cool through molecular hydrogen cooling to Jeans collapse (for SIGOs with a Mach dispersion less than 1) or collapse through the critical density (for SIGOs with a Mach dispersion greater than 1). For all of these regions, we also add a dotted line, showing the timescale for a SIGO of a given mass with median properties (principle axis lengths, distance from the nearest halo, temperature, gas fraction, and velocity dispersion). We also add a black line showing a typical timescale for 2-body relaxation in the simulation at simulation resolution. The figure also shows the age and mass of each SIGO in Run 2vH2 at $z = 20$. The age is computed by comparing the first redshift at which each object meets the conditions to be identified as a SIGO (outlined in Section 3.2) to the time of the plot at $z = 20$. Note that SIGOs with an age of 0.45 Myr are those that have just formed in the simulation. The age is calculated by assuming that their formation time is halfway between the penultimate and final snapshots.

As one can see from the plot, the cooling timescale has a strong (roughly $\mathrm{M}^{-2.6}$) dependence on mass. Because of this, molecular cooling alone is not sufficient to efficiently cool SIGOs below about $\sim 10^5\ \mathrm{M}_\odot$ to collapse. In essentially all cases with such SIGOs, the cooling timescale will be dramatically longer than the collapse and fall-back timescales, so the SIGO will fall into a halo before forming stars.

Above this mass limit, one must also compare the timescales for a SIGO to fall back into a nearby halo (which may be its parent halo or another gravitationally dominant nearby

---

[2] Here we use multiple snapshots in order to limit false detections of SIGOs and better constrain their dynamical properties.

halo) or collapse gravitationally to determine SIGOs' fates. In isolation, high-mass SIGOs are frequently capable of collapsing to form stars; however, because of the wide variance of these timescales, it is possible for the fall-back timescale for a SIGO to be shorter than the collapse timescale. Because of this, the fall-back and collapse timescales must be individually compared, to determine which SIGOs will form stars before accretion onto nearby DM objects. Such SIGOs are marked in the plot with red stars.

One final consideration in studying the evolution of SIGOs apparent from this plot is the simulation's relaxation timescale. At the resolution of this study, even 10^5 $M_\odot$ SIGOs dissipate due to 2-body relaxation within 80 Myr–a similar length of time to that between the first SIGO's formation and our $z = 20$ snapshot. Because of this relaxation, we are undercounting the lowest-mass SIGOs at $z = 20$, as well as potentially underestimating the densities of those that exist. In order to eventually follow SIGOs through the process of forming stars and merging with halos at later times, future studies will need to work at higher resolutions (~ 200 $M_\odot$ or better), increasing the relaxation timescale.

3.6 Discussion

Understanding the behavior of supersonic turbulence in and around SIGOs is key to understanding their evolution and potential for star formation (Chiou et al., 2019). Here we show how supersonic turbulence acts together with molecular cooling in SIGOs to form density peaks, with high enough densities to become sites of star formation. This process is similar to that in GMCs, where turbulence has a scale-dependent effect on star formation, boosting densities on the sonic scale (e.g., Krumholz & McKee, 2005; Burkhart, 2018).

The structure and population-level properties of SIGOs are also similar to those of GMCs. For example, we find that SIGOs have similar substructures and masses to GMCs. In particular, we find that SIGOs have core-like structures and filaments (Figure C.2, or see the evolved SIGO in Figure 3.8), similar to those in GMCs (Mac Low & Klessen, 2004; Krumholz

& McKee, 2005; Mocz & Burkhart, 2018). Further, we find that the mass spectrum of SIGOs is consistent with the mass spectrum of GMCs in the outer Milky Way (see Figure 3.5 in §3.4).

In particular, we show for the first time in a simulation that SIGOs form on scales comparable to the sonic scale (Figure 3.4). Supersonic turbulence aid in the formation and collapse of SIGOs, with SIGOs forming as high density peaks. Molecular cooling lowers these objects' Jeans scale. When this process is sufficiently efficient, this lowers the Jeans scale below the sonic scale and permits collapse of these newly formed SIGOs. This may permit star formation in SIGOs.

Chiou et al. (2019) found that most SIGOs should form stars through atomic cooling alone. However, this result was called into question by Schauer et al. (2021), which included atomic and molecular cooling and did not find star formation sites outside of DM halos in a simulation. Nakazato et al. (2022), on the other hand, followed a SIGO in a zoom-in simulation and confirmed that the SIGO experienced Jeans collapse. The modifications in this paper to the prescriptions in Chiou et al. (2019) are a step towards reconciling these disparate results. Correcting the prescription for collapse in SIGOs based on their asphericity, we come to the conclusion that SIGOs should not form stars when considering only atomic hydrogen cooling. With molecular hydrogen cooling accounted for, SIGOs should be able to form stars. However, not all SIGOs undergo global collapse, exceeding the Jeans scale on the physical scale of the entire SIGO rather than in substructure (neglecting metal line cooling, which has the potential to enhance star formation in SIGOs but has not yet been studied in a simulation; see e.g. Schauer et al., 2021).

At $z = 20$, when most SIGOs have not yet cooled, about 5 SIGOs will be capable of undergoing Jeans collapse. Thus, at lower redshifts, we expect many SIGOs to collapse. We estimate a star-forming SIGO abundance of 0.63 Mpc^{-3} at $z = 20$. This abundance is comparable to the present-day local density of low-metallicity globular clusters (i.e., 0.44

Mpc^{-3}, Rodriguez et al. 2015). Because of the high Poisson uncertainty inherent to this figure, this suggests an explanation for the lack of star-forming SIGOs in Schauer et al. (2021). A smaller simulation box size and σ_8, combined with lower star-forming SIGO abundances than could be inferred from the literature at the time, may allow a $< 1.5\sigma$ Poisson fluctuation to yield a simulation box with no star-forming SIGOs outside of halos, neglecting metal line cooling. However, these results indicate that the abundance of star-forming SIGOs in the Universe may be quite high in regions with significant stream velocities. Because our local neighbourhood has an estimated 1.75σ stream velocity fluctuation (Uysal & Hartwig, 2022), which is 87.5% of the value of the stream velocity studied in this work, this is potentially an important contributor to the early structure of our own Local Group.

The effect of σ_8 on the redshift of SIGO formation must also be considered in contextualizing these results. Here, we have assumed an elevated σ_8, representative of a $\sim 2\sigma$ density fluctuation. This elevated value of σ_8 yields more power; i.e., it helps structure develop earlier (e.g., Greif et al., 2011; Park et al., 2020). As shown in Figure C.2 (see Appendix C), the comoving density of SIGOs before they collapse is not significantly dependent on redshift, so the pre-collapse physical density and therefore the cooling rate of SIGOs is dependent on their redshift of formation. In overdense regions such as those that form galaxy clusters, then, SIGOs can more efficiently cool and collapse than in underdense regions. This factor explains why SIGOs are not seen in regions of the Universe outside of galaxies: SIGOs formed too late to efficiently cool in the underdense regions of the early Universe that gave rise to these volumes. On the other hand, SIGOs form earlier and are more prone to form star clusters in overdense regions.

As highlighted in Figure 3.6, the resolution limitation yields an artificial age limit on SIGOs, which should not exist in the case of gas objects(e.g., Naoz & Narayan, 2014). Thus, we use a semi-analytical formulation to show that the evolution of SIGOs generally proceeds to one of two states: either gravitational collapse to form stars outside of halos (followed by accretion onto halos as a cluster), or fall back into a nearby DM object, where the SIGO can evolve as substructure within the halo. This can be expressed as a timescale problem: SIGOs with a shorter collapse timescale and cooling timescale than fall-back timescale (to the nearest DM object) should form stars outside of halos. Otherwise, we suggest that the SIGO can form a GC-like object, since GCs are often found near the edges of DM halos. Star formation may take place in this overdense gas at the outskirts of the halo. An example evolutionary path of a SIGO that has fallen back into a nearby halo is shown in Figure 3.7. This SIGO had a low mass and an unusually low fall-back timescale (owing to a massive nearby halo), and therefore was unable to collapse outside of a DM halo. We also show, however, that a low initial mass may not prevent a SIGO from reaching the masses needed from star formation outside of halos. In Figure 3.8, we show an example of a SIGO that grows in mass by a factor of nearly 10 over the course of the simulation, lowering its cooling time by 2 orders of magnitude. This presents the possibility that even SIGOs that form outside of halos at lower masses may be capable of cooling to form stars outside of halos through growth, and suggests the need in future work for an analytic model for understanding SIGO growth rates.

Future studies aiming to investigate the collapse of SIGOs should carefully consider the cooling timescale (Figure C.1) in determining a final redshift and resolution for their simulations. While individual SIGOs with unusually short cooling timescales may collapse prior to $z = 20$, a typical early SIGO forming at $z = 25$ with a typical cooling timescale of 100 Myr will not fully collapse until around $z = 17$ (or potentially later with lower values of σ_8). For similar reasons, it is important to account for the two-body relaxation timescale in simulations beyond $z = 20$, which may be comparable to 100 Myr at similar resolutions to those in this paper.

Explicit treatments of star formation with metal line cooling are needed to study the further evolution of SIGOs, including the size of their stellar populations. Exciting observational predictions may become possible with this next step, such as determining a half-light radius for SIGOs, and comparing it not only to present day GCs, but also to expected higher-redshift GC observations and potentially GC progenitor observations from JWST. These simulations will also begin to establish the efficiency of star formation in SIGOs, allowing us to understand characteristic maximum stellar population masses from the process. Stellar feedback, which is an important aspect of the later evolution of SIGOs, must also be considered, though it is beyond the scope of the current work. As these simulations develop, zoom-in simulations including radiative feedback should also be run to the present day, in order to give us a picture of the entire evolutionary history of a SIGO, allowing us to much more firmly connect these high-redshift objects to the structures we see in today's Universe.

CHAPTER 4

Star Formation in Early Star Clusters without Dark Matter

4.1 Introduction

Globular clusters (GCs) are very old ($\sim$ 13 Gyr, Trenti et al., 2015) and dense structures whose formation mechanism has long been debated (see e.g., Gunn, 1980; Peebles, 1984; Ashman & Zepf, 1992; Harris & Pudritz, 1994; Kravtsov & Gnedin, 2005; Mashchenko & Sills, 2005a; Saitoh et al., 2006; Gray & Scannapieco, 2011; Bekki & Yong, 2012; Kruijssen, 2015; Mandelker et al., 2018). Observations indicate that GCs are likely enriched in baryons relative to the overall Universe (e.g., Heggie & Hut, 1996; Bradford et al., 2011; Conroy et al., 2011; Ibata et al., 2013), and that older GCs possess properties that may distinguish them from younger GCs (see for a review Bastian & Lardo, 2018). These properties create some uncertainty regarding the formation of GCs within the hierarchical picture of structure formation. To address this uncertainty, several formation scenarios have been proposed in the literature.

One such formation mechanism for globular clusters posits that they formed at the high-efficiency end of normal galactic star formation, evolving from particularly dense giant molec-

---

The contents of this chapter appeared in Lake et al. (2023b) ular clouds (GMCs) (e.g., Elmegreen & Efremov, 1997; Kravtsov & Gnedin, 2005; Shapiro et al., 2010; Grudić et al., 2022). This is supported by observations of massive young clusters in the merging Antennae system (Whitmore & Schweizer, 1995; Whitmore et al., 1999). This picture is attractive in part because it naturally explains why GCs tend to be very old, from a time in the Universe when these particularly dense GMCs were more common. However, it is not obvious whether the age distribution of GCs from this model is compatible with observations, nor why the GC luminosity function appears similar across environments given this model.

A second theory suggests that GCs form inside DM halos, as suggested by Peebles (1984), which were then stripped by the tidal field of their present-day host galaxies (e.g., Bromm & Clarke, 2002; Mashchenko & Sills, 2005a; Saitoh et al., 2006; Bekki & Yong, 2012; van Donkelaar et al., 2023). The primary strength of this theory is its natural connection between the properties (for example, the total mass) of a galaxy's GCs with its dark matter. It intuitively explains the scaling of these properties with the halo mass, and explains GC ages. However, this theory struggles to explain the observed presence of stellar tidal tails from some GCs, as the extended DM halos this model predicts should help to shield the formed clusters from tidal effects (e.g., Grillmair et al., 1995; Moore, 1996; Odenkirchen et al., 2003; Mashchenko & Sills, 2005a).

In this work, we explore star formation within a third formation mechanism for GCs proposed by Naoz & Narayan (2014). In this theory, at the time of recombination, as baryons decoupled from the photon field and cooled, the average sound speed in the Universe dropped precipitously. This drop caused the average relative velocity[1] between baryons and DM in the Universe (about 30 km s^{-1}) to become highly supersonic (Tseliakhovich & Hirata, 2010; Tseliakhovich et al., 2011). Following recombination, in the standard model of structure formation, baryon overdensities began to collapse, driven by existing DM overdensities that

---

[1] Also known as the streaming velocity due to its coherence on few-Mpc scales

by this time were about 10^5 times larger than baryon overdensities (e.g. Naoz & Barkana, 2005b). The significant relative velocity between baryons and DM complicated this process. Naoz & Narayan (2014) showed analytically that a sufficiently large relative velocity between baryons and their parent DM halo would create a spatial offset between the collapsing baryon overdensity and its parent halo. This effect is particularly important to understanding our local Universe, as Uysal & Hartwig (2022) recently estimated that our Local Group formed in a region of the Universe with a high ($\sim 1.7\sigma$) value of the streaming velocity. In certain instances (especially at high gas masses, such as $\mathrm{M_{gas}} >$ few $\times\, 10^6\ \mathrm{M_\odot}$), the spatial offset produced by the effect is smaller than the virial radius of the parent DM halo, leading to an offset between the centers of gas and DM within halos. The resulting objects are known as Dark Matter + Gas Halos Offset by Streaming (DM GHOSTs Williams et al., 2023), and have unique morphological and kinematic properties.

In other instances, especially when $\mathrm{M_{gas}} \lesssim$ few $\times\, 10^6\ \mathrm{M_\odot}$, the spatial offset between DM and gas within these overdensities is large enough that the baryon overdensity collapses outside the virial radius of its parent halo. Naoz & Narayan (2014) showed that this process would create a gas object with a characteristic mass of 10^4 – a few $\times\, 10^6\ \mathrm{M_\odot}$, which would be depleted of dark matter. This would put the formed objects squarely in the mass range of globular clusters, and given their early Universe nature and presumably low metallicities (any metals they have before star formation must originate from pollution from nearby halos), is suggestive of a connection to the low-metallicity sub-population of GCs (Lake et al., 2021). As SIGOs form from pristine gas, this also could lead to star clusters formed partially or entirely of Population III stars. Because our Local Group likely formed in a region with a large streaming velocity, the evolved forms of these objects are theoretically expected to be present in the Milky Way (Uysal & Hartwig, 2022).

Follow-up studies of these objects–known as Supersonically Induced Gas Objects (or SIGOs)–found them in simulations (Popa et al., 2016; Chiou et al., 2018, 2019, 2021; Lake et al., 2021), and predicted distinctive observational signals from these objects (Lake et al., 2021). However, the connection between SIGOs and GCs is still not firmly established and depends in no small part on the star formation efficiency and stellar properties of SIGOs. Work by Nakazato et al. (2022) using hydrodynamic simulations established that SIGOs are indeed capable of forming stars, following one such SIGO in a zoom-in simulation to Jeans collapse and demonstrating that molecular hydrogen cooling is sufficient to form stars in SIGOs. Lake et al. (2023a) expanded upon this, providing initial estimates of the abundance of star-forming SIGOs and of the timescales important to their ability to form stars outside of halos. However, this work left open many questions about the properties of star-forming SIGOs, such as the efficiency of star formation in SIGOs that do form stars, and the fraction of SIGOs that form stars at redshifts later than z = 20.

When considering star formation in SIGOs, it is vital to consider molecular hydrogen cooling (e.g., Glover, 2013; Schauer et al., 2021; Nakazato et al., 2022). H_2 cooling allows the temperature of primordial gas clouds to lower to ~ 200 K, which lowers their Jeans masses to $\sim 1000\ M_\odot$, potentially allowing these primordial gas clouds to collapse and form stars (Yoshida et al., 2008). With these factors in mind, in the present *letter* we present the results of a suite of `AREPO` simulations including the streaming velocity and incorporating molecular hydrogen cooling, with the aim of constraining some of the properties of star-forming SIGOs, including their star-formation efficiency.

This *letter* is organized as follows: in Section 4.2 we detail the simulation setup. In Section 4.3 we discuss the bulk properties of star formation in SIGOs, as well as comparing star formation in SIGOs to more classical star formation within DM halos. Lastly, in Section 4.5 we summarize our work, as well as discuss avenues for future work to build on these results.

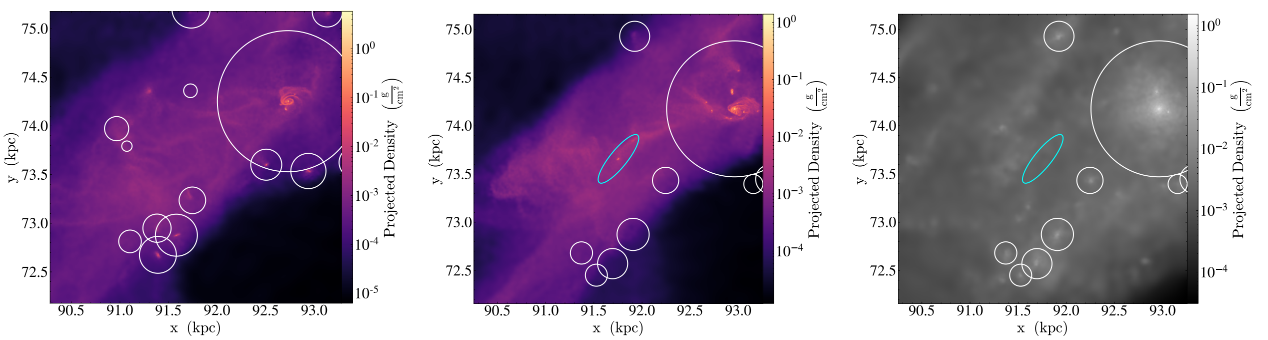
For this work, we have assumed a ΛCDM cosmology with $\Omega_\Lambda = 0.73$, $\Omega_M = 0.27$, $\Omega_B = 0.044$, $\sigma_8 = 1.7$, and $h = 0.71$.

4.2 Methodology

We use the moving-mesh code `AREPO` (Springel, 2010) for our simulations. We present two simulations with a 2.5 Mpc box size, 768^3 DM particles with mass $\mathrm{M_{DM}} = 1.1 \times 10^3\ \mathrm{M_\odot}$, and 768^3 Voronoi mesh cells with gas mass $\mathrm{M_B} = 200\mathrm{M_\odot}$, evolved from $z = 200$ to $z = 12$. Gas mesh cells in the simulation become eligible for collapse to form stars when their mass exceeds the Jeans mass on the scale of the cell. Subsequently, using the stochastic procedure described in Marinacci et al. (2019), eligible gas cells are converted into star particles on the free-fall timescale. Star particles are implemented as collisionless particles with the mass of the gas cell that gave rise to them.

We use a modified version of `CMBFAST` (Seljak & Zaldarriaga, 1996) to generate transfer functions for our initial conditions, incorporating first-order scale-dependent temperature fluctuations (Naoz & Barkana, 2005b), and the effect of the streaming velocity. Following the methods of Chiou et al. (2019, 2021); Lake et al. (2021); Nakazato et al. (2022); Lake et al. (2023a), we generate initial conditions setting $\sigma_8 = 1.7$. This choice allows us to simulate a rare, overdense region where structure forms early, which increases our statistical power. This environment is similar to those that form galaxy clusters. As discussed in Park et al. (2020), this choice is also similar in effect to increasing the redshift of structure formation compared to the Universe overall by a factor of $\sqrt{2}$. We would, for example, expect corresponding structures to those in our simulation at $z = 20$ to form in a region with the bulk properties of the Universe at $z \sim 14$.

We present 2 simulations in this paper, labelled as 2v and 0v. We use 2v and 0v here to indicate the stream velocity in the simulation. 2v simulations use a streaming velocity of $2\sigma = 11.8\ \mathrm{km\ s^{-1}}$ at the initial redshift $z = 200$, applied as a uniform boost to the x velocity of the baryons, as in Popa et al. (2016). 0v runs use the same initial conditions, but do not include a streaming velocity.



Gas Density without Streaming Velocity

Gas Density with Streaming Velocity

DM Density with Streaming Velocity

Figure 4.1: The gas density field around a star-forming SIGO (marked with a cyan ellipse) and local dark matter halos (marked with white circles) in our simulations at z = 15. The color scale shows the column mass density of matter in the box to a depth of 3 kpc, centered on the center of mass of the SIGO in run 2v. From left to right, we consider: **(a)** gas density without the streaming velocity (0v), **(b)** gas density with the streaming velocity (2v), and **(c)** the dark matter density with the streaming velocity (2v). All three panels show the same region. As can be seen in the middle panel, a SIGO is embedded in a larger shock (the high-density region of gas) and has a central high-density region/star formation site. The central SIGO here has a first generation of about 2×10^4 $M_\odot$ of stars by z = 15 in run 2v. The SIGO does not exist in run 0v (left).

We use the chemistry and cooling library `GRACKLE` (Smith et al., 2017; Chiaki & Wise, 2019) to track non-equilibrium chemical reactions and their associated radiative cooling explicitly in the gas. This includes molecular hydrogen and HD cooling, as well as chemistry for 15 primordial species: e^-, H, H^+, He, He^+, He^{++}, H^-, H_2, H_2^+, D, D^+, HD, HeH^+, D^-, and HD^+. The cooling rate of molecular hydrogen includes both rotational and vibrational transitions (Chiaki & Wise, 2019).

In this *letter*, we use the object classifications from Chiou et al. (2018) to identify SIGOs and DM halos. DM halos are identified using an FOF algorithm with a linking length that is 20% of the mean DM particle separation, or about 650 cpc. This algorithm calculates the location and virial radius of DM halos in the simulation output, assuming sphericity for simplicity (although DM halos at these times are distinctly aspherical e.g., Sheth et al., 2001; Lithwick & Dalal, 2011; Vogelsberger & White, 2011; Schneider et al., 2012; Vogelsberger et al., 2020). The same FOF algorithm run on the gas component of the output then allows us to identify gas-primary objects. Star particles are associated with the gas-primary object that their nearest-neighbor gas cell belongs to. We require these objects to contain a minimum of 100 gas cells and star particles to be considered as SIGOs (Chiou et al., 2021).

SIGOs are generally quite elongated in gas streams, so each gas-primary object is next fit to an ellipsoid, by identifying an ellipsoidal surface that encloses every particle in the object (Popa et al., 2016). These ellipsoids are tightened by shrinking their axes by 5% until either the ratio of the axes lengths of the tightened ellipsoid to those of the original ellipsoid is greater than the ratio of the number of gas cells contained in each, or 20% of their particles have been removed, following Popa et al. (2016). We then apply a final filter requiring that SIGOs be located outside the virial radius of all DM objects, and have a gas+stars mass fraction of above 60%, as in Nakazato et al. (2022) and Lake et al. (2023a) [2]. This limits

---

[2] Below the gas+stars mass fraction threshold of 60%, there are a variety of objects falsely identified by the FOF as SIGOs outside of halos in the no-streaming-velocity case (which may be nuclear star clusters, or compact gas objects in extended halos). At and above this threshold, many objects present in the run

false detections of SIGOs, as the filamentary nature of gas in runs with molecular hydrogen cooling enabled tends to result in the misidentification of SIGOs with a lower gas+stars fraction.

4.3 Star Cluster Formation in SIGOs

In the classical description of structure formation (i.e., no streaming velocity) stars often form inside, and at the centers of, DM halos (e.g., Tegmark et al., 1997). However, the streaming velocity acts to separate gas and eventually stars from the centers of these halos (e.g., Tseliakhovich & Hirata, 2010; Williams et al., 2023). In Lake et al. (2023a), we suggested that it may be straightforward for SIGOs to form stars. In this section, we analyze the evolution of a characteristic SIGO as it becomes a star cluster in the early Universe (z$\sim 17-12$).

Figure 4.1 shows this SIGO at the redshift ($z = 15$) of its peak star formation in the two runs, i.e., the streaming velocity run (2v, middle and right panels) and the no streaming velocity run (0v, left panel) at $z = 15$. As shown, the SIGO is not present in the run without streaming velocity (0v, left panel). While a small gas overdensity is present near the top left of the frame, it is associated with a DM overdensity missed by the FOF and the baryon fraction within 0.15 kpc of the center of the gas overdensity is 22%, confirming that it is not a spurious object related to the SIGO. At all snapshots before and during star formation in the streaming velocity runs, the SIGO contains less than 200× the mean DM density. The SIGO begins forming stars at $z \sim 16$ in run 2v.

As depicted in Figure 4.1, the predominant DM halos exist in both runs. However, as mentioned, the SIGO only exists in the presence of the streaming velocity and is embedded in the gas stream (e.g., Nakazato et al., 2022; Lake et al., 2023a; Williams et al., 2023). The

---

with the streaming velocity but absent in the run without it (true SIGOs) are removed, but very few objects falsely identified as SIGOs remain to be removed in the no-streaming-velocity run, as shown by Lake et al. (2023a).

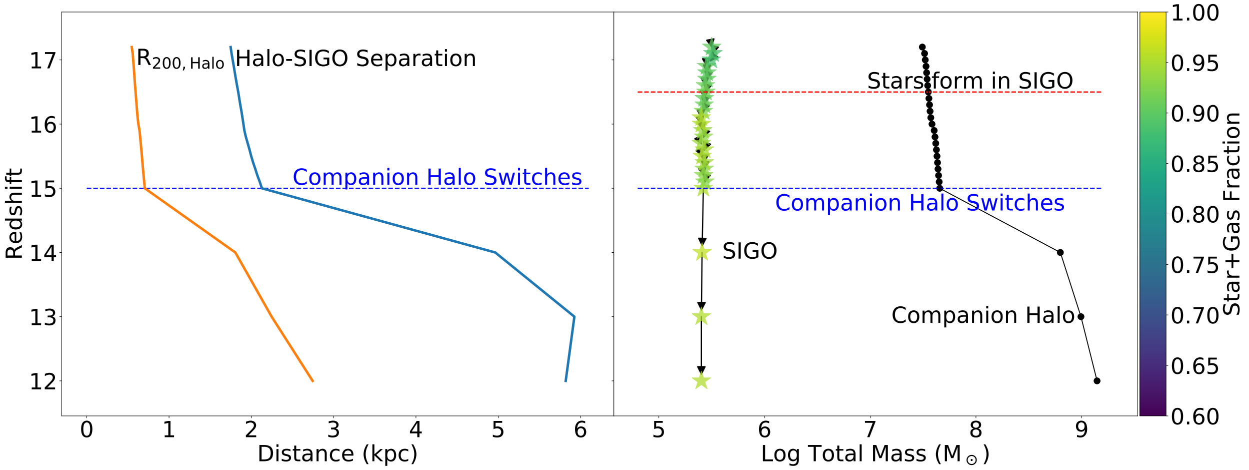


Figure 4.2: **The Evolution of a Star-Forming SIGO:** Here in the left panel we show the evolution of the R_{200} of the companion halo to the SIGO in Figure 4.1 (defined as the nearest halo to the SIGO as a function of the halo's R_{200}). We also show the separation between the centers of mass of the SIGO and companion halo. In the right panel, we show the evolution of the total mass of the SIGO and the companion halo. The SIGO is indicated with a star symbol in this panel. It begins forming stars at $z \sim 16.5$, as shown in the right panel. The black dots and lines in this panel indicate the mass evolution of the halo most closely associated with this SIGO at each redshift. The colors show the mass fraction of stars and gas within the SIGO compared to its total mass, showing that it maintains a high gas fraction throughout its evolution. Just after redshift 15, the companion halo that had previously been closest to the SIGO, as a function of its R_{200}, is supplanted by a slightly more distant, but significantly larger, halo. This transition is marked with a blue horizontal line labelled "Companion Halo Switches" in both panels.

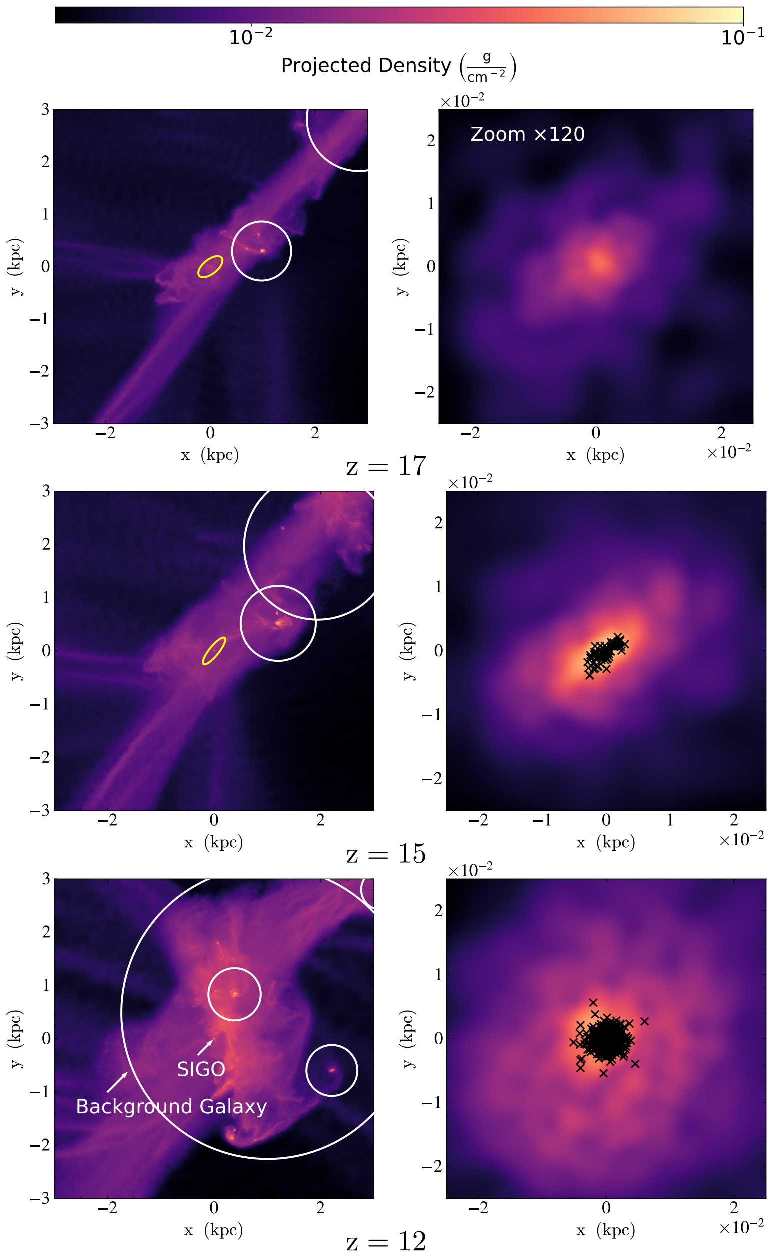


Figure 4.3: **The Formation of a SIGO-Derived Globular Cluster:** Here we show successive snapshots of a SIGO (yellow ellipse) becoming a star cluster interacting with DM halos (white circles) in our 2σ streaming velocity run. Stars are marked in the right, zoom-in panel of each snap with black Xs. In the top panel at $z = 17$, the SIGO is outside any DM halo and has not formed stars. By $z = 15$ in the middle panel, the SIGO is undergoing star formation. In the bottom panel at $z = 12$, the SIGO (indicated by a white arrow) is no longer identified as an independent gas object, caught on the edges of a protogalaxy merger. This SIGO's fate is likely as a bound star cluster in this newborn galaxy (marked as background galaxy, and also referred to as a protogalaxy in the text). Two other smaller halos are also present in the bottom panel.

SIGO in Figure 4.1 contains 2×10^4 $\mathrm{M}_\odot$ of stars in run 2v at z = 15. It is also apparent that the gas is shifted between the left and middle panels, similarly as noted by Nakazato et al. (2022) and Williams et al. (2023).

The SIGO's mass evolution can be seen in the right panel of Figure 4.2, and its surface density evolution is shown in Figure 4.3 in context. In Figure 4.2, the example SIGO (yellow-green star in the right panel) is associated with progressively larger nearby halos (in black), as the halos undergo a process of hierarchical mergers and accretion. The companion halo, defined as the nearest halo to the SIGO as a function of the halo's R_{200}, changes just after $z = 15$, marked in the Figure with a blue horizontal line. The SIGO's mass is comprised of both gas and stars and is only slightly shrinking, due to the loss of some gas (likely due to 2-body interactions between particles at our simulation's limited resolution, e.g. Lake et al., 2023a). We note that the stars and gas are gravitationally bound.

The left panel of Figure 4.2 shows the evolution of the physical separation between the centers of mass of the SIGO and its companion halos (in blue), as well as the evolution of the R_{200} of the SIGO's companion halos, defined as the radius around the halo that encloses 200 times the critical density of the Universe. As one can see, the SIGO is slightly drifting away from its first companion halo at and before $z = 15$, but the SIGO begins to fall into its second, much larger (and growing) companion halo by redshift 12. At $z = 12$, the SIGO is outside this halo at a center-of-mass separation $\Delta\mathrm{R} = 2.1\mathrm{R}_{200,\mathrm{Halo}}$, or about 6 kpc.

The SIGO and companion DM halos' evolutionary processes are visualized in the left column of Figure 4.3, which shows 6×6 physical kpc boxes, left column, and a zoomed-in region (50×50 pc) in the right column. The different rows show three redshifts: $\mathrm{z} = 17, 15$, and $\mathrm{z} = 12$. As shown in this Figure, the nearby large DM halo at $\mathrm{z} = 17$ is in the process of merging with a larger DM halo located at the right top corner of the image. This tidally separates the SIGO from the nearby DM halo: we see this DM halo slightly further away from the SIGO at $\mathrm{z} = 15$. This process may give the SIGO more time to cool and form stars outside of the halo before accretion, as well as limiting tidal forces from the halo on this SIGO. Stars start forming at z = 16.5 (see Figure 4.2), when the central 10 pc has a gas surface density of about 880 $M_\odot$ pc^{-2}. Significantly, tidal forces on this SIGO from its companion halo at this time are more than an order of magnitude smaller than the forces of self-gravity within this SIGO, allowing this SIGO's collapse to occur mostly unaffected by tides (e.g. Jog, 2013). These results are subject to the exact configuration of a SIGO, and it may be possible that tides impact the collapse of other SIGOs, though that is beyond the scope of this paper.

Subsequently, the halo is accreted onto another nearby protogalaxy (also referred to as the second companion halo), soon after z=12. In the bottom panel of Figure 4.3, the star cluster is located just outside of this large protogalaxy but has remained intact with the longest axis radius (from an ellipsoid fit) on the order of 6.6 pc, and shortest axis of 3.4 pc. This nascent star cluster is comprised in its entirety of stars that originated within the SIGO, and is now gravitationally bound to the 10^9 $M_\odot$ (total mass) protogalaxy, which also has stars of its own. The star cluster contains no dark matter and has a stellar mass of 7.4×10^4 $M_\odot$. We subsequently refer to this as a globular-cluster-like object.

In order to answer whether the SIGO is expected to be hosted by the protogalaxy, we ran a two-body simulation of the subsequent evolution of the SIGO and protogalaxy including cosmological expansion, confirming that the SIGO's orbital path enters the virial radius of the protogalaxy on a bound orbit. This analysis suggests that the SIGO is likely to fall within the virial radius of this protogalaxy within 100 Myr of the end of our `AREPO` hydrodynamic N-body simulation.

The end state of this SIGO as a cluster residing within a halo is commonplace. As shown in Lake et al. (2023a), star-forming SIGOs are expected to be accreted onto nearby halos shortly after forming stars. The hierarchical merging of these halos allows the most massive halos to accrete SIGO-derived star clusters. The protogalaxy which accretes this particular SIGO-derived star cluster is one of the largest protogalaxies in the simulation at all snapshots and could potentially host several SIGO-derived objects at later redshifts as it accretes nearby systems. On Gyr timescales, based on these results and those of Lake et al. (2023a), we expect more massive halos to host more SIGO-derived star clusters, gained through hierarchical formation. This particular cluster, as well as other low-mass star clusters derived from this process, eventually likely disperses through relaxation within its host halo on these Gyr timescales, while more massive clusters may survive (e.g., Naoz & Narayan, 2014).

An additional important property of this cluster is the degree of rotational support compared to random motion within its constituent stars. It is general consensus that local globular clusters are supported by random motion rather than ordered rotation, and this SIGO is not exceptional in that regard. Similarly to Chiou et al. (2018), we express the spin parameter as

$$\lambda_{\rm SIGO} = \frac{J_*}{\sqrt{2} M_* v_c R_{\rm max}}, \tag{4.1}$$

where M_* is the total stellar mass in the star cluster, v_c is the circular velocity at a distance $R_{\rm max}$ from the center of the cluster, and $R_{\rm max}$ is the maximum axis radius of the star cluster determined by an ellipsoid fit. At $z = 12$, this SIGO's star cluster has $\lambda_{\rm SIGO} \approx 0.072$, which is comparable to the spin parameter of many present-day globular clusters (e.g., Kamann et al., 2018).

Note that the streaming velocity for the 2v case is 0.7 km s^{-1} at z=12 (comparable to 59.5 km s^{-1} at the time of recombination, and 1 km s^{-1} at z$\sim$ 17, when the SIGO's overdensity initially formed). This streaming velocity injects turbulence into the gas, forming the SIGO (Lake et al., 2023a). As such, one expects a signature of this turbulence to be left in the SIGO's velocity dispersion. The velocity dispersion of the stars in the cluster is estimated as 2.3 km s^{-1} at z=12 (comparable to that of present-day globular clusters, e.g., Kamann et al., 2018), suggesting that other sources of velocity dispersion also play major roles in the cluster.

4.4 Detectability by JWST

At high redshifts, one primary mechanism by which objects like these SIGO-derived star clusters could be detected is UV emission from young massive stars (e.g., Sun & Furlanetto, 2016; Hegde & Furlanetto, 2023; Senchyna et al., 2023). As a proof of concept examining whether SIGOs could be observed with JWST, we use a semi-analytical model based on the example SIGO in this simulation to examine the flux from these young, massive stars in SIGOs forming at later redshifts. As a proof of concept, we model the emission from these test SIGOs, assuming that they share the star formation rate of our example SIGO[3], but form at various redshifts immediately prior to and during Reionization ($z = 6 - 10$). The UV luminosity of these SIGOs is given approximately by Sun & Furlanetto (2016):

$$\dot{M}_{\rm SFR} = \mathcal{K}_{\rm UV,1500} \times L_{\rm UV,1500} \ , \tag{4.2}$$

where we take $\dot{M}_{\rm SFR}$ to be the average star formation rate in our example SIGO in the 20 Myr period following the start of its star formation (after which, in a simulation with feedback, star formation may be quenched). See Appendix D for further discussion of the effect of this time duration. $L_{\rm UV,1500}$ is the rest-frame UV luminosity at 1500 Å. $\mathcal{K}_{\rm UV,1500}$ here is a fiducial constant, which following Sun & Furlanetto (2016) we set to be approximately $\mathcal{K}_{\rm UV,1500} = 1.15 \times 10^{-28}\ {\rm M_\odot yr^{-1}/ergs\ s^{-1}Hz^{-1}}$, which assumes a Saltpeter IMF. It is important to note here that this constant most likely underestimates the luminosity at a given star formation rate, as SIGOs have extremely low metallicities and would be likely to have a top-heavy IMF.

The solid lines in Figure 4.4 show the flux from test SIGOs with this star formation history, placed at varying redshifts. Three solid lines are shown, representing three different gravitational lens magnifications ($\mu = 10, 100$ and 1000, bottom to top). Characteristic

---

[3] Though note that the trend of star formation rates with redshift in SIGOs is not established, and could be impacted by decreasing gas densities with redshift.

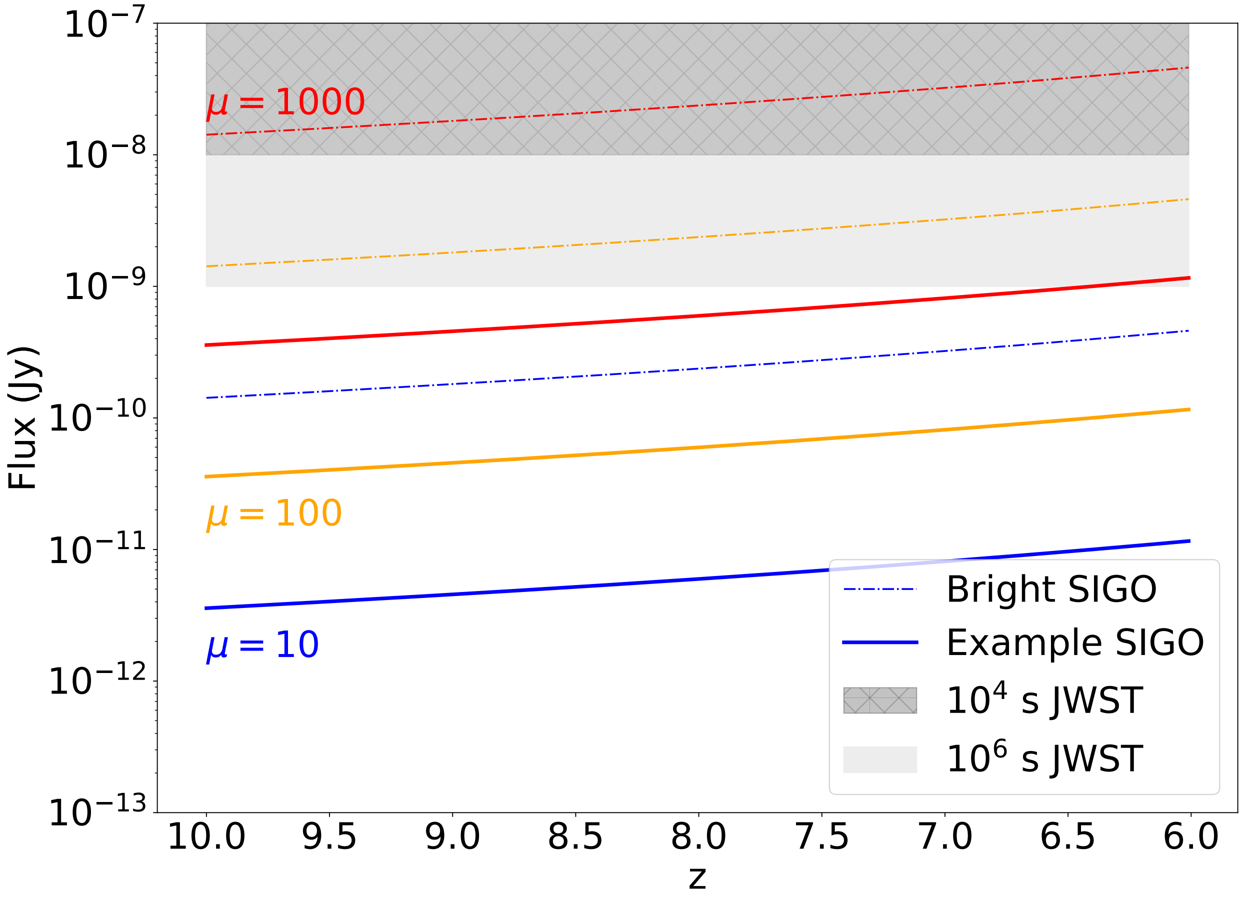


Figure 4.4: **SIGO flux as a function of redshift of initial star formation, and JWST detectability.** We consider the possibility of observing a lensed SIGO beginning star formation at later redshifts, similarly to the observation of a faint, lensed source in Welch et al. (2022). Solid lines show the modeled 1500 Å flux adopting the star formation rate estimate directly from our simulation (i.e., Figure D.1, in Appendix D). However, as discussed in Lake et al. (2023a), larger SIGOs more readily form stars; therefore, we also consider a more massive, larger SIGO, with a gas mass of $\sim 10^7$ $M_\odot$. We scale the star formation rate linearly as a function of mass (as implied from the SFR-mass relation, e.g., Lada et al., 2012). This example is shown in dot-dashed lines, labelled as 'Bright SIGO'. For the two examples we show three possible magnifications, of 10 (blue),100 (orange), and 1000 (red), from bottom to top. The light (dark) shadowed regime shows characteristic JWST sensitivity after 10^4 (10^6) seconds exposure time, with an SNR of about 10, in agreement with the JWST Exposure Time Calculator (Pontoppidan et al., 2016).

detectability from JWST at two exposure times is shown as gray-shaded regions. As seen in the Figure, an object with properties of our example SIGO, forming stars at the end of Reionization and magnified by a factor of a thousand or more, would only just be observable in a JWST field. UV emission from SIGOs similar to our example SIGO would likely not be observable with JWST at all.

However, the SIGO we study in this *letter* is likely not the most luminous possible SIGO, so to understand whether any SIGO would be detectable by JWST requires a model for the most luminous possible SIGO. The most massive SIGOs found have gas masses approaching 10^7 $M_\odot$ (Lake et al., 2023a), exceeding the gas mass of this SIGO at the start of star formation by a factor of 40. We argue that an approximate value for the star formation rate of the most luminous SIGO can, then, be given by multiplying the star formation rate of our example SIGO by this factor of 40 ratio in the SIGOs' gas masses (based on the SFR-mass relation, e.g., Lada et al., 2012). This assumes that the star formation efficiency of SIGOs of different masses is the same.

We overplot the simulated observed flux from this characteristic most luminous SIGO in Figure 4.4 with dashed lines. As one can see, in a very deep JWST field, strong lensing with a magnification of around 100 or better may allow a particularly luminous SIGO to be observable even at very high redshift, given sufficient exposure time. This result is not surprising, in the context of the recent observations of even individual stars or star systems at high redshift in such lensed fields (e.g., Welch et al., 2022).

4.5 Discussion

SIGOs (Supersonically Induced Gas Objects) are a natural consequence of early structure formation in ΛCDM (Naoz & Narayan, 2014). These gas objects form in the early Universe ($z \gtrsim 10$), with little to no dark matter, in the patches of the Universe with non-negligible streaming velocity between the dark matter and the baryons. Interestingly, it was recently suggested that our own Local Group formed in a region with a large streaming velocity (Uysal & Hartwig, 2022), implying that the small-scale structure in our vicinity was greatly impacted by the streaming velocity. Investigating the star formation of these objects is critical for predicting future local and high redshift observations. While previous studies expected that these objects would eventually form stars (e.g., Chiou et al., 2019; Nakazato et al., 2022; Lake et al., 2023a), until now, no study investigated the formation of stars in these objects.

Here we present, for the first time, a study of the outcomes of star formation in SIGOs. We estimate the stellar mass of SIGOs, and follow the evolution of a SIGO after star formation. We present a 2σ streaming velocity run 2v, and a control run without the streaming velocity (0v) for comparison. See Figure 4.1 for a comparison of these runs.

We find that some SIGOs form stars. As expected (see Lake et al., 2023a), many other SIGOs accrete onto nearby DM halos prior to forming stars, forming diffuse structures akin to DM GHOSts (Williams et al., 2023). In total, out of 5325 unique SIGOs found at integer redshifts, 9 SIGOs formed stars outside of a DM halo before z = 12 [4], and others form DM GHOSt analogs or may form stars later (Williams et al., 2023). We explore the population and overall star formation efficiency in Lake et al. in prep.

Here we focused on an example SIGO as a detailed case study of the formation and evolution of a GC-like star cluster. Figure 4.1 depicts this process in context, showing that simulation runs with the streaming velocity effect form a gas object in a location where there is none in the no-stream-velocity runs, and that the object is capable of cooling to form stars while fully outside of nearby DM halos.

Figure 4.3 depicts the evolution of a SIGO and its vicinity as it forms stars. By z = 12,

---

[4] The latter number is visually verified, ensuring that the SIGO is outside of concentrations of dark matter missed by the FOF and that the candidate SIGO did not originate/form stars in the nucleus of a halo at earlier times. We also verify that the star particles formed within the SIGO.

the resultant GC-like star cluster is bound to a large, 10^9 $M_\odot$ protogalaxy (although it lies beyond its R_{200}, as shown by Figure 4.2). Based on a simple two-body simulation, the SIGO is likely to merge with the halo to form a bound cluster within the halo on a timescale of about 150 Myr following the end of our `AREPO` hydrodynamic N-body simulation.

The SIGO formed stars at $z = 16.5$ while still outside its closest DM halo. The SIGO was separated from this nearby halo through a tidal interaction with a third, larger halo, possibly allowing it to remain outside of the nearby halo as it formed stars. Subsequent evolution may yield a cluster that resembles a cluster resultant from a more classical evolution (e.g. Sameie et al., 2022). The SIGO-derived clusters may have some differing characteristics, such as their velocity profiles (e.g., Williams et al., 2023). The resulting cluster could also have a high galactocentric distance compared to clusters that formed locally, owing to its accreted nature.

Note that in this study we do not include feedback or the effects of metal enrichment on the SIGO. Although radiative feedback will act to reduce the efficiency and duration of star formation within the SIGO, even low levels of metal enrichment can significantly increase cooling rates within it (such as from pair-instability supernovae in nearby halos, Schauer et al., 2021). In addition, the limited mass resolution of these simulations results in a significant underestimation of the effectiveness of molecular hydrogen cooling in SIGOs (Nakazato et al., 2022). Furthermore, our softening scale of 2.2 pc is a significant fraction of the radius of the star cluster, which is expected to lower our simulated star formation rates, acting together with our underestimation of the effectiveness of molecular hydrogen cooling to make star formation in SIGOs seem slower than it is. Taken together, these processes will have an ambiguous effect on the stellar mass of SIGOs. Higher stellar masses will increase both the longevity of star clusters derived from SIGOs as well as the potential for binary black hole mergers and gravitational wave events as these clusters evolve.

Already, early JWST observations have found potential candidates for high-z clusters, some with high galactocentric distances (e.g. Pascale et al., 2022; Senchyna et al., 2023). Figure 4.4 explores this possibility, considering whether a SIGO-derived star cluster with properties similar to the cluster explored in this letter would be observable by JWST in the Reionization epoch and whether a SIGO that is especially massive and bright may be observable in the same epoch. We find that a SIGO with properties similar to the one in this letter would likely be unobservable with JWST (even considering lensing). However, a more massive SIGO similar to the most massive SIGOs in simulations may be detectable via lensing, presenting the possibility of direct detections.

CHAPTER 5

Early Star Formation with the Streaming Velocity

5.1 Introduction

The first stars in the Universe, known as Population III (Pop III) stars, formed at zero metallicity and are thought to have very different properties from the stars we see today (Bromm, 2013). These metal-poor stars formed from gas clouds cooled primarily via radiative transitions of molecular hydrogen (H_2). Through this cooling mechanism, primordial gas clouds could lower their temperatures to ~ 200 K, corresponding to a Jeans mass of ~ 1000 $M_\odot$. Sufficiently massive primordial gas clouds could then collapse to form stars (Yoshida et al., 2008). Because H_2 cooling is less efficient than high-metallicity cooling at very low temperatures, this limits the minimum halo mass at which early Universe structures can form stars (Haiman et al., 1996; Tegmark et al., 1997; Abel et al., 2002; Bromm & Clarke, 2002; Yoshida et al., 2003b; Lake et al., 2023a). These stars were the primary sources of the first metals in the Universe, and as such were vital to later galaxy formation (Ferrara et al., 2000; Madau et al., 2001). Understanding the environments in which these first stars formed, and when they formed, will inform our understanding of these later galaxies.

Upcoming and ongoing observations from JWST may have the potential to observe some

---

The contents of this chapter appeared in Lake et al. (2024b) of these first stars, or the pair-instability supernovae they may produce (Johnson, 2010; Whalen et al., 2012; Visbal et al., 2016; Lake et al., 2023b). Indirect signatures of Pop III stars, such as the 21-cm line, also provide a promising means by which to explore Cosmic Dawn (Mebane et al., 2020; Magg et al., 2022; Hassan et al., 2023). Theoretical models allow us to interpret these observations and maximize their benefit, so there is an immediate need for detailed theoretical models of Pop III star formation (Menon et al., 2024).

A key environmental effect affecting the formation of Pop III stars is the relative motion of baryons and dark matter in the early Universe, known as the streaming velocity (Tseliakhovich & Hirata, 2010; Naoz et al., 2013; Kulkarni et al., 2021; Hegde & Furlanetto, 2023). Prior to Recombination, while the growth of baryon overdensities was suppressed by the field of photons, dark matter exhibited bulk flows towards large scale overdensities. These flows were coherent on few-Mpc scales, and their magnitude varied following a Maxwell-Boltzmann distribution. At the time of Recombination, as the temperature of the baryons dropped precipitously, these bulk flows (around $\mathrm{v_{bc,rms}} = 30\ \mathrm{km\ s^{-1}}$ at $\mathrm{z}= 1100$) became highly supersonic (Tseliakhovich & Hirata, 2010; Tseliakhovich et al., 2011). In regions with high values of this streaming velocity, gas is capable of advecting from its parent halo, and even when accreted, may exhibit lower densities and star formation rates (Tseliakhovich & Hirata, 2010; Naoz & Narayan, 2014; Williams et al., 2023). Low-mass ($\mathrm{M_{DM}} < 10^8 \mathrm{M_\odot}$) halos in these regions at high redshifts (also known as Dark Matter + Gas Halos Offset by Streaming, or DM GHOSTs) may have their star formation suppressed, as well as exhibiting unique morphological and rotational properties (Williams et al., 2023). As higher values of this streaming velocity are correlated with larger overdensities, this will preferentially impact galaxies and galaxy cluster progenitors at high redshift.

A further impact of this streaming velocity at high redshift is the tendency for gas overdensities to advect out of their parent halos, allowing the formation of gas-enriched objects known as Supersonically Induced Gas Objects, or SIGOs (Naoz & Narayan, 2014; Popa et al., 2016;

Chiou et al., 2018; Lake et al., 2021). These SIGOs are capable of star formation outside of dark matter halos (Chiou et al., 2019, 2021; Nakazato et al., 2022; Lake et al., 2023a,b). They also form in great abundance in the early Universe, with SIGO abundances approaching the abundance of present-day low-metallicity globular clusters by the epoch of Reionization (Lake et al., 2021). As these SIGOs are composed of nearly pristine gas, this could be an additional source of Pop III star formation prior to Reionization (Lake et al., 2023a,b).

In this paper, we aim to study the population-level effects of streaming on halos and SIGOs by examining the Kennicutt-Schmidt relation and star formation rate densities in a pair of small-box simulations with and without streaming. The paper is organized as follows: in Section 5.2 we discuss the setups of the simulations used. In Section 5.3, we discuss the effects of streaming on DM halo star formation, including the Kennicutt-Schmidt relation. In Section 5.4 we show and contextualize population-level statistics on star formation in SIGOs. Lastly, in Section 5.5 we summarize our results and suggest future avenues of exploration for star formation under the influence of streaming.

For this work, we have assumed a ΛCDM cosmology with $\Omega_\Lambda = 0.73$, $\Omega_{\rm M} = 0.27$, $\Omega_{\rm B} = 0.044$, $\sigma_8 = 1.7$, and $h = 0.71$.

5.2 Methodology

In this paper, we present the results of two `AREPO` (Springel, 2010) simulations, each with 768^3 DM particles with mass $M_{\rm DM} = 1.1 \times 10^3$ $M_\odot$, and 768^3 Voronoi mesh cells with gas mass $M_{\rm B} = 200 M_\odot$. Gas cells become eligible to form stars when their mass exceeds the Jeans mass on the cell's scale. Eligible gas cells are converted into star particles on the free-fall timescale. When the free-fall timescale is longer than the simulation timestep, this is implemented as a stochastic process as described in Marinacci et al. (2019). The formed star particles have the mass of the gas cell that gave rise to them, and are collisionless.

The simulations have a 2.5 Mpc box size, and are evolved from $z = 200$ to $z = 12$. Initial conditions are generated using transfer functions from a modified version of CMBFAST (Seljak & Zaldarriaga, 1996), which incorporates first-order scale-dependent temperature fluctuations (Naoz & Barkana, 2005b) and the streaming velocity. In line with the methods of Chiou et al. (2019, 2021); Lake et al. (2021); Nakazato et al. (2022); Lake et al. (2023a), we use $\sigma_8 = 1.7$ to generate our initial conditions, simulating a rare overdense region where structure forms early, similar to regions that form galaxy clusters. This enhances our statistical power. One simulation uses a $2\sigma_{\rm vbc} = 11.8$ km s^{-1} streaming velocity at the initial redshift $z = 200$, applied as a uniform boost to the x velocity of the baryons, as in Popa et al. (2016). The other simulation does not include a streaming velocity effect.

Both simulations explicitly include non-equilibrium molecular hydrogen chemistry and its associated radiative cooling, using the chemistry and cooling library GRACKLE (Smith et al., 2017; Chiaki & Wise, 2019). This includes molecular hydrogen and HD cooling, as well as chemistry for 15 primordial species: e^-, H, H^+, He, He^+, He^{++}, H^-, H_2, H_2^+, D, D^+, HD, HeH^+, D^-, and HD^+. The cooling rate of molecular hydrogen includes both rotational and vibrational transitions (Chiaki & Wise, 2019). It is important to note that we do not include metal cooling, which enhances cooling, especially in larger halos that continue to form stars after supernovae. We also do not include Lyman-Werner, radiative, or supernova feedback, which lowers star formation rates by order of magnitude, especially in low-mass halos (see e.g., Xu et al., 2016). Thus, halos smaller than $\sim 10^8$-10^9 $M_\odot$ may undergo one burst of star formation. Notably, the highest halo mass in our simulations (few $\times 10^9$ $M_\odot$) indicates the overall trend we expect. We highlight that the comparison to the observations is done at the high-mass end.

We use the object classifications from Chiou et al. (2018) to identify SIGOs and DM halos. We use a friends-of-friends (FOF) algorithm with a linking length that is 20% of the mean DM particle separation, or about 650 cpc, to identify DM halos. This gives us the locations and virial radii of DM halos in the simulation, assuming sphericity for simplicity (although it is important to note that DM halos at these times can be ellipsoidal e.g., Sheth et al., 2001; Lithwick & Dalal, 2011; Vogelsberger & White, 2011; Schneider et al., 2012; Vogelsberger et al., 2020). We also run the same FOF algorithm on the gas component of the output, with stars as a secondary component. This identifies gas-primary objects. In order for these objects to be considered as possible SIGOs, we additionally require that they contain at least 100 combined gas and star particles (Chiou et al., 2021).

Following (Popa et al., 2016), we fit gas-primary objects to an ellipsoid, by calculating the smallest ellipsoidal surface that encloses every particle in the object. This is necessary because these gas objects are generally quite elongated within highly non-spherical gas streams. We tighten the ellipsoids by shrinking their axes by 5% until either the ratio of the axes lengths of the tightened ellipsoid to that of the original ellipsoid is greater than the ratio of the number of gas cells contained in each, or 20% of their particles have been removed. Finally, in order to be identified as a SIGO we require that the center of mass of these ellipsoids must be located outside the virial radius of nearby DM halos, and that the object must have a baryon fraction above 60%. These constraints are necessary to effectively distinguish SIGOs from other classes of gas objects (Nakazato et al., 2022; Lake et al., 2023a).

Finally, because there is some degeneracy between true star-forming SIGOs and errors in the FOF algorithm, we verify each star-forming SIGO visually. The gas and DM density fields around each candidate SIGO are visualized side-by-side, and we visually verify that the gas overdensity associated with each SIGO is spatially offset from nearby DM overdensities. We also visualize the same spatial region from the run without streaming, ensuring that the star-forming SIGO is not present in that (control) run. This gives us confidence that we are identifying true star-forming SIGOs.

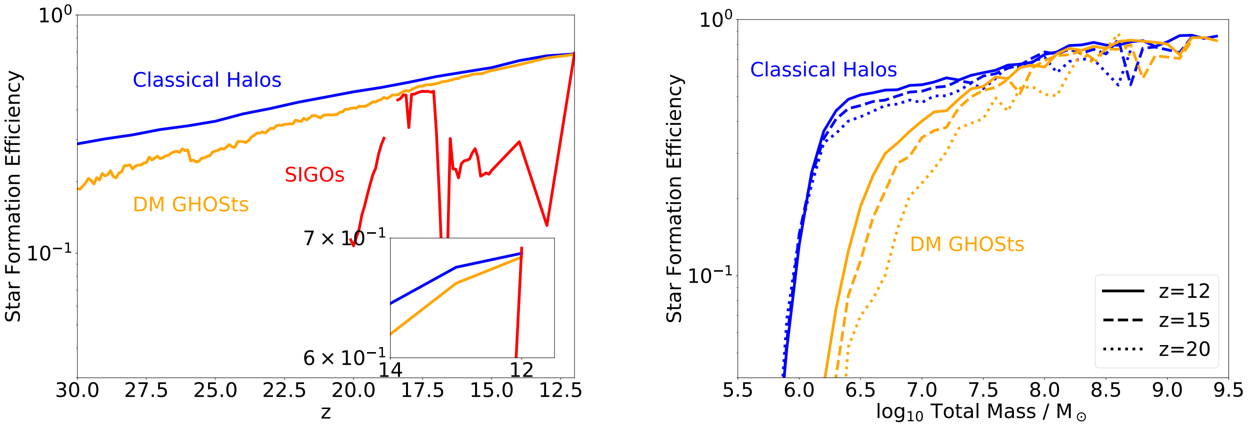
We classify all non-SIGO objects that form in our simulations with streaming as DM GHOSts. Corrections for the enhanced σ_8 in halos both with and without streaming for our star formation rate density plots are made via analytic calculations of the abundance of halos at various masses and redshifts using the methods described in Lake et al. (2021). These analytic calculations are subsequently used to estimate a redshift-, streaming-, and mass-varying abundance correction to $\sigma_8 = 0.826$, which is applied to our SFR densities. Gas objects are matched to nearby host halos, and these halos' masses are used in the calculations.

In order to identify stellar components of halos, we run a gas-and-star-primary FOF, and match objects from it to their nearest halos. This allows us to compute stellar masses of halos by summing over associated objects, ensuring that we are identifying even star particles that are just outside their parent halos. In all objects, star formation rates are determined by the formation time of star particles in the simulations. Each new star particle is matched to its host object at the first snapshot after it forms. The implied mass of new stars formed in a given object is divided by the time between snapshots to determine a star formation rate. This represents an average SFR over the timeframe between snaps.

5.3 Star Formation in DM GHOSts

5.3.1 Star formation efficiency

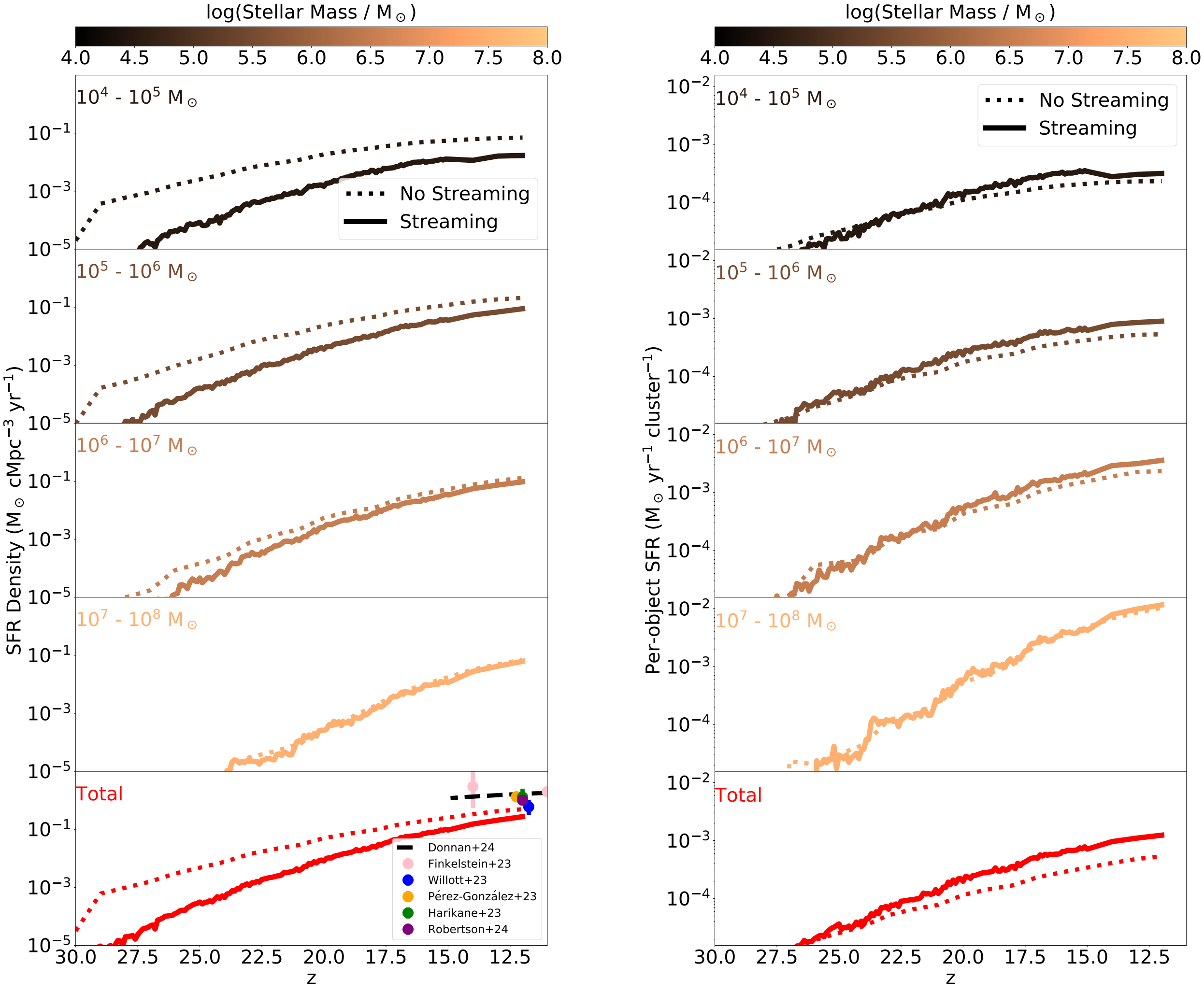
DM GHOSts are the counterparts to classical halos under significant streaming velocities and potentially are the ancestors of the oldest dwarf galaxies in our Local Group (Williams et al., 2023, 2024a). In particular, using observations of nearby dwarf galaxies, it has been pointed out that the local environment probably represented a high streaming patch of the Universe (Uysal & Hartwig, 2022). DM GHOSts comprise a stellar mass range that extends far beyond that of SIGOs and, therefore, are readily detectable by today's JWST. Predictions of their star formation history may soon be directly testable. The common prediction in the literature suggests that the streaming velocity suppresses the star formation in DM GHOSts, for DM halo masses $\lesssim 10^{8-9}$ $M_\odot$ (e.g., Tseliakhovich & Hirata, 2010; Fialkov et al., 2012; O'Leary & McQuinn, 2012; Bovy & Dvorkin, 2013; Tanaka et al., 2013; Tanaka & Li, 2014;



(a) Star formation efficiency binned by redshift

(b) Star formation efficiency binned by mass

Figure 5.1: **Star formation efficiency in various object classes:** here, we show the star formation efficiency of classical halos (halos in our no-streaming simulation, blue), DM GHOSts (halos in our $2\sigma_{\rm vbc}$ streaming simulation, orange), and in the left panel, SIGOs (in our $2\sigma_{\rm vbc}$ streaming simulation, red) as a function of redshift. In the right panel, we show the efficiency of star formation in classical halos (blue) and DM GHOSts (orange) at $z = 12$ (solid lines), $z = 15$ (dashed lines), and $z = 20$ (dotted lines) as a function of mass, binned by 0.1 segments in $\log_{10}$ of total mass of all types of matter. At high redshifts, DM GHOSts have substantially suppressed star formation compared to classical halos, but they catch up to the overall star formation rates in halos by $z = 12$ (owing in part to their suppression of the abundance of small-scale structure in comparison to larger halos). SIGOs have generally lower overall star formation efficiencies, likely owing to their lack of dark matter.



(a) Summed SFR Density in Mass Bin

(b) Average SFR Density in Mass Bin

Figure 5.2: **Star formation rate density in stellar objects of different masses:** Here we show plots comparing star formation rate densities with (solid lines) and without (dotted lines) streaming at different object masses. The left panel shows population-level statistics and the right panel shows averaged (per-object) statistics. In the left panel, only low-mass object SFRs are substantially lowered by streaming at the lowest redshifts studied on the population level, up to our low-mass resolution limit of about 10^5 M$_\odot$. Observational constraints from Donnan et al. (2024), Finkelstein et al. (2023), Willott et al. (2023), Pérez-González et al. (2023), Harikane et al. (2023), and Robertson et al. (2023) are included for comparison in the bottom panel of the left figure, which shows the total SFR density in all objects up to the highest-mass object in our box, representing halo masses from 10^5 M$_\odot$ up to the highest-mass halo in our box, 2×10^9 M$_\odot$. In the right panel, we see that at later times, star formation is actually more rapid in objects of a given stellar mass with streaming than without, as the objects catch up to their no-streaming counterparts. Note that the SFR curves in this Figure were normalized to represent $\sigma_8 = 0.826$ using Equation (E.3). The comparison to the observed SFR for the $\sigma_8 = 1.7$, used in the simulation in shown in Figure E.1. We draw the attention of the reader to the striking agreement between the observations and simulation in the higher σ_8 case. See Figure 5.3 for further analysis.

Park et al., 2020; Schauer et al., 2023; Hegde & Furlanetto, 2023; Conaboy et al., 2023).

However, Williams et al. (2024a) recently showed an additional, more subtle, effect of streaming on star formation. Although in DM halos with halo masses $\lesssim 10^7$ M$_\odot$ the number of stars and the star formation efficiency is suppressed at z$\sim$ 12 at the population level, the abundance of stars in larger DM halos is not suppressed. This implies a comparable star formation efficiency and a higher star formation rate in high-mass DM GHOSTs compared to classical structures at z$\sim$ 12.

In Figure 5.1, we depict the star formation efficiency, defined as the mass in stars divided by the baryonic mass in a halo, i.e., $M_\star/M_{\rm baryon}$. The left panel of this figure shows the average efficiency over all structure masses as a function of redshift. As shown, the star formation efficiency in halos at $z \sim 30$ is about twice as low in halos formed in a simulation with a $2\sigma_{\rm vbc}$ streaming velocity than in classical halos. This suppression originates from the lower gas densities within halos resulting from streaming at high redshifts, which subsequently suppresses the formation of molecular hydrogen, and thus cooling in low-mass halos.

In the left panel of Figure 5.1, the star formation efficiency is comparable at $z = 12$ in DM GHOSts to that of classical halos (consistent with Williams et al., 2024a). To explain this further, in the right panel of this figure we show the star formation efficiency in DM GHOSts (orange) and classical halos (blue) at $z = 20$ (dotted lines), $z = 15$ (dashed lines) and $z = 12$ (solid lines). In the simulation with streaming, the streaming velocity preferentially suppresses the formation of low-mass structures (as expected from analytical work by e.g. Tseliakhovich & Hirata, 2010). As a result, higher-mass halos, which have higher star formation efficiencies and, with streaming, may have higher star formation rates (e.g. Kravtsov et al., 2018; Williams et al., 2024a), hold a higher weight in the overall star formation efficiency, leading to comparable or even potentially higher star formation efficiencies as a function of the overall mass of gas in halos. These results are also sensitive to feedback, which is not included here but preferentially suppresses star formation in low-mass halos, as well as significantly lowering the overall star formation efficiency. This feedback is likely to significantly affect the minimum mass cutoff in Figure 5.1 by $z = 12$ and later redshifts, but in analytic modelling is still important at higher redshifts such as $z = 20$ (Hegde & Furlanetto, 2023).

5.3.2 Star formation rate

One of the key results from recent JWST observations (e.g., Harikane et al., 2023; Pérez-González et al., 2023; Willott et al., 2023; Finkelstein et al., 2023; Donnan et al., 2024) is the presence of unexpectedly high star formation rates in the early Universe. The differing slopes of the star formation efficiencies with and without streaming in Figure 5.1 hint that streaming may increase the star formation rate at intermediate redshifts, but the bias towards high-mass objects with streaming in this plot serves as a confounding factor.

In order to better understand this effect, we turn to population-level star formation rate density statistics in Figure 5.2. This figure bins gas objects by their stellar mass (into factor-of-ten solar mass bins), corrects for our enhanced σ_8 in our simulations (for comparisons to an average-density volume of the Universe; see Appendix E.1 for more details), and plots their overall comoving star formation rate density across cosmic time. The left panel of this figure shows summed SFR densities in the mass bin, while the right panel shows averaged (per-object) SFR densities. At later redshifts (after $z \sim 20$), the population of objects in or above the 10^6 $\mathrm{M}_\odot$ bin have comparable star formation rates with and without streaming. In fact, objects in the 10^7 $\mathrm{M}_\odot$ bin actually have slightly higher star formation rates with streaming at some redshifts, a fact that may derive from enhanced gas accretion (Williams et al., 2024a). Furthermore, as seen in the right panel, objects in all mass bins exhibited higher star formation rates individually with streaming than without at lower redshifts, starting at $z \sim 20$. This may, again, partially result from enhanced gas accretion. In addition, because star formation is delayed with streaming, objects in a given stellar mass bin here tend to be associated with higher-mass halos with streaming than without it: the higher halo masses

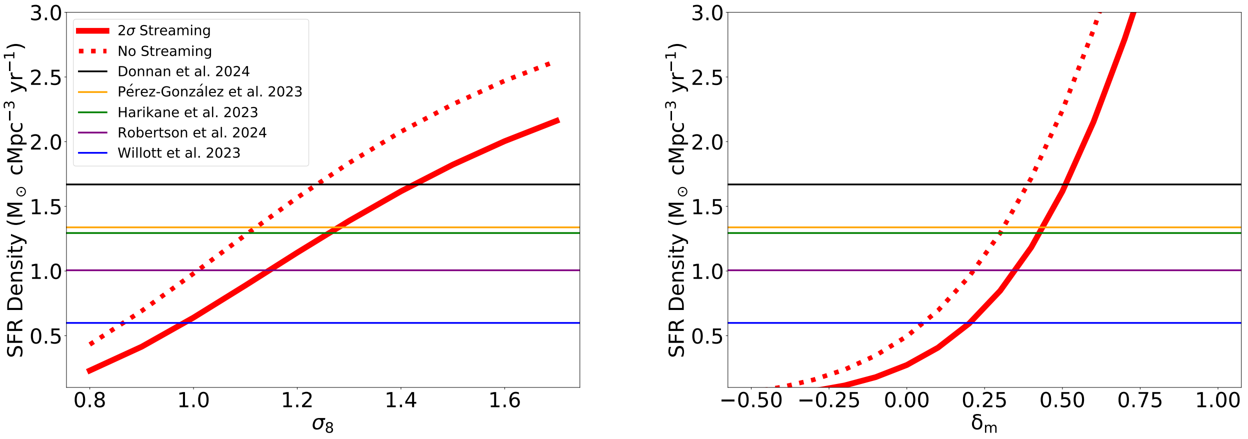


Figure 5.3: **Difference between our simulations and various observations as an overdensity effect:** In the this figure, we show the dependence of our simulations' star formation rate densities at $z = 12$ on σ_8 and the matter overdensity $\delta_{\rm m}$, applying the corrections for varying σ_8 and $\delta_{\rm m}$ described in Appendix E.1. The left panel assumes $\delta_{\rm m} = 0$ and the right panel assumes $\sigma_8 = 0.826$. This shows the effect of observing more- or less-dense regions on the predicted star formation rate. Equivalently, this shows a model for the properties of a region of the Universe expected to produce a given observation in these simulations, with and without streaming. To avoid clutter, we've omitted uncertainties, as at present they are large compared to the size of the y axis. Note that this is linearly scaled, while Figure 5.2 shows the same quantity but is log-scaled.

with streaming could lead to higher star formation rates for a given stellar mass.

Despite this, the suppression of low-mass halo star formation efficiency is still significant, even at $z = 12$ (though, again, at lower redshifts, the mass at which this is significant shifts to lower values). Under the influence of a $2\sigma_{\rm vbc}$ streaming velocity, star formation efficiencies are suppressed by a factor of ~ 2 or more at halo masses below $\sim 10^{7.5}$ M$_\odot$ at $z = 12$, as seen in the right panel of Figure 5.1. At higher redshifts, as seen in Figure 5.2, this gap only grows. This may act in concert with the reduced abundances of low-mass halos (Tseliakhovich & Hirata, 2010; Williams et al., 2024a) to even further suppress the abundance of low-luminosity dwarf galaxies at these redshifts. As the star formation density is constrained to higher redshifts, it is more and more strongly affected by streaming: future JWST and next generation observations may start to constrain this quantity tightly enough at high enough redshift for this to be a good probe of the effects of streaming.

In the left panel of Figure 5.2, we also over-plot several current JWST observational constraints to compare the results of our simulations to the literature. Our results, which are normalized to $\sigma_8 = 0.826$, are consistent with Willott et al. (2023) as shown, but under-predict star formation relative to the majority of observational constraints. This may be attributed to reduced molecular hydrogen cooling in our simulation, a resolution effect discussed later, and the lack of metal line cooling.

However, as we highlight in Figure 5.3, another possibility for this inconsistency is that the observed SFRs are, in fact, coming from an overdense regime (see for an example discussion regarding this point Willott et al., 2023). Therefore, in Figure 5.3, we show the full dependence of our $z = 12$ SFR density in each simulation (solid which includes streaming, and dashed without streaming), as a function of the matter overdensity, normalized using σ_8. The matter overdensity δ_m is defined in terms of the mean density of matter in the Universe $\rho_{\rm m,mean}$ as

$$\delta_m = \frac{\rho_m}{\rho_{\rm m,mean}} - 1. \tag{5.1}$$

The σ_8 and $\delta_{\rm m}$ correction is specified in Appendix E.1. We also over plot the star formation rate density observations from Figure 5.2 at $z \sim 12$, allowing us to convert their observed star formation rates to an equivalent σ_8 or $\delta_{\rm m}$.

As one can see, the results of Willott et al. (2023) are consistent with regions of the Universe with properties relatively similar to the overall Universe in our simulation without streaming (with an effective σ_8 value of 0.86). However, if these observations are derived from a $2\sigma_{\rm vbc}$ patch of the Universe, they are consistent with an overdense regime. Furthermore, at face value, we infer that the Harikane et al. (2023) and Donnan et al. (2024) observations are consistent with $\sigma_8 \approx 1.1$ (1.26) and 1.23 (1.42) without (and with $2\sigma_{\rm vbc}$) streaming. Pérez-González et al. (2023) and Robertson et al. (2023) have similarly overdense inferred fields with $\sigma_8 \approx 1.11$ (1.27) and 1.01 (1.14) without (and with $2\sigma_{\rm vbc}$) streaming.

Naturally, other factors, such as metal line cooling or observational corrections, may contribute to the uncertainty of this analysis. Regardless, we suggest a strategy that can either extract the underlining density of the star-forming region or, in case the latter is found via other means, can constrain the value of streaming. Lastly, Figure 5.3 suggests that the aforementioned observations are consistent with overdense regions in the Universe, apart from Willott et al. (2023) which is consistent with a near-mean-density region without streaming.

5.3.3 Kennicutt-Schmidt relation

A priori, the aforementioned result of enhanced star formation for a given object seems counterintuitive, given the injection of turbulence associated with streaming. However, the DM halo provides a deep potential well for the gas in the DM GHOSts, which overcomes the gas pressure. This effect is then reflected in the dependency of the SFR with the gas surface density, otherwise known as the Kennicutt-Schmidt relation (KSR, Schmidt, 1959). In Figure 5.5, we compare the KSR within DM GHOSts in our $2\sigma_{\rm vbc}$ streaming run (orange crosses) to that in our run without streaming (blue dots) at $z = 12$. Gas surface densities are taken within an ellipsoid described by the smallest ellipsoid capable of containing all of a given gas object's star particles. In this proof of concept, the SFR surface densities considered are those of star particles formed between $z = 13$ and $z = 12$. Halos whose star-forming regions convert more than 25% of their gas mass to stars in this timespan are omitted, as we do not have the time resolution to adequately trace their star formation as a function of a defined, constant gas density.

As one would expect, we see a trend of increasing SFR density with gas surface density with and without streaming. To differentiate between the SFRs of the two cases at given gas densities, we use linear regression to subtract the rising trend in the data. As mentioned, one effect of streaming on these objects is to prevent low-mass/low-density objects from forming stars at these redshifts: this creates a bias towards low-density behavior in our run without streaming, which could be different from intermediate-density behavior. Furthermore, our

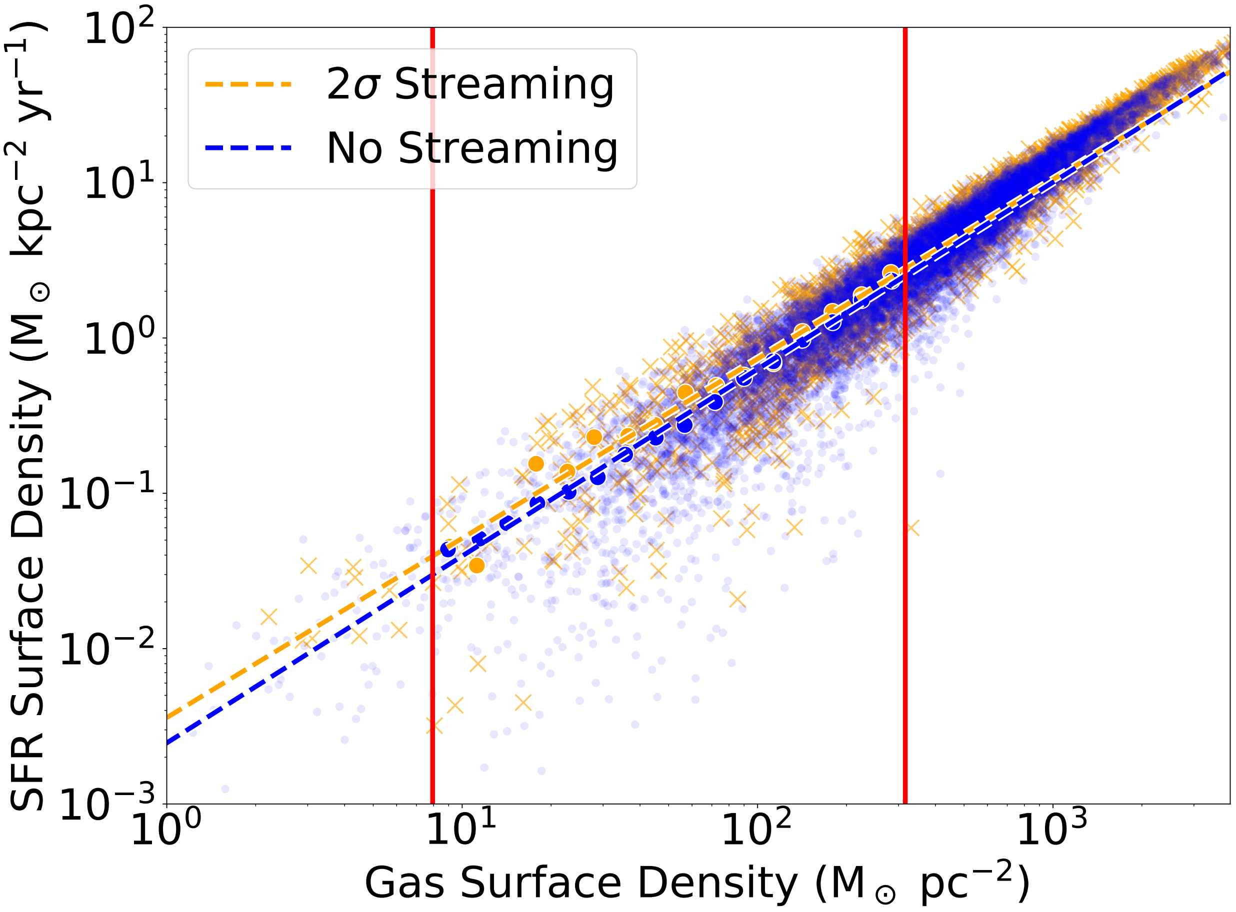


Figure 5.4

Figure 5.5: **Kennicutt-Schmidt relation (KSR) in DM GHOSts:** In the this figure, we show the relation between the star formation rate surface density and gas surface density in halos with (orange) and without (blue) the influence of baryon-dark matter streaming, at $z = 12$. Individual halos' star-forming regions are plotted as individual points, and trend lines are linearly fit and over-plotted with dashed lines of matching colors. Low-gas density halos are suppressed by streaming, necessitating a low-density cutoff to minimize sampling bias. Furthermore, the highest-density halos are impacted by our time resolution, with a KSR slope that approaches 1 as the minimum density considered is raised. However, when comparing only moderate-density halos (region between the red lines), the KSR has consistent slopes with and without streaming. The range of our fit is shown as a red-outlined central region.

limited time resolution and lack of feedback cause high-density objects to tend towards a slope of 1, representing the full conversion of gas to stars. In an effort to lessen the impact of missing feedback, we address these issues by introducing high- and low-density cutoffs in the data (see Appendix E.2 for more details), and bin the remaining data into bins of 0.1 dex width in gas surface density space. Gas surface density values and SFR surface density values are calculated from the mean of the objects in each bin, and SFR surface density errors are taken as the error of this mean. To account for uncertainties derived from the binning process, the gas surface density errors are taken to be half the width of the bin or 0.05 dex. Using linear regression, we then find a best-fit slope α in each case that fits the relation

$$\log\left(\Sigma_{\mathrm{SFR}}\right) = \alpha \times \log\left(\Sigma_{\mathrm{gas}}\right) + C \ . \tag{5.2}$$

In Figure 5.5, we examine the moderate-density star-forming gas objects left after the cuts and show our fit for these objects. Our high- and low-density cuts are shown as red lines. We overlay lines of best fit for the KSR in our $2\sigma_{\mathrm{vbc}}$ ($0\sigma_{\mathrm{vbc}}$) streaming case in blue (orange). These represent $\alpha = 1.20 \pm 0.02$ for the case without streaming and $\alpha = 1.16 \pm 0.05$ for the case with $2\sigma_{\mathrm{vbc}}$ streaming: the difference between these two slopes is merely the standard error. In fact, the only noticeable difference is that the moderate-density star formation regions with streaming may have somewhat higher star formation rates for a given gas density. With the slope estimates subtracted and marginalizing over the independent errors in both slopes and assuming that the residuals are normally distributed, we compute a probability of $p = 0.0012$ that the difference in the means of the residuals with and without streaming is due to random chance. Thus, there is sufficient evidence that the star formation rate density in gas regions of the same density is higher with streaming than without it. While surprising, this can perhaps be explained using the results of Williams et al. (2024a): gas which would have been accreted onto low-mass halos at high redshift was instead advected out in the run

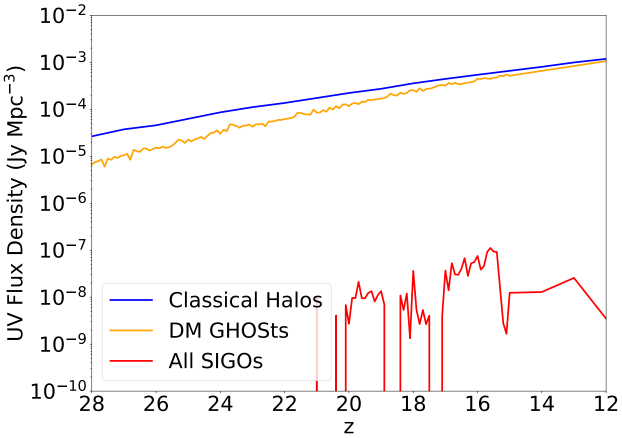


Figure 5.6: **Contributions of object classes to the UV luminosity function** Here, we show the UV flux density of classical halos (halos without streaming, blue), DM GHOSts (halos with a $2\sigma_{\mathrm{vbc}}$ streaming effect, orange), and SIGOs (red, with $2\sigma_{\mathrm{vbc}}$ streaming). This relies on new stars–over time. Although the DM GHOSts form stars more slowly at high redshifts, the influence of streaming fades and they begin to form stars as rapidly as classical halos. Meanwhile, SIGOs form stars several orders of magnitude slower, owing to their relatively low masses, and are a small contribution to the overall UV continuum at high redshift.

with streaming. This gas, which is still generally localized in high-density regions, enhances accretion flows onto halos at lower redshifts, potentially resulting in enhanced star formation in higher-density regions around $z = 12$ under the influence of streaming compared to the no streaming case. This manifests itself in stronger star formation rates for a given gas surface density with streaming in this plot.

Most of the flux from these early objects takes the form of UV emission from young, massive stars (e.g., Sun & Furlanetto, 2016; Hegde & Furlanetto, 2023; Senchyna et al., 2023). We can estimate this flux semi-analytically using the method of Sun & Furlanetto (2016):

$$\dot{M}_{\mathrm{SFR}} = \mathcal{K}_{\mathrm{UV,1500}} \times L_{\mathrm{UV,1500}} \ , \tag{5.3}$$

where we take $\dot{M}_{\mathrm{SFR}}$ to be the average star formation rate in a given class of objects within a Δz $= 0.1$ period preceding the redshift examined (effectively showing the averaged-out flux from a given class of objects on large scales). $L_{\mathrm{UV},1500}$ is the rest-frame UV luminosity at 1500 Å. $\mathcal{K}_{\mathrm{UV},1500}$ here is a fiducial constant, which following Sun & Furlanetto (2016) we set to be approximately $\mathcal{K}_{\mathrm{UV},1500} = 1.15 \times 10^{-28}$ $\mathrm{M_{\odot} yr^{-1}/ergs\ s^{-1} Hz^{-1}}$, which assumes a Saltpeter IMF. At these early redshifts, this may well underestimate luminosities, as the metallicities of these early objects are likely quite low and could permit a top-heavy mass function.

Figure 5.6 shows the total flux density from classical halos and DM GHOSts as a function of redshift, in blue and orange respectively, compared to the flux from SIGOs in red (the SIGOs are discussed in Section 5.4). These UV fluxes are calculated with Equation 5.3. SIGOs are only included as long as they have not merged with halos – several SIGOs fall into halos over the course of this simulation (analogous to accreted clusters), and once this happens they are included as part of the DM+G object flux for the purposes of this figure.

Under the influence of streaming, UV fluxes from DM GHOSts at very high redshifts ($z > 20$) are nearly an order of magnitude lower than UV fluxes from classical halos. However, as with star formation efficiencies, this gap closes over time, and by $z = 12$, nearing the era of reionization, this gap has nearly closed completely. This is an even stronger statement of the fading effect of streaming than the narrowing gap in star formation efficiency seen above: here, this result does not depend on the gas mass in halos. Since as seen in Figure 5.3 the star formation rate in these objects is still a sensitive probe of streaming at $z = 12$, this implies that future, higher-redshift observations will be even more sensitive.

5.4 Star Formation in SIGOs

As has been recently shown through both simulations and theory (Lake et al., 2023a,b; Nakazato et al., 2022), SIGOs are capable of forming stars, even outside of dark matter halos. These star clusters may be accreted by forming protogalaxies, becoming bound clusters analogous to globular clusters (Lake et al., 2023b). In this section, we analyze population-

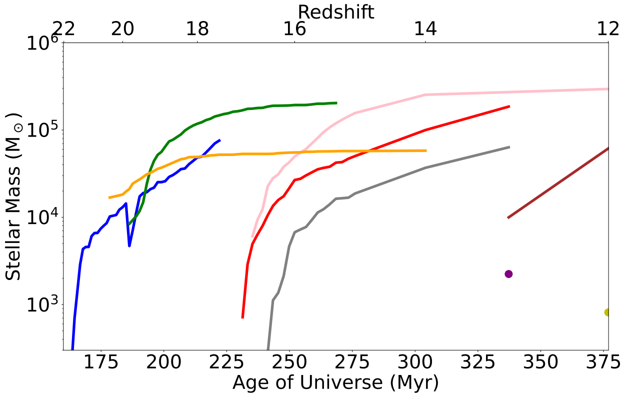


Figure 5.7: **The star formation history of SIGOs:** This plot shows the star formation history of all visually confirmed SIGOs which form stars in our $2\sigma_{\rm vbc}$ simulation. The first of these SIGOs to form stars does so before $z = 21$, and the final SIGO to begin star formation does so at the final timestep of our simulation, $z = 12$. SIGOs are removed from this plot when they merge with nearby dark matter halos. Two of these SIGOs (colored in blue and green) fall all the way to the center of the nearest halos, becoming structures akin to nuclear star clusters.

level properties of these clusters, aiming to estimate their mass range and abundances.

In Figure 5.7, we show the star formation history of each star-forming SIGO in our simulation. Each SIGO is represented by a different-colored line, and (two) SIGOs that are only found with stars at a single snapshot are represented with points. Each of these SIGOs is visually verified by visualizing their gas and dark matter density evolutions, verifying that the gas overdensities that give rise to the SIGOs originate outside of DM halos.

9 SIGOs in this simulation form stars by $z = 12$, an abundance of 0.58 cMpc^{-3}. Keeping in mind that these simulations do not include feedback and thus tend to convert all gas to stars on long timescales in star-forming objects, the characteristic masses of stars formed quickly (i.e. on $\sim$ 10 Myr timescales) in these objects fall within the range $M_* \sim 10^3 - 10^5$ M$_\odot$. These SIGOs are diverse in their evolutionary outcomes: for example, of the 9 star-forming SIGOs in our simulation, 2 fall into the centers of dark matter halos and appear to form objects analogous to nuclear star clusters, while 7 remain visually distinguishable baryon objects by the end of our simulation, $z = 12$.

As many of these SIGOs will disperse through processes such as two-body relaxation before the present day (e.g., Naoz & Narayan, 2014; Lake et al., 2023a), observational campaigns for SIGOs must focus on high redshifts, where these objects are abundant, albeit faint. As shown by Lake et al. (2023b), massive SIGOs may be detectable by current instruments, so it is important to constrain the properties of these objects at high redshift, to aid in distinguishing them from other classes of object and contextualizing observations.

As seen in Figure 5.6, the UV flux density from SIGOs in the simulation is several orders of magnitude lower than the UV flux from DM+G objects, owing to their lower abundance, and the significantly higher maximum mass of DM+G objects, compared to the maximum stellar mass of SIGOs of about 10^5 $M_\odot$. As noted by Lake et al. (2023b), SIGOs are not likely to be detectable by current generation instruments at these simulation redshifts ($z \geq 12$) owing to these low luminosities. However, similar SIGOs forming later, just preceding Reionization, may be detectable (Lake et al., 2023b), making it important to understand the process of star formation within SIGOs.

As SIGOs form outside of dark matter halos and follow a relatively unique evolutionary pathway to form stars as a result, it is interesting to compare the efficiency of their star formation to that of star formation in DM GHOSTs (gas components of halos in regions affected by significant streaming; in this case, halos in our $2\sigma_{\rm vbc}$ streaming velocity simulation) and classical halos (halos without streaming). To this end, we can revisit Figure 5.1, which shows the efficiency of star formation. The gap in the SIGO data in this figure between $z = 20$ and $z = 19$ corresponds to redshifts where no star-forming SIGOs are found outside of halos (prior star-forming SIGOs have merged into halos).

Owing to small number statistics, SIGOs have the most variability in their star formation efficiencies as a function of redshift, but generally have the smallest star formation efficiencies of the three object classes. Potentially owing to their lack of dark matter and reliance on molecular hydrogen cooling, SIGOs tend to form stars more slowly and less efficiently than classical halos (e.g. Lake et al., 2023a). In addition, in this simulation that does not include radiative feedback, older objects tend to exhibit artificially high star formation efficiencies. Because older SIGOs tend to merge with nearby halos and drop out of this figure (Lake et al., 2023a), while older DM+G objects tend to survive and keep forming stars, the calculated efficiencies of star formation in classical halos and DM GHOSts may perhaps be biased to be relatively higher than in our observed Universe compared to SIGOs. This effect will be particularly pronounced for low-mass DM GHOSts.

5.5 Conclusions

The streaming velocity has a variety of significant effects on star formation at high redshifts, from suppressing star formation in low-mass halos to allowing the formation of small-scale gas structures outside of halos and even potentially enhancing star formation at certain redshifts in high-mass halos (e.g., Naoz & Narayan, 2014; Fialkov & Barkana, 2014; Chiou et al., 2019; Schauer et al., 2021; Lake et al., 2023b; Hegde & Furlanetto, 2023; Williams et al., 2024a). Notably, it has been suggested that our own Local Group may have formed in a region with a particularly high streaming velocity, increasing the significance of streaming's effects for our understanding of our own galaxy (Uysal & Hartwig, 2022). In this work, we focused on two components of these effects: the population-level contribution of supersonically induced gas objects (SIGOs) to early star formation and the impacts of streaming on star formation rates in early halos.

In order to contextualize the role of the streaming velocity in early star formation, we present the first population-level study of star formation in SIGOs and DM GHOSts in a simulation with explicit star particles. We estimate the stellar masses of these objects and study how they change over time. We systematically compared them to classical objects (i.e., no stream velocity). For example, as seen in Figures 5.1 and 5.2, significant differences in star formation efficiencies and rates persist between low-mass halos to as late as $z = 12$ in these simulations with and without streaming. Streaming tends to advect gas out of low-mass halos, an effect which persists even in $\lesssim 10^7$ M$_\odot$ halos at $z = 12$. Regions of the Universe with relatively high values of the stream velocity, such as our Local Group, may then have critical masses for star formation from the molecular hydrogen cooling threshold that rival those from Lyman-Werner radiation even at these relatively late redshifts (e.g., Hegde & Furlanetto, 2023). This is clearly visible in the right panel of Figure 5.1, where the mass at which the star formation efficiency in DM GHOSts drops off is systematically much higher than in classical halos.

At $z \sim 12-20$, accretion from the surrounding gas is enhanced with streaming for high-mass halos, resulting in slightly higher overall SFR densities in halos with streaming than without it after $z = 20$. This effect takes place even for lower-mass objects which are classically treated as being inhibited by streaming. The SFR per stellar object is higher in the case of streaming in all mass bins at $z = 12$, as depicted in the right panel of Figure 5.2. In Figure 5.5, we also see that the SFR surface density (i.e., the Kennicutt-Schmidt relation) is enhanced in moderate-gas-density (roughly, lower mass) halos with streaming compared to without it at $z = 12$.

These effects are studied in the case of a $2\sigma_{\rm vbc}$ streaming velocity; however, 1σ streaming velocities are more common in the Universe, and while rarer, 3σ streaming velocities are present and present a more extreme case. To infer the differences between these cases and the present case, and with an eye on the accretion-driven origin of this SFR enhancement, we must understand when accretion is enhanced for different values of the streaming velocity. This has previously been studied in the form of the gas fraction in halos (Naoz et al., 2013; Hirano et al., 2023), which catches up to the no-streaming case at different redshifts depending on the value of the streaming velocity in the region. In the more common lower-streaming cases, the suppression of the gas fraction in halos is weaker and ends at higher redshifts (e.g. $z \sim 15$). Therefore, we expect weaker impacts on star formation rates in this case that are more pronounced at higher redshifts. On the flipside, in higher-streaming regions, star formation rates in individual halos could be even more enhanced at $z \sim 12$, with effects that persist into the epoch of reionization.

It is important to note that our simulation is resolution-limited and does not include feedback or metal line cooling, which is likely to cause a general lowering and flattening with redshift of the SFR density relations we show, as feedback is more important at lower redshifts. These competing effects likely lead to an underestimation of the efficiency of molecular hydrogen cooling, as well as that of cooling more generally (Nakazato et al., 2022). The lack of feedback also likely leads to an overestimation of the star formation efficiency in these objects, especially at later times and at low masses. This is particularly pronounced in the minimum mass at z=12 and in the low-mass star formation rates at z=12. Thus, we are most sensitive to star formation in halos in these simulations in the DM mass range of a few×10^6 $\mathrm{M}_\odot$ or higher, where our results are less sensitive to molecular hydrogen cooling rates (instead, they depend on the filtering mass, e.g., Hegde & Furlanetto, 2023).

As shown in Figure 5.2, some current observational constraints from JWST may already be compatible with our simulations without streaming, given their margins of error (Willott et al., 2023). However, our SFRs may be considered upper limits due to the lack of feedback, thus exacerbating the discrepancy between simulated and observed star formation rates. In principle, as suggested by Figure 3, comparisons between observed and simulated SFRs can permit inferences on the environments of observed systems, such as their large-scale over- or under- density and their streaming conditions.

We note that the SFR shown in Figure 5.2 is normalized to the Planck 2013 $\sigma_8 = 0.826$ (Planck Collaboration et al., 2014), chosen for consistency with previous studies of DM GHOSts. However, a smaller σ_8 seems to have similar effects to a decreased large-scale overdensity. We analyze the effect of the dense environment (i.e., the “local” σ_8 or δ_m) in Figure 5.3, where the SFR density at $z = 12$ for varying large-scale overdensities (see Appendix E.1), are shown with (solid) and without (dashed) streaming. This Figure highlights the potential of comparing observations with simulations. Specifically, constraints on the over- or under-density of the observational volume and its star formation rate may serve as an indication of the stream velocity. For example, Willott et al. (2023) speculated that their observations may take place in an underdense region. This observation is consistent with our no-streaming simulations. On the other hand, high SFRs such as the one reported by Donnan et al. (2024) may require a $\sigma_8 = 1.42$ ($\sigma_8 = 1.23$) with (without) streaming. Once the relation between today's observations and simulations is better understood, a similar strategy can be used to extract information about the density and streaming value of future survey fields.

Finally, the streaming velocity produces SIGOs. Figure 5.7 shows the star formation history of all SIGO-derived star clusters in the simulation. Out of 5325 unique SIGOs found in the simulation at all redshifts, only 9 confirmed SIGOs formed stars outside of halos, showing the rarity of this process as a function of the overall abundance of SIGOs (which is not surprising for low-mass objects outside of DM halos). However, as SIGOs are common objects before reionization, this still represents an abundance of SIGO-derived star clusters of 0.58 cMpc^{-3}, which is higher than the local abundance of low-metallicity globular clusters (0.46 cMpc^{-3}, Rodriguez et al., 2015). Typical stellar masses for SIGOs in this simulation within 10 Myr after the beginning of star formation[1] fall on the order of $10^3 - 10^5$ $\mathrm{M}_\odot$; however, it is important to remember that this is likely sensitive to the efficiency of molecular hydrogen cooling, which as noted above is underestimated in our simulation (Lake et al., 2023a). It is, therefore, possible that this is an underestimate of the final stellar mass of SIGOs.

Given that SIGOs form stars, it is interesting to ask whether the total stellar flux from SIGOs is significant compared to the flux from classical objects at high redshift. The simple

---

[1] Feedback will dominate after this period and so we think it is important to emphasize the results of this initial burst of star formation, which has a more limited star formation efficiency.

answer is that, owing to their small stellar masses, UV emission from SIGOs is relatively small (Figure 5.6). However, we suggest that as SIGOs sit within the gas streams which feed early halos, feedback from these SIGOs may have important effects on gas accretion onto these early halos, potentially lowering the star formation efficiency in halos when feedback is considered in conjunction with streaming.

Also visible in Figure 5.6 is a comparison of UV flux from more classical halos with (orange) and without (blue) streaming. An important feature of streaming can be observed here: while there is a near order-of-magnitude dropoff in UV emission with streaming at $z = 30$, the effect of streaming falls off with redshift (as the streaming velocity decays), such that the UV emission with and without streaming is similar by $z = 12$, as star formation rates have caught up within halos in the streaming velocity simulation. Similarly, for older stars, Figure 5.1 shows that the overall star formation efficiencies in classical halos and DM GHOSts have begun to converge by $z = 12$. This also results in part (right panel of Figure 5.1) from the increasing contribution of high-mass halo star formation, which is minimally affected by streaming, to the overall rate of star formation in the Universe at lower redshifts.

As this paper shows, observations have, for the first time, begun to directly probe and place constraints on the era where the streaming velocity is critical to structure formation. As the JWST dataset grows and observational constraints at high redshifts tighten, the potential for statistical comparisons at these high redshifts to isolate the effect of streaming grows with it. Because the streaming velocity is a prediction of ΛCDM not present in many alternative models, this presents an exciting opportunity for coming high-redshift observations to test ΛCDM structure formation.

CHAPTER 6

The Stellar Initial Mass Function of Early Dark Matter-free Gas Objects

6.1 Introduction

The cosmic microwave background demonstrated that the Universe at early times, known as the era of Recombination, was remarkably uniform, apart from small overdensities to the tune of one part in 100,000 (Bennett et al., 1996). Well before Recombination, dark matter fell towards these primordial inhomogeneities, allowing them to grow over time to form the seeds of the galaxies we see today. However, until Recombination, the gas was tightly coupled to radiation, which suppressed the growth of gas overdensities (Naoz & Barkana, 2005a). Because the motion of baryons was suppressed by the radiation field, this gave rise to baryon-dark matter relative motions, known as baryon-dark matter streaming or as the streaming velocity (Tseliakhovich & Hirata, 2010; Tseliakhovich et al., 2011). These motions are highly supersonic (~ 30 km s^{-1} at $z = 1100$, or about 5 times the speed of sound) and generate spatial offsets between gas clumps and the early dark matter overdensities that give rise to them, depleting the newborn clumps of dark matter (Naoz et al., 2012; Naoz & Narayan, 2014). The gas clumps, devoid of dark matter (like globular clusters, or GCs), form

---

The contents of this chapter appeared in Lake et al. (2024a) and have been submitted to Nature baryon-enriched structures with masses below 10^7 $M_\odot$ known as supersonically induced gas objects (SIGOs), which have been shown to form star clusters through simulations and theory (Popa et al., 2016; Chiou et al., 2018, 2019, 2021; Lake et al., 2021; Williams et al., 2023; Lake et al., 2023a; Nakazato et al., 2022; Lake et al., 2024b). These objects form in similar abundance to low-metallicity globular clusters in the early Universe (Lake et al., 2021), and because the Milky Way may have formed in a region of the Universe with an elevated value of the streaming velocity (Uysal & Hartwig, 2022), the local Universe is expected to have formed a higher abundance of SIGOs.

A model of such an object is shown in the top panel of Figure 6.1. Specifically, we show a SIGO extracted from a cosmological simulation using `AREPO` (Springel, 2010) at $z = 15$. The top left panel shows a self-gravitating gas-dominated structure at kpc scales, while the right shows the lack of a DM component in the same object[1]. This depiction of a SIGO shows it just before star formation begins at its core.

Studies of these SIGOs in Milky Way-scale cosmological simulations have determined their abundances, some observational signatures, and some population-level properties (Popa et al., 2016; Chiou et al., 2018, 2019, 2021; Lake et al., 2021, 2023a,b; Lake et al., 2024b; Williams et al., 2024b). However, all these simulations studied SIGOs at the inherently limited resolution required for Mpc-scale simulations. This restricted their ability to probe the interior of the SIGOs, as the resolution of these runs ($\sim 10^2$ $M_\odot$) is not sufficient to resolve the objects' cores. This level of resolution also artificially limited the effect of molecular hydrogen chemistry in the cooling and subsequent gravitational collapse of SIGOs, as well as under-resolving the gravity and hydrodynamics. One way to avoid these numerical issues is by conducting zoom-in simulations (Nakazato et al., 2022). Yet, explicitly modelling star formation that follows the gravitational collapse of SIGOs to form stars remains a challenge even using the improved resolution of the aforementioned numerical methodology. Moreover,

---

[1] See Methods 6.2.1 for simulation details

Name	Metallicity ($Z_\odot$)	Jets?	$R_{\rm eff}$ (pc)	Σ ($M_\odot$ pc^{-2})
M6NJ	10^{-6}	No	0.025	1.5×10^5
M6J	10^{-6}	Yes	0.19	1.0×10^3
M4NJ	10^{-4}	No	0.05	3.7×10^4
M4J	10^{-4}	Yes	0.03	4.9×10^4

Table 6.1: Description of simulation runs performed. Runs have differing metallicities and either contain or neglect jet feedback. Names for different runs are given. Effective radii are calculated as the half-light radius 100 kyr after the formation of the first star, and stellar surface densities are calculated using the same effective radii.

none of these studies included feedback, and none probed the stellar properties of SIGOs (i.e., the Initial Mass Function, IMF). This limited the interpretation of the results, leaving open key questions such as the star formation efficiency of SIGOs and their ultimate luminosities and observability by JWST and future instruments.

Here, guided by the properties of the star-forming SIGOs in our lower-resolution cosmological simulations (Lake et al., 2023b), we investigate the star-formation process at substantially higher resolutions. We utilize small-scale, very-high-resolution ($\sim 10^{-2}$ $M_\odot$) `STARFORGE` simulations (Grudić et al., 2021a) (see Methods 6.2.2 for details) performed with the `GIZMO` code. These simulations resolve the evolution of a single SIGO analogue through its cooling, collapse, and resolved star formation, including jet feedback [2]. This allows us to determine the properties of the SIGO, i.e., the first dark matter-free star cluster. Specifically, we ascertain its IMF and stellar mass, and estimate its star formation efficiency. In Figure 6.1 we illustrate the resolution achieved in this type of simulation. Specifically, in the lower panels, we show an example of one of our `STARFORGE` simulations; for illustration purposes, here we display the star-forming region of the SIGO 13 kyr after the formation of the first star, from our simulation with 10^{-6} $Z_\odot$ metallicity without jet feedback (run M6NJ, see also Table 6.1).

---

[2] The contribution of line-driven stellar winds is negligible at the present metallicities, and the timescales covered by our simulation are shorter than the explosion times of the first supernovae.

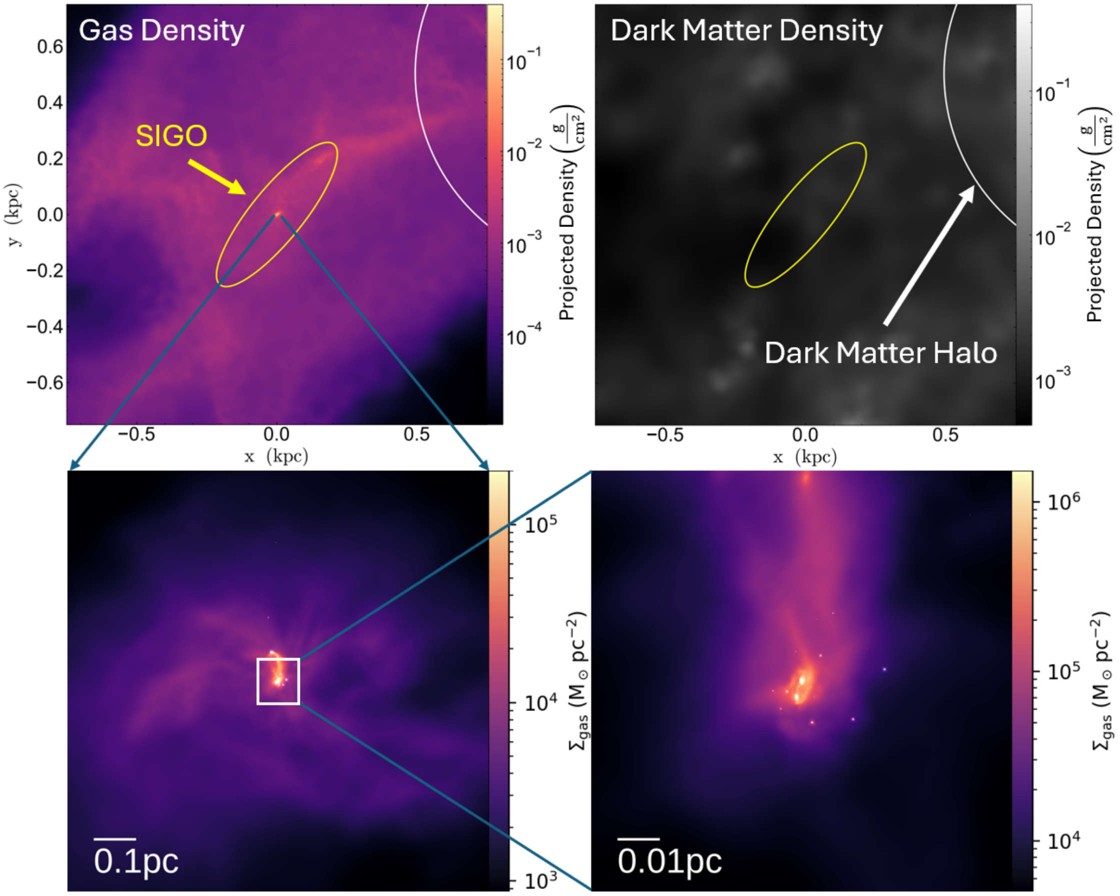


Figure 6.1: A star-forming SIGO (yellow) in `AREPO` and `STARFORGE` simulations. In the top line (`AREPO` simulations), it is near a DM halo (white). The left top panel shows the local gas density, while the right top panel shows the local DM density (and lack thereof, in the SIGO). The bottom panels show a similar SIGO (i.e., not a zoom simulation from the top panel) in a metal-poor `STARFORGE` simulation, with much higher resolutions and on much smaller scales.

6.2 Methods

In this paper, we follow the evolution and star formation of an isolated, ultra-low-metallicity gas cloud with the `STARFORGE` numerical simulation framework (Grudić et al., 2021a). This allows us to simulate the physical processes of star formation across a wide dynamical range. Here, we use it to model several aspects of stellar feedback, including protostellar outflows, stellar winds, and supernovae. We run these simulations using the `GIZMO` hydrodynamic code (Hopkins, 2015), with an initial condition based on a set of low-resolution `AREPO` simulations.

As mentioned, SIGOs have consistently low dark matter densities and are not associated with extended dark matter halos. Because of this, our `STARFORGE` simulations do not include dark matter. The lack of dark matter lowers the virial temperature of the object and thus its average temperature compared to classical Pop III star formation in minihalos (Schauer et al., 2021; Kulkarni et al., 2021; Hegde & Furlanetto, 2023; Lenoble et al., 2024). We take the temperature of our present SIGO to be 225 K, representative of the example SIGO from the `AREPO` simulation (see Section 6.2.3 for a discussion of the simulation setup as well).

The early star formation process is sensitive to the gas composition. Metals enable more efficient cooling, facilitating gravitational collapse and star formation. SIGOs do not evolve in isolation, and so we must consider whether nearby supernovae can impact their gas metallicities. More particularly, large protogalaxies can host SIGOs (Lake et al., 2023b), and thus, star formation in host galaxies may contribute small amounts of metals to the SIGO. This may "pollute" the SIGO with metals, allowing it to form stars from gas that is not truly pristine, suggested by (Schauer et al., 2021). Therefore, here we consider 2 scenarios for external metal pollution: one with low levels of metals ($\mathrm{Z_{SIGO}} = 10^{-6}\ \mathrm{Z_{\odot}}$) in which metallicity effects are minimal (Omukai et al., 2010), and one with relatively high levels of metal enrichment that take the SIGO to the boundaries of Pop III to Pop II star formation ($\mathrm{Z_{SIGO}} = 10^{-4}\ \mathrm{Z}\odot$). Additionally, we consider two extreme scenarios for stellar feedback. Specifically, present-day star formation is observed to produce protostellar jets, which carry about 30% of the mass accreted in the protostellar phase (Shu et al., 1988; Pelletier & Pudritz, 1992; Cunningham et al., 2011). The existence and strength of jets in early, low-metallicity protostars is currently unknown (Machida et al., 2008; Machida & Doi, 2013; Sharda et al., 2020; Sadanari et al., 2021; Prole et al., 2022; Sadanari et al., 2024). Thus, for each run at different metallicities, we perform two runs, where the jets are "on" or "off." We do not include radiative feedback because radiative trapping near the protostars may mitigate the feedback effect (Jaura et al., 2022), see Methods 6.2.2 for more discussion. See Table 6.1 for the full set of simulations carried out in this work.

6.2.1 AREPO

The initial conditions for our simulations are inspired by a set of small-box (2.5 cMpc) `AREPO` simulations (Springel, 2010) performed and analyzed in Lake et al. (2023b) and Lake et al. (2024b). These simulations included 768^3 DM particles with mass $\mathrm{M_{DM}} = 1.1 \times 10^3\ \mathrm{M_\odot}$, and 768^3 Voronoi mesh cells with gas mass $\mathrm{M_B} = 200\ \mathrm{M_\odot}$, and were run from $z = 200$ to $z = 12$. Initial conditions for these simulations used transfer functions from a modified version of `CMBFAST` (Seljak & Zaldarriaga, 1996), which incorporates first-order scale-dependent temperature fluctuations (Naoz & Barkana, 2005a) and the streaming velocity (Tseliakhovich & Hirata, 2010). The streaming velocity was implemented as a uniform boost based on its 2σ value at $z = 200$ of 11.8 km $\mathrm{s^{-1}}$. The simulations also included non-equilibrium molecular hydrogen cooling based on `GRACKLE` (Smith et al., 2017; Chiaki & Wise, 2019). This package includes non-equilibrium molecular hydrogen and deuterium chemistry. However, it does not include metal line cooling, nor does it include external radiation such as Lyman-Werner backgrounds.

6.2.2 STARFORGE

As discussed above, the high-resolution simulations utilize the `STARFORGE` high-resolution star formation framework within the `GIZMO` hydrodynamic simulation code (Hopkins, 2015;

Grudić et al., 2021a). These simulations include metal and molecular hydrogen cooling, molecular hydrogen nonequilibrium chemistry, and external heating from an interstellar radiation field (ISRF). The cooling and chemistry is based on FIRE-3 (Hopkins et al., 2023), and includes recombination, thermal bremsstrahlung, metal lines, molecular lines, fine structure, and dust collisional processes.

These `STARFORGE` simulations also explicitly include feedback mechanisms through supernovae and protostellar jets. As we are only simulating these clouds for about 100 kyr after the formation of the first star, no supernovae occur in these simulations. We model the protostellar jets using the mechanism described in the `STARFORGE` methods paper (Grudić et al., 2021a), launching them with a velocity equal to 30% that of the keplerian velocity at the protostellar radius and a mass equal to 30% of the accreted mass by the protostar. Because of our extremely low metallicities, stellar winds are not significant in these clouds and we neglect them (see for a review Klessen & Glover, 2023). Radiative feedback is likely to have an effect on these early, low-metallicity clouds by driving gas outflows and ionization (e.g. Stacy et al., 2012; Guszejnov et al., 2022); however, the exact impact is uncertain due to the unknown extent to which radiation can escape the high-density accretion regions surrounding Pop III stars (see e.g. Jaura et al., 2022; Sharda & Menon, 2024). Thus, the effect of radiative feedback is outside the scope of the present paper and not included.

Star formation is modeled according to the set of prescriptions implemented in `STARFORGE` discussed in the `STARFORGE` methods paper (Grudić et al., 2021a) and reproduced here. To be considered for star formation and become a sink particle, a gas cell must first satisfy a minimum density criterion based on the Jeans scale: in our simulations with a resolution of 0.01 $M_\odot$, this density can be expressed as $\rho_J \approx 4.7 \times 10^{-12}\,(c_s/\mathrm{km\ s^{-1}})^6$ g cm^{-3}, where c_s is the local speed of sound. Secondly, the gas cell must be the densest cell of the cells with overlapping kernel radii (for this purpose, existing sinks have infinite density). Thirdly, the gas cell's density must be increasing with time. Fourthly, the cell must be gravitationally unstable at the resolution scale. Fifth, the tidal tensor at the position of the gas cell must be fully compressive. Finally, the free fall timescale for the gas cell must be shorter than both the free fall and orbital timescales to nearby sink particles. Upon becoming a sink particle, the new star can accrete gas particles based on the conditions in the methods paper(Grudić et al., 2021a).

6.2.3 Initial Conditions

Our simulations are a proof-of-concept motivated by the properties of SIGOs in `AREPO` simulations (Lake et al., 2023b). The simulations begin with a uniform sphere of gas with a mass of 3×10^5 $\mathrm{M}_\odot$ generated using `MakeCloud`[3], representing a SIGO with sufficient mass to form stars, but with a relatively low mass for such an object. This was chosen for computational feasibility, allowing us to explore several cloud setups with different feedback and metallicity prescriptions. The cloud has a radius $R_{\mathrm{cloud}} = 130$ pc and initial temperature of 225 K, consistent with molecular hydrogen cooling in pristine gas and with SIGOs in our AREPO simulations. The cloud is surrounded by warm, diffuse gas to ensure that the SIGO is in thermal pressure equilibrium with its surroundings. The cloud's initial turbulent energy is specified using α_{turb} (Bertoldi & McKee, 1992):

$$\alpha_{\mathrm{turb}} = \frac{5||v_{\mathrm{turb}}||^2 R_{\mathrm{cloud}}}{3\mathrm{GM}}, \tag{6.1}$$

where v_{turb} is the turbulent velocity field, which is initialized using a gaussian random field using a k^{-2} power spectrum. We take $\alpha_{\mathrm{turb}} = 0.5$, which is approximately its median value for star-forming SIGOs in simulations(Lake et al., 2023b).

In order to simulate the relatively pristine, early environment of these SIGOs, we take their metallicity to be either 10^{-6} or 10^{-4} $\mathrm{Z}_\odot$, which allows for small variations in metal pollution from nearby structures, but still represents Pop III star formation. Thus, the primary

---

[3] https://github.com/mikegrudic/MakeCloud

coolant is molecular gas, which we trace explicitly using the methods of Hopkins et al. (2023). We instantiate the cloud with a molecular hydrogen mass fraction of 10^{-3} throughout, corresponding to the median value observed in star-forming SIGOs in simulations (Lake et al., 2023b), and typical of early star-forming molecular clouds (Klessen & Glover, 2023). The ambient radiation field is chosen to mimic a $z = 20$ radiation background for the mean Universe.

6.3 Results

In Figure 6.2, we show the central star-forming region of the evolved gas cloud (SIGO) in the `STARFORGE` runs at three different times: at the formation of the first star $t_{\rm fs}$, 10 kyr after $t_{\rm fs}$, and 100 kyr after $t_{\rm fs}$ (columns from left to right). These visualizations are centered on the most massive star in each simulation. The low-metallicity 10^{-6} $Z_\odot$ run is shown in the top two rows (without jets, top row, and with jets, second row), and the 10^{-4} $Z_\odot$ run is shown on the bottom two rows (without jets, third row, and with jets, bottom row). For reference, we list the half-light radius and stellar mass surface density (using this half-light radius[4]) in Table 6.1. Because the high-mass and therefore high-luminosity stars are centrally concentrated in this initial burst of star formation, the stellar surface densities are very high, orders of magnitude above those of local star formation, and comparable to some of the highest known star cluster densities found in high-redshift stellar clusters observed with JWST (Adamo et al., 2024).

As is apparent from the rightmost column of Figure 6.2, the higher-metallicity run has a tendency towards fragmentation on 10^{-2}-10^{-1} pc scales, regardless of feedback. This feature is more prominent for the higher metallicity run, owing to its enhanced cooling and lower temperatures on these sub-pc scales. However, this fragmentation makes little difference

---

[4] This uses simulated masses, but note that direct comparison to observations is subject to observational uncertainty, especially because these stars would be UV-bright compared to a Kroupa or Salpeter IMF and short-lived due to their low metallicities and high masses.

to the total mass of stars formed in each run. In both runs without jets, the total stellar mass at 10^5 yrs is about ~ 600 M$_\odot$. On the other hand, the metallicity may have an impact on the maximum mass of the stars formed. Specifically, at 10^5 yrs, the highest mass reached at the end of the simulation in run M6NJ is about 70 M$_\odot$, compared to about 50 M$_\odot$ in run M4NJ. The less-fragmented accretion features in the lower-metallicity run permit substantially greater maximum masses of individual stars, even at these very low metallicities (as depicted in Figure 6.3). Assuming a star formation timescale of up to one Myr (after which, even if radiation has not quenched star formation, supernovae will), our results suggest a star formation efficiency in this SIGO of order 1%, with or without jets (though note that we do not account for the effects of radiative feedback). It is important to note that on the scale of star-forming SIGOs, this is a small object, which forms a low-mass and low-density star formation region. Thus, it is likely that this is a lower bound on the star formation efficiency of high-mass SIGOs.

Given that jet feedback plays a minor role in the overall star formation efficiency in these clouds, we can next ask whether it impacts the masses of the stars formed. In Figure 6.3, we show the stellar IMF of the 4 simulations. The low-metallicity runs are shown in the left panel, and the 10^{-4} Z$_\odot$ runs are shown in the right panel. Runs with jets are indicated with solid lines, and runs without jets are indicated with dashed lines. We also report fits to the high-mass end of the IMF of the form $dN/dM \propto M^{\alpha}$. For the low-metallicity runs, we find $\alpha = -0.32 \pm 0.12$ (-0.20 ± 0.05) for the case without (with) jets. For the 10^{-4} Z$_\odot$ runs, we find $\alpha = -0.16 \pm 0.05$ (-0.40 ± 0.05) for the case without (with) jets. Through a K-S test, we find no significant differences in α for the low-metallicity case ($p = 0.35$ that the differences between the two distributions are due to random chance), but find support for significant differences in the high-metallicity case ($p = 0.03$). In the high-metallicity case, we find a significantly steeper slope in the IMF with jets than without, owing to enhanced low-mass star abundances as the jets disrupt accretion flows.

In all cases, we find power-law stellar IMFs that are top-heavy, as is expected in Pop III star formation in other environments. For comparison, a Salpeter IMF is shown on each panel of Figure 6.3, indicating just how much more top-heavy these star clusters are than present-day stellar groups. This will likely result in low mass-to-light ratios in SIGOs: their enhanced luminosities could aid in observation (especially with the aid of gravitational lensing) in instruments such as JWST (Lake et al., 2023b). If present, bipolar outflows do play a role: we find a stellar mass-to-light ratio in run M6NJ of $10^{-4} M_{\odot}/L_{\odot}$, 10× lower than the $10^{-3} M_{\odot}/L_{\odot}$ in run M6J. This result is reflected to a lesser degree in runs M4NJ and M4J, with the mass-to-light ratio in M4J being 30% higher than that in M4NJ. Also note that because the SIGO may form stars at later times, there is uncertainty in the final/true IMF–we can only provide an estimate at 100 kyr.

Pop III star formation has been investigated before, both in zoom-in cosmological simulations (Stacy & Bromm, 2013; Hirano et al., 2014, 2015; Prole et al., 2023) and in small cloud simulations (Wollenberg et al., 2020; Jaura et al., 2022; Prole et al., 2022). The former often have either limited time evolution or mass resolution, while the latter is often envisioned as a small ($\lesssim 3000$ M$_{\odot}$) cloud within a bigger classical minihalo. For the first time, inspired by the DM-free structures seen in our cosmological simulations, we report a top-heavy IMF in a high-redshift, GC-like large cloud.

6.4 Conclusions

Our simulations have demonstrated that early systems without dark matter are able to naturally form stars, including models of thermochemistry and resolving fragmentation and individual stars. Further, we have shown that jet feedback has a minor impact on the low-metallicity high-mass IMF and does not substantially affect the star formation efficiency, although it may slightly steepen the final stellar mass function in SIGOs that are externally metal-enriched. We have also shown that the process of star formation in SIGOs results in a top-heavy stellar mass function, producing bright stars with short lifetimes. This result is robust to external enrichment of the SIGOs, with even substantial levels of metal enrichment producing top-heavy IMFs. SIGOs also form stars with very high stellar mass surface densities (even without invoking methods of reaching higher densities, such as ram pressure during accreting onto a minihalo), comparable to anomalous observations at high redshift (Adamo et al., 2024). These formed clusters may be bright enough to be observed with JWST, if they form up to the Reionization era at comparable star formation efficiencies, especially with top-heavy IMFs (Lake et al., 2023b). High-mass star-forming SIGOs may be able to form larger star clusters than the example SIGO here, resulting in UV-bright objects in JWST fields, which may be hosted within larger protogalaxies. These could manifest similarly to the star clusters already observed at high redshift, for example the Cosmic Gems clusters(Adamo et al., 2024) as well as others(Messa et al., 2024; Vanzella et al., 2024), which exhibit small sizes and low metallicities as would be expected of a DM-deficient star cluster through this process. Clusters before Reionization, to be detected by future observations, may also originate from high-redshift SIGOs

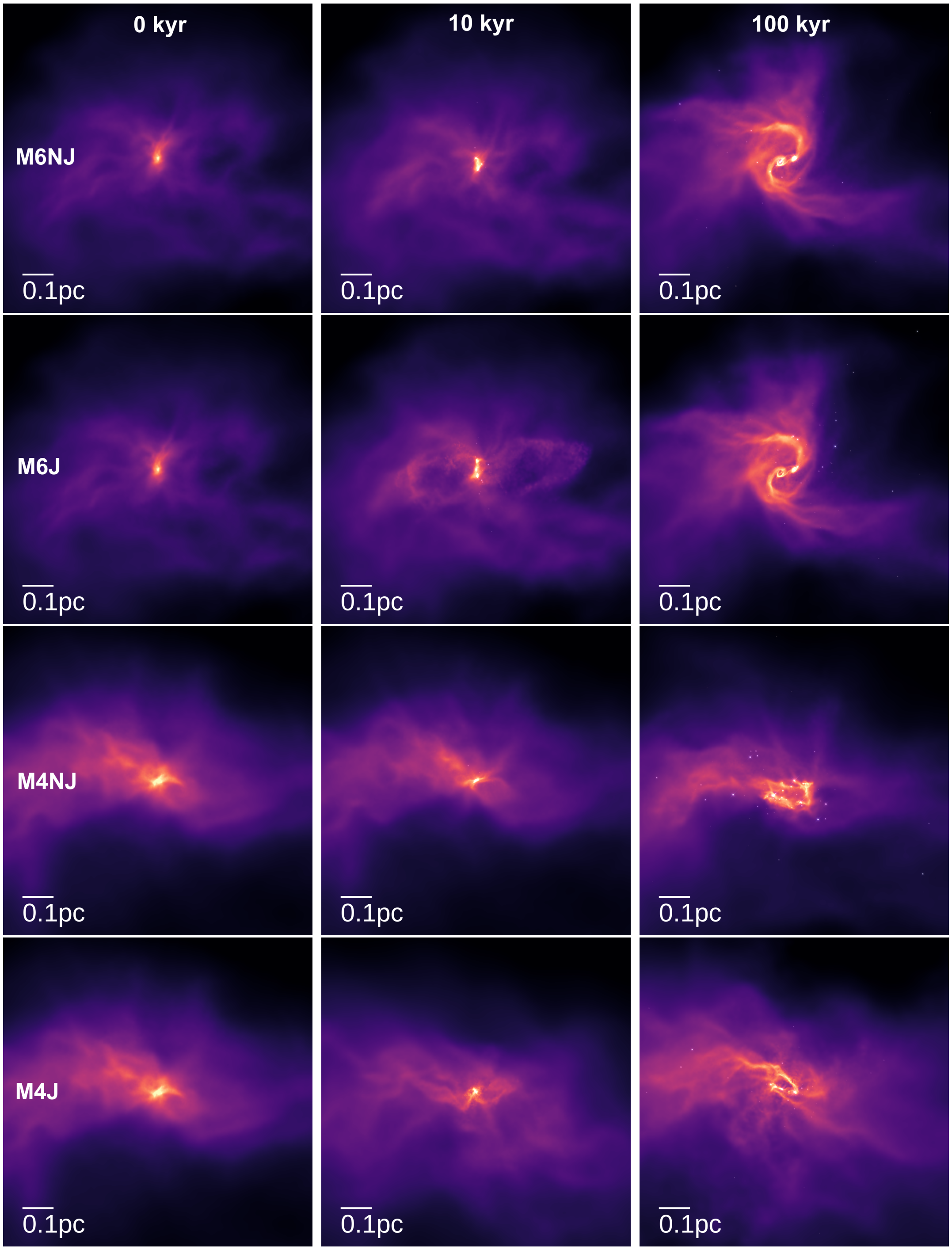


Figure 6.2: Visualizations of the gas density in the star-forming region in our `STARFORGE` simulations. The columns show (from left to right) the different systems 0, 10, and 100 kyr after the formation of the first star. The rows (from top to bottom) are 10^{-6} $Z_{\odot}$ without (M6NJ) and with (M6J) jets, and 10^{-4} $Z_{\odot}$ without (M4NJ) and with (M4J) jets, as labelled in Table 6.1.

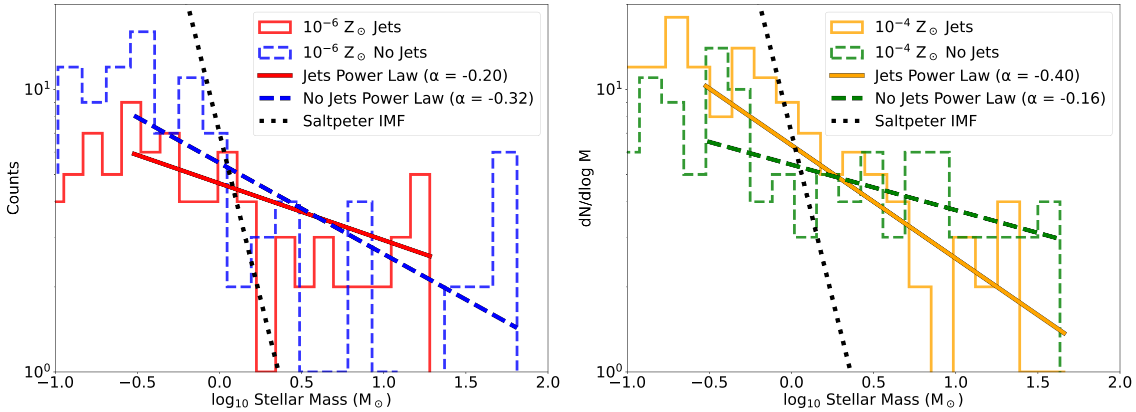


Figure 6.3: The stellar mass function of SIGOs at different metallicities with and without feedback. The left panel shows simulations with $10^{-6}Z_\odot$, and the right panel shows the $10^{-4}Z_\odot$ cases. Dashed lines indicate simulations without jets, whereas solid lines indicate simulations with jets. The dotted lines represent a Salpeter mass function for comparison.

CHAPTER 7

Conclusions

I have presented a study on the effects of the streaming velocity on the formation of the first small-scale structures. I used AREPO and GIZMO, as well as STARFORGE, to numerically simulate SIGOs, as well as Population III star-forming halos, and created theoretical models to underpin my results. I also created semi-analytic models of the Universe to discern observational properties of these objects on large scales.

My semi-analytic model revealed that SIGO abundances vary significantly with local streaming velocity, producing spatial patterns on scales larger than those of typical large-scale density fluctuations. Furthermore, I showed that SIGOs were abundant, even under unfavorable model assumptions (underestimated cooling and condensation of structure)–so abundant, in fact, that they compare to today's globular cluster abundances.

Following on this work, I aimed to create a theoretical basis for SIGO star formation, studying the process of molecular hydrogen cooling within the SIGOs. I showed that high-mass SIGOs tend not to evolve in isolation, explaining why, if SIGOs form some globular clusters, we don't see isolated globulars. I also showed that SIGOs need a minimum mass of 10^5 $M_\odot$ in order to form stars, but that SIGOs above that mass could generally form stars. Furthermore, I showed that SIGOs can grow through accretion, presenting another pathway for their evolution.

Given these theoretical results, the next step was to simulate SIGO collapse with explicit star formation. I showed that SIGOs are capable of forming dense star clusters, with properties similar to proto-globular-clusters. Furthermore, I showed that these SIGO-derived clusters may be detectable with JWST, presenting an exciting pathway for the observation of a SIGO. These simulations also lent themselves to studies of the effect of streaming on dwarf galaxy formation: I showed that streaming acts to increase star formation rates in individual dwarf galaxies even as it lowers the overall abundance of dwarf galaxies (at a given mass).

Next, it was important to begin to model feedback processes–specifically jet feedback–to understand how easily star formation in a SIGO could be quenched. I studied this at very high resolution, allowing me to simultaneously probe the inner regions of the SIGO and understand the spatial extent of the SIGO's star cluster. This also permitted a study of the stellar mass function of the SIGOs, which revealed that their IMF was top-heavy, implying very high light-to-mass ratios.

In sum, these results suggest that SIGOs are a promising route for star cluster formation and a useful class of objects for further study. If SIGOs commonly form massive star clusters, my results show that this will imprint a signal in the spatial variations of globular cluster abundances on scales of tens of Mpc, a promising pathway for the indirect observation of SIGOs in the immediate future. My results also show that SIGO-derived clusters may be observable even at the high redshifts when they formed. Even if this process only rarely forms massive clusters (not contributing significantly to the present-day globular cluster population), the identification of even a single SIGO is evidence for the streaming velocity and thus ΛCDM. Further research into unique observational signatures of these clusters may provide a powerful new tool to test our cosmological models.

As data from JWST's high-redshift observations of the Universe continues to stream in, small-scale structures are revealed at earlier times than ever before. Theoretical modelling of the process of formation of the first dwarf galaxies and their companion structures is timely, to contextualize the anomalous behavior of star formation in high-redshift observations. Moving forward, simulations and observations of the first galaxies and of SIGOs with streaming have the potential to constrain non-ΛCDM cosmologies, and provide a sensitive probe of our Universe's properties.

APPENDIX A

Procedure for normalizing analytic model

There are two steps in producing the normalization factor A. First, we match the analytical results to each simulation run at each given value of $v_{\rm bc}$ using Equation (2.6). For completeness, we show the equation again here, with terms labelled as coming from simulations or analytic models:

$$\left.\frac{dN(u_{\rm vbc})}{dM}\right|_{\rm SIGO,Sim} = A(u_{\rm vbc})\left(\left.\frac{dN(u_{\rm vbc}=0)}{dM}\right|_{\rm DM,Analytic} - \left.\frac{dN(u_{\rm vbc})}{dM}\right|_{\rm DM,Analytic}\right) \tag{A.1}$$

where we remind the reader that $u_{\rm vbc} = v_{\rm bc}/\sigma_{\rm vbc}$ and where the normalization factor is assumed to take the form:

$$A(u_{\rm vbc}) = C \times u_{\rm vbc}^n \ , \tag{A.2}$$

We use the Levenburg-Marquardt method to find the optimal value of A for each simulated stream velocity: $u_{\rm vbc} = 1$, 2, and 3. We limited this fit to masses above $4 \times 10^4 M_\odot$, as below this mass, the effects of our limited resolution affect our simulation data. These normalization values and their corresponding stream velocities are depicted in Figure A.1 as red points. They are also listed in the table below, with error estimates δA:

Second, we fit the points for all three simulations for the different $v_{\rm bc}$ values, using a linear model fit in log-log space. We find best-fit values of C and n, using the above data for A, we find that $C = 6.3 \times 10^{-4} \pm 2.1 \times 10^{-5}$ and $n = 2.43 \pm 0.04$, shown in Figure A.1 as the

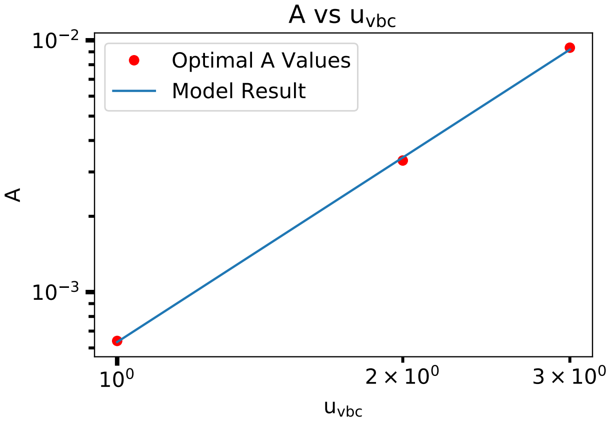


Figure A.1: Plot of the agreement between our model normalization parameter and our optimal normalization parameters obtained from comparison with simulations, as a function of stream velocity.

blue line.

u_{vbc}	1	2	3
A	6.40×10^{-4}	3.33×10^{-3}	9.35×10^{-3}
δA	1.45×10^{-5}	6.12×10^{-5}	6.65×10^{-5}

Table A.1: Optimal values of our normalization factor A with error estimates for each of our simulated stream velocities.

APPENDIX B

Parameters for approximate SIGO number density model

In order to enable future semi-analytic studies of SIGO number densities, we compute an approximate formula to estimate SIGO abundances as a function of stream velocity and mass. This formula provides results corresponding to the output of Equation (2.11).

We found a form for this relation given by Equation (2.12), repeated here for convenience:

$$\log_{10}\left(\frac{N(>M_{\rm SIGO})}{\rm Mpc^3}\right) \;=\; a(u_{\rm vbc})\left\{\log_{10}\left(\frac{M_{\rm SIGO}}{\rm M_\odot}\right)\right\}^{b(u_{\rm vbc})} + c(u_{\rm vbc})\ ,$$

where

$$a(u_{\rm vbc}) \;=\; a_1u_{\rm vbc}^5 + a_2u_{\rm vbc}^4 + a_3u_{\rm vbc}^3 + a_4u_{\rm vbc}^2 + a_5u_{\rm vbc} + a_6\ , \tag{B.1}$$

$$b(u_{\rm vbc}) \;=\; b_1u_{\rm vbc}^5 + b_2u_{\rm vbc}^4 + b_3u_{\rm vbc}^3 + b_4u_{\rm vbc}^2 + b_5u_{\rm vbc} + b_6\ , \tag{B.2}$$

and

$$c(u_{\rm vbc}) \;=\; c_1u_{\rm vbc}^5 + c_2u_{\rm vbc}^4 + c_3u_{\rm vbc}^3 + c_4u_{\rm vbc}^2 + c_5u_{\rm vbc} + c_6\ . \tag{B.3}$$

Using a nonlinear least squares regression, we find the best fit parameters to match this model to our analytic results. The best fit parameters are reported in Table B.1. An example of the agreement between our analytic model and the fit to the model is presented in Figure B.1. The results of Equation (2.11) are shown in blue, giving $N(>M)_{\rm SIGO}$ for M $> 10^5{\rm M}_\odot$, with the result of Equation (2.12) shown in red. Notably, this reported fit holds

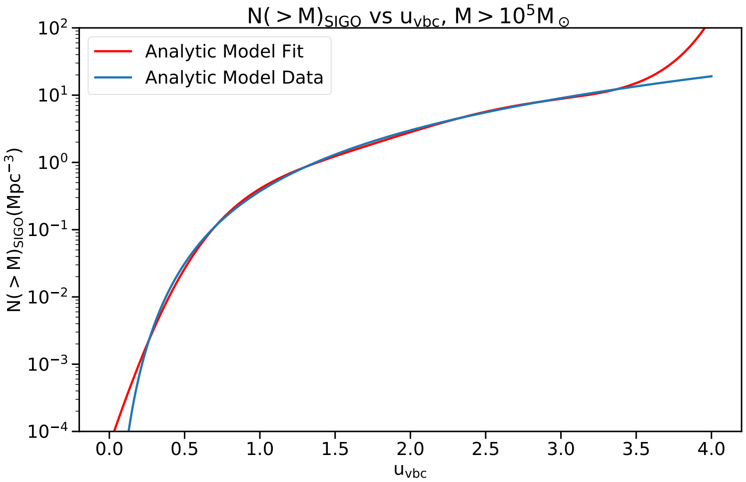


Figure B.1: Plot illustrating the agreement between our analytic model for SIGO abundances (Equation (2.11)) and our reported fit to the model output (Equation (2.12)) at M$>10^5$ M$_\odot$. The two equations agree closely for few $\times 10^4$ M$_\odot$ $<$ M $< 8 \times 10^5$ M$_\odot$, and for $u_{\rm vbc} < 3.4$.

for 2.5×10^4 M$_\odot$ $\lesssim$ M $\lesssim 8 \times 10^5$ M$_\odot$, and for $v_{\rm bc} < 3.4\sigma_{\rm vbc}$.

	1	2	3	4	5	6
a	0.0014	-0.0165	0.0745	-0.159	-0.1554	-0.1066
b	0.0104	-0.1248	0.5661	-1.186	1.051	2.067
c	0.0718	-0.7767	3.239	-6.618	7.284	-1.238

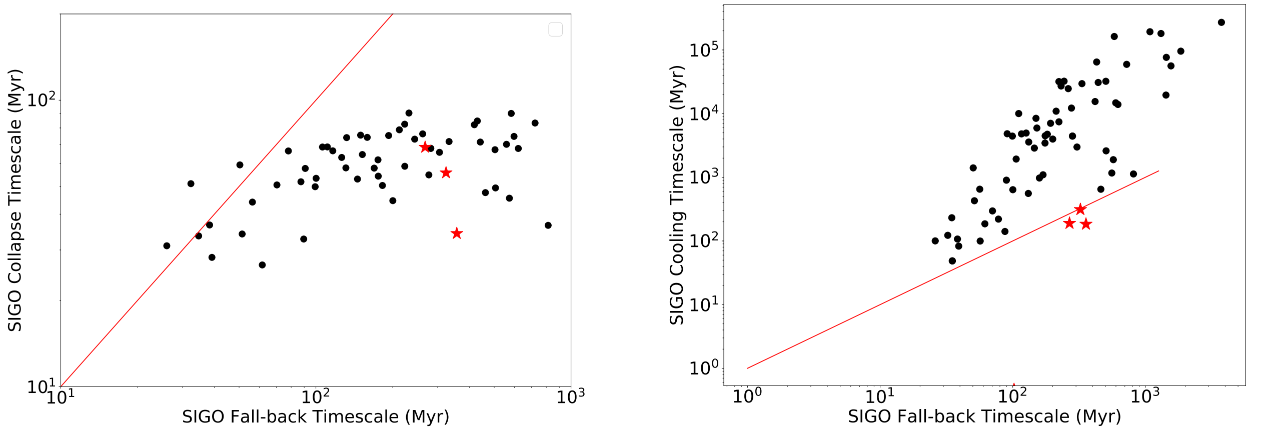
Table B.1: Best fit parameters used in Equation (2.12) that match analytic model results from Equation (2.11), obtained using a nonlinear least squares regression. Coefficients are given to high precision due to the sensitivity of the model to the coefficients.

APPENDIX C

Comparison of Important Timescales

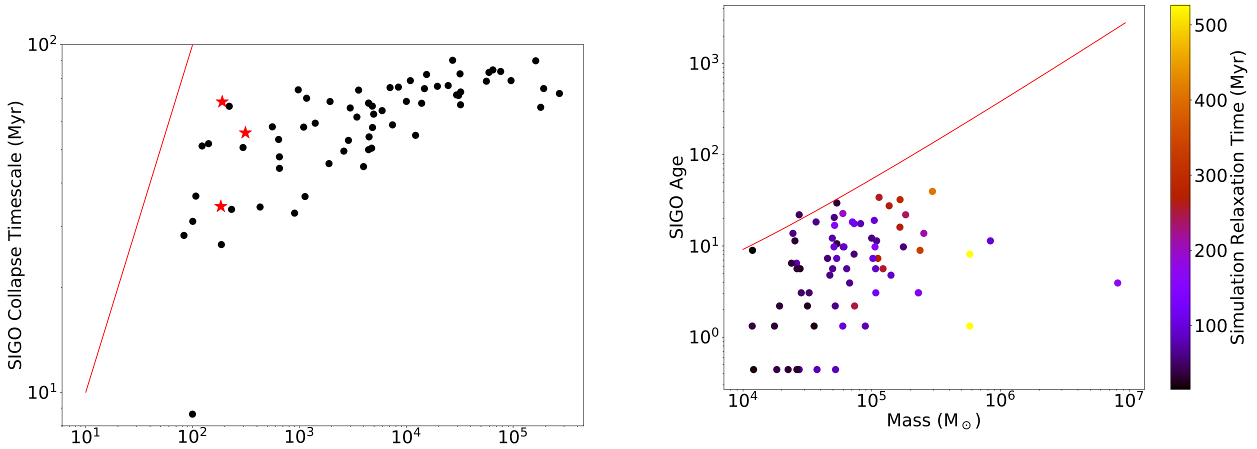
Figure C.1 displays another comparison of all of these timescales, as a proof of concept to aid the reader in gaining a physical intuition of how the various timescales compare. This allows a qualitative prediction of which SIGOs are able to collapse. The top left panel of the Figure shows the collapse timescale of SIGOs in Run 2vH2. This is plotted against the SIGOs' median fall-back timescale (across time) to their nearest halo. This median corrects for errors in the FOF finder for DM halos: if a nearby halo merges with another, creating an ellipsoidal DM density distribution, the spherical treatment of halos in the simulation creates a deceptively long free-fall timescale, so it is preferable for the purposes of physical intuition of this evolution to refer instead to a median, which is more stable and not subject to the variations in halo-SIGO separation caused by mergers. As one can see, most SIGOs have a shorter collapse timescale than median fall-back timescale (and are below the red line in the panel reflecting equal timescales). This indicates that the cooling timescale is key to determining whether or not SIGOs will collapse before falling into nearby DM halos.

Panels (b) and (c) of Figure C.1 show two other comparisons of the timescales involved in this evolution. Panel (b) shows a comparison of the cooling timescale and the median fall-back timescale for all $z = 20$ SIGOs. This panel further demonstrates that while the collapse timescale is important for the evolution of SIGOs, that timescale alone is usually insufficient to determine whether or not a SIGO will collapse independently of a DM halo– the cooling timescale tends to be a bigger driver of such SIGOs' evolution, especially at lower masses.



a) Collapse timescale vs. median fall-back timescale

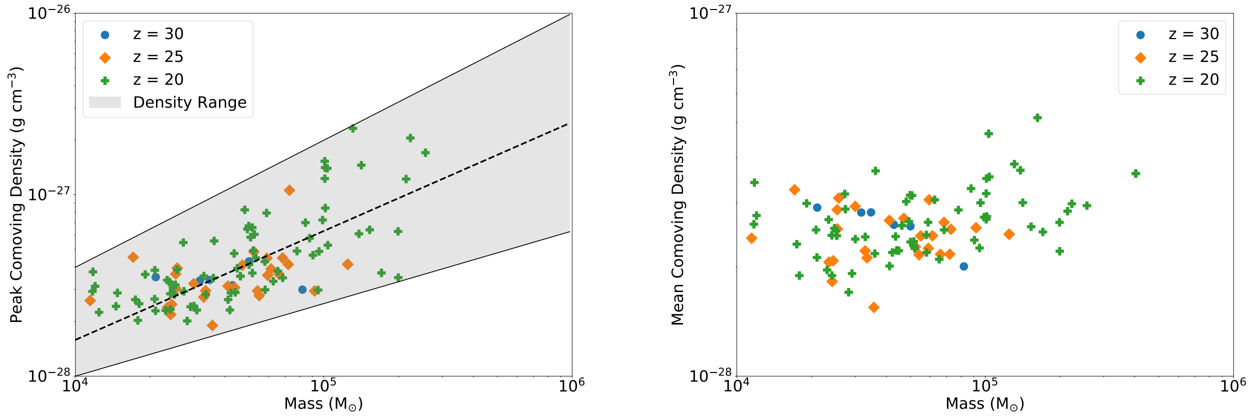
b) Cooling timescale vs. median fall-back timescale



c) Collapse timescale vs. Cooling timescale

d) Relaxation timescale at simulation resolution

Figure C.1: **Timescale Comparisons:** Here we show a number of important timescales to the evolution of SIGOs at $z = 20$ in Run 2vH2. In the first 3 panels (a-c), a black dot indicates a SIGO from the simulation. A red star indicates a SIGO that is likely to form stars outside of a DM halo, with cooling and collapse timescales shorter than its median fallback timescale. Note that 2 star-forming SIGOs could not be pictured in this figure, due to their very short cooling and collapse timescales indicative that collapse has already occurred. The red lines in these figures indicate equality between the two timescales being compared. The final panel (d) shows the ages of $z = 20$ SIGOs at the resolution of the simulation plotted against their masses. The red line shows the relaxation timescale at simulation resolution of a SIGO with typical properties as a function of mass.



(a) (b)

Figure C.2: **Typical SIGO Densities:** here we show the peak (left panel) and mean (right panel) comoving densities of SIGOs at various redshifts in Run 2vH2. In each panel, a mark indicates a SIGO from the simulation. The peak density is calculated as the volume-weighted average density of the most dense Voronoi gas cell in the SIGO and its 9 nearest neighbours. The representative range of peak densities we use in this paper (Eqs. 3.7 and 3.8) is shown in the left panel in gray shading. The linear regression to the $z = 20$ density data is presented as a dashed line.

Panel (c) of Figure C.1 compares the cooling timescale for SIGOs to their collapse timescale, showing that the cooling timescale for SIGOs is, in fact, always longer than the collapse timescale at formation. Note that this cooling timescale includes an isochoric cooling term for the SIGO to begin to collapse (Eq. 3.10): this does not reflect the balance between cooling times and collapse times once SIGOs achieve the Jeans mass and begin to collapse.

Unlike the other panels, Panel (d) of Figure C.1 shows a comparison of the relaxation timescale (Eq. 3.14) of the simulation with mass, showing that we have lost some SIGOs at low masses to relaxation before $z = 20$ in the simulation.

SIGOs generally form in a similar range of densities regardless of mass. Their early evolution occurs primarily through cooling, and sometimes mass growth through gas accretion, rather than through immediate global collapse. However, SIGOs' collapse timescales depend on their mass, as their peak densities in their cores have a mass dependence. Figure C.2 shows the relation between the peak and mean densities in SIGOs as a function of mass at several redshifts, to help illuminate this effect. In the right panel of this Figure, even SIGOs that have evolved for some time at $z = 20$ do not have mean gas densities that exhibit a strong dependence on mass. This is expected: most SIGOs at this time are around 10 Myr old, and nearly all of these SIGOs have not yet had a chance to globally cool and collapse on $\gtrapprox 50$ Myr timescales. However, there are hints of substructure formation in high-mass SIGOs. In the left panel of Figure C.2, we show the trend in peak density with mass, excepting 4 $z = 20$ SIGOs that have already partially or fully collapsed and reached much higher densities. A linear regression yields an approximate formula for the relation between peak comoving density and mass at $z = 20$: $\log_{10}(\rho_{\rm peak,com}\ ({\rm g/cm^3})) = 0.6 \times \log_{10}({\rm M_{SIGO}}({\rm M_\odot}))$ - 30.2. As one can see, there is a clear positive correlation between peak density and mass–one that becomes stronger at later redshifts. At $z = 30$, the Spearman coefficient is -0.37 ($N = 6$, $p = 0.47$), at $z = 25$ it is 0.50 ($N = 28$, $p = 0.007$), and at $z = 20$ it is 0.78 ($N = 85$, $p = 4 \times 10^{-19}$). Notably, this trend remains even if the highest-mass SIGOs at $z = 20$ are excluded, indicating that it is not only a function of increased structure formation. If only SIGOs with a mass lower than the highest mass $z = 25$ SIGO are studied, the Spearman coefficient at $z = 20$ is 0.70 ($N = 69$, $p = 1.7 \times 10^{-11}$).

It is also important to consider the effects of varying the baryon fraction cutoff on SIGO abundances. In Figure C.3, we show the effect of varying this cutoff with and without $\rm v_{bc}$, at a variety of redshifts when SIGOs from our $z = 20$ snapshot formed. Below $\rm f_b = 0.6$, the abundances of SIGOs removed by increasing our $\rm f_b$ cutoff are comparable in our $\rm v_{bc} =$ 0 and 2σ runs, indicating that raising this cutoff helps us to distinguish spurious objects from true SIGOs. However, above $\rm f_b = 0.6$, many more SIGOs are being removed in the 2σ simulation than spurious objects from the 0 $\rm v_{bc}$ simulations, indicating that we do not benefit from further raising the $\rm f_b$ cutoff at these redshifts. Notably, at $z = 25$, this threshold is sufficient to remove all but 1 spurious SIGO. There are still a small fraction of spurious objects within the SIGO population at $z = 20$, however, so the SIGOs which are marked as "potential star-forming SIGOs" in Figure 3.6 were then verified by ensuring that no "SIGOs"

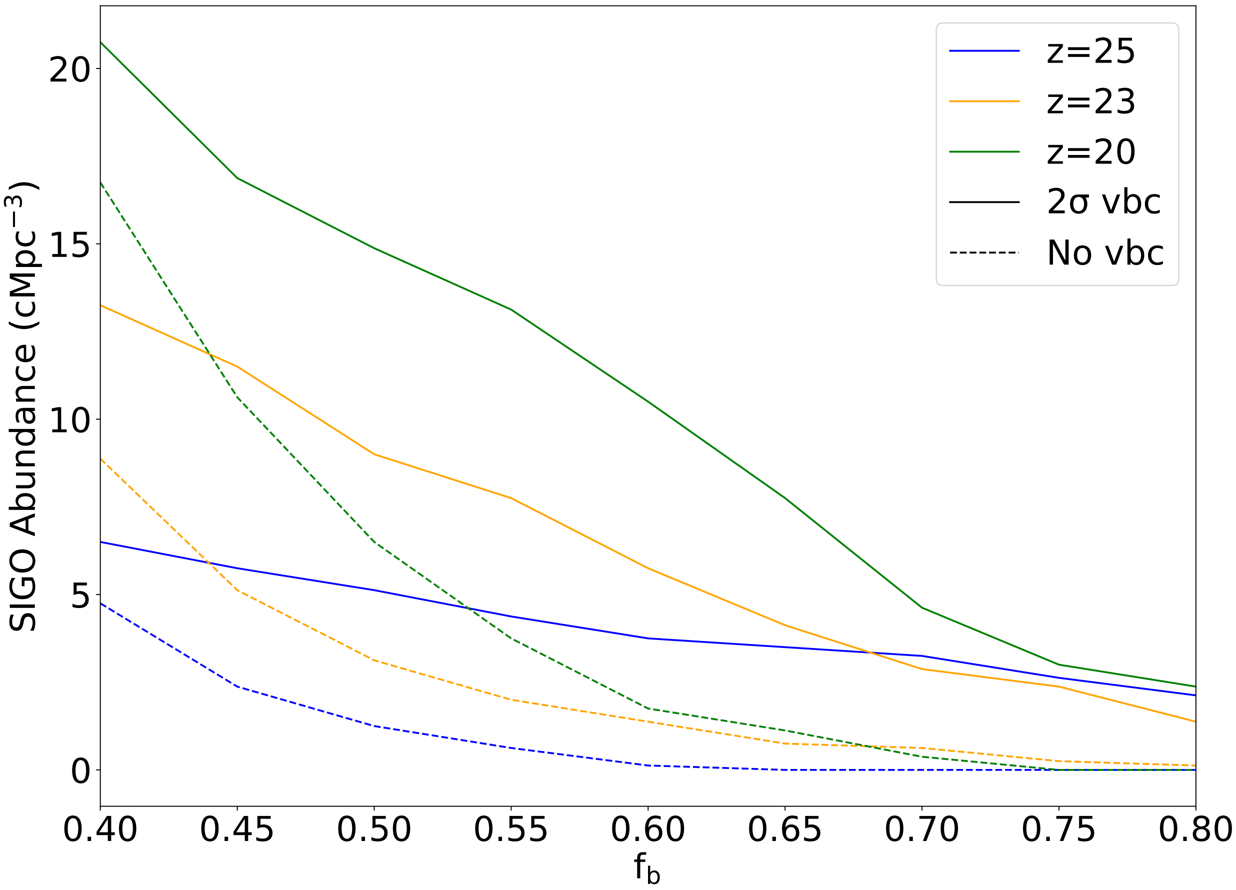


Figure C.3: **Choosing a baryon fraction cutoff:** in this Figure we show the effect of different baryon fraction cutoffs on calculated SIGO abundances at $z = 25$ (blue), $z = 23$ (orange), and $z = 20$ (green). Results from Run 2vH2 are shown with solid lines, whereas results from Run 0vH2 are shown with dashed lines. This Figure aims to justify our choice in f_b cutoff.

in the no stream velocity simulation existed nearby (in the vicinity of the same identified parent halo).

APPENDIX D

Effect of Time on Cumulative Star Formation Rate

We estimate the star formation rate directly from the rate at which star particles are formed in our simulation. This is shown in Figure D.1. We note that we do not include radiative feedback, which eventually acts to quench star formation. As mentioned in Section 4.5, the effects of feedback on our simulation may be balanced on timescales shorter than the quenching timescale by our cooling mechanism, which serves as a lower bound to the true cooling rate in SIGOs for the reasons given in the aforementioned section. The likely quenching timescale is $\sim 10-20$ Myr, so in order to estimate the star formation rate in the SIGO during its initial starburst, we need to select a star formation rate over this period. As shown in Figure D.1, the star formation rate over this period is not a large source of uncertainty, as it varies only by about a factor of two over the timespan. As such, we use a period of 20 Myr following the start of star formation in the SIGO to calculate a star formation rate, 557 $M_\odot$ Myr^{-1}.

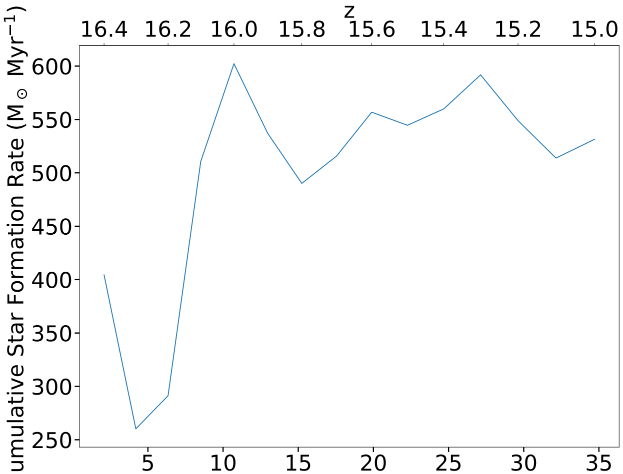


Figure D.1: Plot of the cumulative star formation rate (calculated as total stellar mass divided by time since the start of star formation) in the SIGO as a function of time. Regardless of time period considered, the star formation rate within the SIGO is the same to within a factor of about 2.

APPENDIX E

Modelling the Effect of Streaming Varying on Cosmological Parameters

E.1 Adjusting for different values of σ_8

As discussed in Section 5.3.2, we use an enhanced value of σ_8 in our simulations to increase our statistical power and simulate a rare, overdense region of the Universe. For the purposes of Figure 5.2 and Figure 5.3, we adjust our results to a Universe with different values of σ_8, using the following analytical correction.

The comoving number density of halos of mass M at redshift z is given by

$$\frac{dn}{dM} = \frac{\rho_0}{M} f_{ST} \left| \frac{dS}{dM} \right| \ , \tag{E.1}$$

where we use the Sheth et al. (2001) mass function that both fits simulations and includes non-spherical effects on the collapse. The function f_{ST} is the fraction of mass in halos of mass M:

$$\begin{aligned} f_{ST}(\delta_c, S) &= A' \frac{\nu}{S} \sqrt{\frac{a'}{2\pi}} \left[1 + \frac{1}{(a'\nu^2)^{q'}} \right] \exp\left(\frac{-a'\nu^2}{2} \right) \qquad \text{(E.2)} \\ \text{and} \quad \nu &= \frac{\delta_c}{\sigma} = \frac{\delta_c}{\sqrt{S}} \ . \end{aligned}$$

We use best-fit parameters $A' = 0.75$ and $q' = 0.3$ (Sheth & Tormen, 2002). We can then

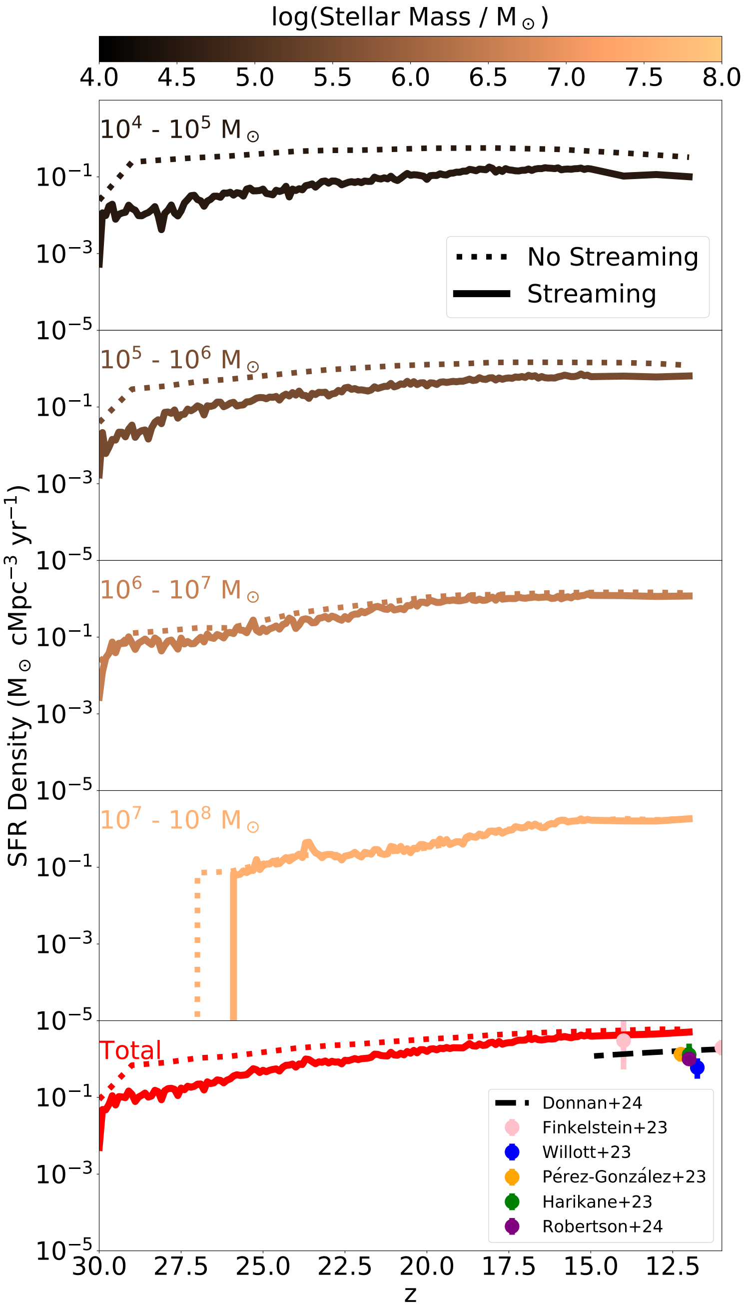


Figure E.1: **Figure 5.2 without adjusting for σ_8:** This plot shows star formation rate densities with (solid lines) and without (dotted lines) streaming at different object masses, with our enhanced simulation parameter $\sigma_8 = 1.7$.

calculate the relative fraction $f_{0.826}$ of DM halos with said mass M with $\sigma_8 = 0.826$ compared to $\sigma_8 = 1.7$ as

$$f_{0.826} = \left(\frac{dn}{dM}\right)_{\sigma_8=0.826} \left(\frac{dn}{dM}\right)^{-1}_{\sigma_8=1.7} . \tag{E.3}$$

We apply this adjustment to re-weight the abundances of halos of different masses and match stellar objects to their nearest host halos in order to re-weight these objects in turn (thus producing an adjusted SFR density and number density for each object). Corrections to different values of σ_8 follow the same process.

E.2 Determining density cutoffs for the Kennicutt-Schmidt Relation

In order to calculate the Kennicutt-Schmidt relation, we fit data using the linear relation Equation 5.3.3. Using this and the full data set, we find $\alpha = 1.340 \pm 0.004$ for the case without streaming and $\alpha = 1.299 \pm 0.007$ for the case with $2\sigma_{\text{vbc}}$ streaming. However, our estimates of the slope of the Kennicutt-Schmidt relation are systematically biased by differing behavior at low, medium, and high gas densities, as well as by the unbalanced populations of low-density star-forming gas objects with and without streaming. This is shown in Figure E.2. As mentioned in Section 5.3, the timescale on which a significant fraction of the gas in the highest-gas-density objects are converted to stars is comparable or less than our timestep between snapshots. This drops the slope of the Kennicutt-Schmidt relation to 1 at the highest densities, corresponding to complete conversion of gas to stars (up to an efficiency factor). We can calculate the point at which this becomes significant by introducing a low-density cutoff to our data (Figure E.2) and systematically raising it by the gas density intervals present in our data. This allows us to isolate the region of data which is impacted by the changing slope. We take the high-density cutoff to be the point at which the slope began to decrease monotonically (given some unavoidable noise) to 1, where the errors due to time resolution dominate all other sources of error. This cutoff is calculated

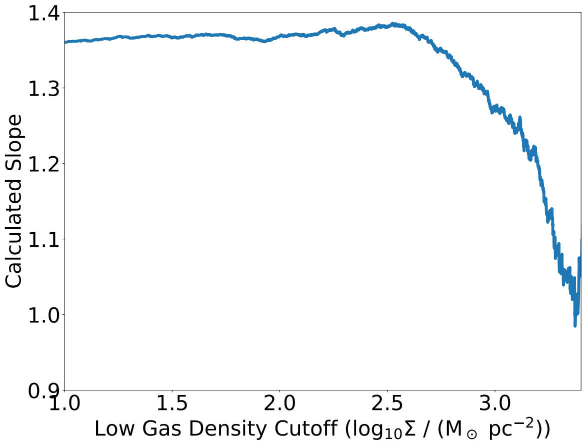


Figure E.2: **Kennicutt-Schmidt Relation Density Cutoff Selection:** This plot shows the calculated slope of the Kennicutt-Schmidt relation (through a linear regression) as we filter the data with and without streaming using a varying density cutoff (on the x axis). We use the limiting behavior of this plot at high densities to determine an appropriate range of densities over which our results can be taken to be representative of the effect of streaming. Only the more limiting of the two simulations is shown.

to be $\Sigma_{\mathrm{high}} = 10^{2.5}$ $\mathrm{M}_\odot$ pc^{-2} (rounded to the nearest 0.1 dex) in the no-streaming case, and $\Sigma_{\mathrm{high,streaming}} = 10^{3.0}$ $\mathrm{M}_\odot$ pc^{-2} in the less-limiting case with streaming. We select the more limiting of the two cutoffs.

The second complication in our Kennicutt-Schmidt slope is the sparsity of low-density gas objects compared to higher-density gas objects in our data with streaming compared to our data without it. If the Kennicutt-Schmidt relation in our simulation takes on a different slope at low densities, the over-representation of low-density objects in the simulation without streaming would produce a seemingly different slope in a linear KSR fit, even if the behavior of the KSR with compared to without streaming is the same at low and intermediate densities. To minimize this effect, we bin the data into bins of width 0.1 dex as described in Section 5.3, and calculate a low-density cutoff as the point at which the sparser low-density data set (the simulation with streaming) has 3 or fewer objects per bin, at which point we have insufficient statistics, which sets our cut at $\Sigma_{\mathrm{Gas}} = 10^{0.9}$ $\mathrm{M}_\odot$ pc^{-2}. Our results then describe and are valid for the intermediate density range $10^{0.9}$ $\mathrm{M}_\odot$ $\mathrm{pc}^{-2} \leq \Sigma_{\mathrm{Gas}} \leq 10^{2.5}$ $\mathrm{M}_\odot$ pc^{-2}.

APPENDIX F

Testing STARFORGE

F.1 Convergence Tests

As `STARFORGE` has not previously been used to simulate Pop III or low-metallicity Pop II star formation, we run a suite of 4 simulations at 0.005, 0.01, 0.02, and 0.05 $M_\odot$ resolution at Z= 10^{-4} $Z_\odot$ with jet feedback, allowing us to understand the effects of resolution on our results. These correspond to 6×10^7, 3×10^7, 1.5×10^7, and 6×10^6 particles, respectively. We tested the total mass and cumulative stellar IMF of each run, in order to check convergence of the results at decreasing resolutions. These tests are shown in Figure F.1. The simulation with 0.005 $M_\odot$ mass resolution is shown in blue, 0.01 $M_\odot$ in orange, 0.02 $M_\odot$ in green, and 0.05 $M_\odot$ is shown in red. The left panel shows the cumulative stellar IMF 10^5 yrs after the first star, and the right panel shows the total mass vs time following the formation of the first star.

In the left panel of this Figure, one can see that the stellar IMF is well-converged at mass resolutions below 0.02 $M_\odot$, with the caveat that low-mass stars (those resolved with fewer than 30 or so gas particles) are not converged, as their collapse is not well-resolved. Conclusions can, therefore, be drawn from the high-mass IMF. Similarly, in the right panel, we see that the total mass in stars is well-converged at 10^5 yrs between the 0.005 $M_\odot$ and 0.01 $M_\odot$ runs. Discrepancies occurring before this time may be attributed to our choice to track the time since the formation of the first star, as that is a somewhat stochastic event but made

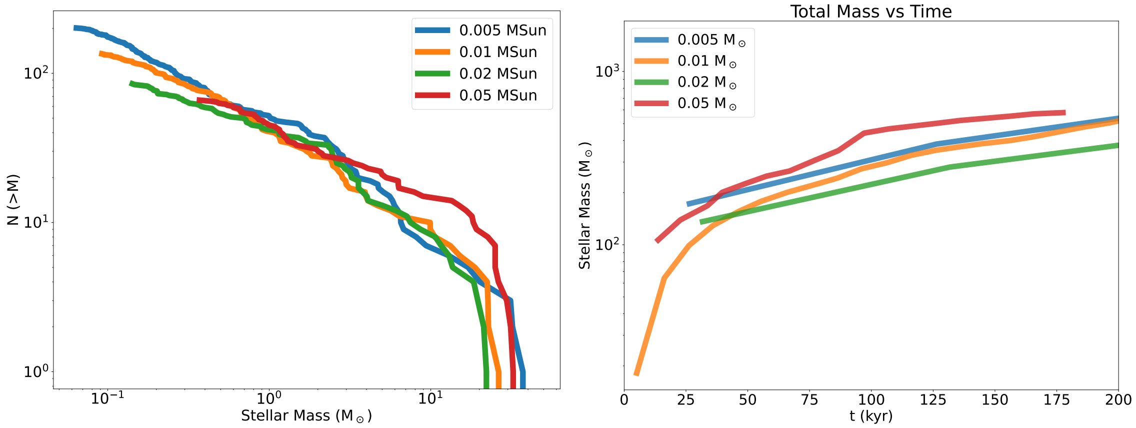


Figure F.1: Convergence tests of the IMF (left) and total stellar mass (right) in simulations of 0.005 $M_\odot$ (blue), 0.01 $M_\odot$ (orange), 0.02 $M_\odot$ (green), and 0.05 $M_\odot$ (red) resolution. The left panel depicts the IMF at 10^5 years after the formation of the first star, and demonstrates convergence of the high-mass end of the IMF (stars resolved with at least 30 gas particles in each run) for resolutions better than 0.02 $M_\odot$.

necessary by the relatively long pre-star-formation evolution of the cloud. These tests give us confidence that, in terms of the high-mass end of the IMF and total stellar masses, 0.01 $M_\odot$ resolution is sufficiently accurate for the present suite of Pop III star formation simulations.

F.2 Phase Diagram

Owing to their lack of dark matter and thus relatively low masses, the early collapse process in SIGOs is characterized by relatively low virial temperatures of only a few hundred K. As seen in Figure F.2, this results in a temperature peak in the early collapse of about 300 K at n$\approx$ 10 cm^{-3}. At this point, molecular hydrogen cooling allows the gas to collapse while cooling to just under 200 K by n$\approx$ 10^5 cm^{-3}, which is slightly lower than traditional Pop III star formation (Klessen & Glover, 2023), perhaps owing to the relatively long free-fall timescale in this system.

Following this initial collapse, characterized by lower temperatures than dark matter halos with a similar gas mass, the SIGO's collapse becomes more similar to that of molecular cool-

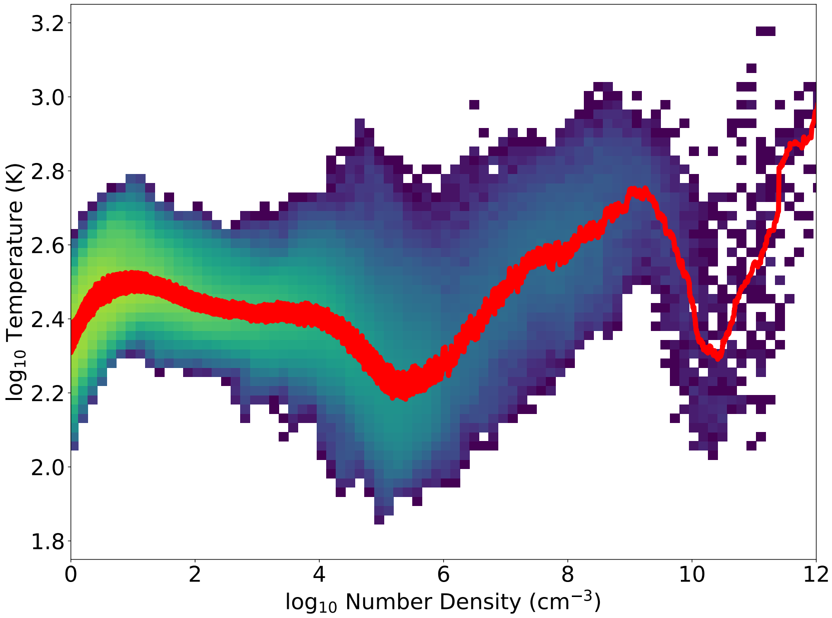


Figure F.2: The temperature and density of collapsing gas in our 10^{-6} $Z_\odot$ STARFORGE simulation at the time of formation of the first star. The red line is a moving average of 100 points nearest to the given number density and the yellow-violet background is a 2D histogram of all of the gas in our simulation at or above a number density of 1 cm^{-3}.

ing minihalos. Protostellar collapse between about $10^5\ \mathrm{cm}^{-3} < n_\mathrm{H} < 10^9\ \mathrm{cm}^{-3}$ is dominated in halos by PdV work from gas accreting onto the core, and not from heating through the halo potential, and in halos of a similar gas mass this causes heating to about 1000 K by $n_\mathrm{H} \sim 10^9\ \mathrm{cm}^{-3}$ (Schauer et al., 2021; Klessen & Glover, 2023). As expected, because this is not driven by dark matter, we see a similar heating process here.

At $n_\mathrm{H} \sim 10^9\ \mathrm{cm}^{-3}$, the 3-body reaction for molecular hydrogen formation becomes important, and as in halos, the gas becomes nearly fully molecular. This creates a brief period during which enhanced molecular hydrogen cooling permits the gas to cool sharply, followed by a return to PdV heating at $n_\mathrm{H} \sim 10^{10}\ \mathrm{cm}^{-3}$ during the runaway collapse to protostellar formation. In this study, we form stars at about few$\times 10^{12}\ \mathrm{cm}^{-3}$, and do not further trace the gas evolution.

Bibliography

Abbott B. P., et al., 2016, Phys. Rev. Lett., 116, 061102

Abbott B. P., et al., 2017a, Classical and Quantum Gravity, 34, 044001

Abbott B. P., et al., 2017b, Phys. Rev. Lett., 119, 161101

Abbott R., et al., 2021a, Physical Review X, 11, 021053

Abbott R., et al., 2021b, Phys. Rev. D, 104, 022005

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

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

Adamo A., et al., 2024, Nature, 632, 513

Antonini F., Gieles M., 2020, Phys. Rev. D, 102, 123016

Antonini F., Barausse E., Silk J., 2015, ApJ, 806, L8

Appel S. M., Burkhart B., Semenov V. A., Federrath C., Rosen A. L., 2022, ApJ, 927, 75

Asaba S., Ichiki K., Tashiro H., 2016, Phys. Rev. D, 93, 023518

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

Barkana R., Loeb A., 2004, ApJ, 609, 474

Bastian N., Lardo C., 2018, Annual Review of Astronomy and Astrophysics, 56, 83

Baumgardt H., Kroupa P., 2007, MNRAS, 380, 1589

Bekki K., Yong D., 2012, MNRAS, 419, 2063

Belczynski K., Sadowski A., Rasio F. A., Bulik T., 2006, The Astrophysical Journal, 650, 303–325

Belczynski K., Holz D. E., Bulik T., O'Shaughnessy R., 2016, Nature, 534, 512

Benedict G. F., et al., 2002, AJ, 123, 473

Bennett C. L., et al., 1996, ApJ, 464, L1

Bergin E. A., Tafalla M., 2007, ARA&A, 45, 339

Bertoldi F., McKee C. F., 1992, ApJ, 395, 140

Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton University Press

Böker T., Sarzi M., McLaughlin D. E., van der Marel R. P., Rix H.-W., Ho L. C., Shields J. C., 2004, AJ, 127, 105

Bovy J., Dvorkin C., 2013, ApJ, 768, 70

Bradford J. D., et al., 2011, ApJ, 743, 167

Breivik K., Chatterjee S., Andrews J. J., 2019, ApJ, 878, L4

Breivik K., et al., 2020, ApJ, 898, 71

Bromm V., 2013, Reports on Progress in Physics, 76, 112901

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

Brown I. A., 2013, Journal of Cosmology and Astroparticle Physics, 2013, 003–003

Brown T. M., et al., 2014, ApJ, 796, 91

Burkhart B., 2018, ApJ, 863, 118

Burkhart B., 2021, PASP, 133, 102001

Burkhart B., Lazarian A., 2012, ApJ, 755, L19

Burkhart B., Falceta-Gonçalves D., Kowal G., Lazarian A., 2009, ApJ, 693, 250

Cain C., D'Aloisio A., Iršič V., McQuinn M., Trac H., 2020, ApJ, 898, 168

Chaboyer B., 1999, in Heck A., Caputo F., eds, Astrophysics and Space Science Library Vol. 237, Post-Hipparcos Cosmic Candles. p. 111 (`arXiv:astro-ph/9808202`), doi:10.1007/978-94-011-4734-7_7

Chatterjee S., Rodriguez C. L., Rasio F. A., 2017, ApJ, 834, 68

Chen S., Richer H., Caiazzo I., Heyl J., 2018, ApJ, 867, 132

Chiaki G., Wise J. H., 2019, MNRAS, 482, 3933

Chiou Y. S., Naoz S., Marinacci F., Vogelsberger M., 2018, MNRAS, 481, 3108

Chiou Y. S., Naoz S., Burkhart B., Marinacci F., Vogelsberger M., 2019, ApJ, 878, L23

Chiou Y. S., Naoz S., Burkhart B., Marinacci F., Vogelsberger M., 2021, ApJ, 906, 25

Clauset A., Shalizi C. R., Newman M. E. J., 2009, SIAM Review, 51, 661

Conaboy L., Iliev I. T., Fialkov A., Dixon K. L., Sullivan D., 2022, arXiv e-prints, p. arXiv:2207.11614

Conaboy L., Iliev I. T., Fialkov A., Dixon K. L., Sullivan D., 2023, MNRAS, 525, 5479

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

Cunningham A. J., Klein R. I., Krumholz M. R., McKee C. F., 2011, ApJ, 740, 107

Dalal N., Pen U.-L., Seljak U., 2010, J. Cosmology Astropart. Phys., 11, 007

Donnan C. T., et al., 2024, arXiv e-prints, p. arXiv:2403.03171

Druschke M., Schauer A. T. P., Glover S. C. O., Klessen R. S., 2020, MNRAS, 498, 4839

Elmegreen B. G., 2010, ApJ, 712, L184

Elmegreen B. G., Efremov Y. N., 1997, ApJ, 480, 235

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

Ferrara A., Pettini M., Shchekinov Y., 2000, Monthly Notices of the Royal Astronomical Society, 319, 539

Fialkov A., Barkana R., 2014, MNRAS, 445, 213

Fialkov A., Barkana R., Tseliakhovich D., Hirata C. M., 2012, MNRAS, 424, 1335

Finkelstein S. L., et al., 2023, arXiv e-prints, p. arXiv:2311.04279

Fragione G., Kocsis B., 2018, Phys. Rev. Lett., 121, 161103

Galli D., Palla F., 1998, A&A, 335, 403

Glover S., 2013, in Wiklind T., Mobasher B., Bromm V., eds, Astrophysics and Space Science Library Vol. 396, The First Galaxies. p. 103 (`arXiv:1209.2509`), doi:10.1007/978-3-642-32362-1_3

Gray W. J., Scannapieco E., 2011, ApJ, 742, 100

Greif T. H., Springel V., White S. D. M., Glover S. C. O., Clark P. C., Smith R. J., Klessen R. S., Bromm V., 2011, ApJ, 737, 75

Grillmair C. J., Freeman K. C., Irwin M., Quinn P. J., 1995, AJ, 109, 2553

Grudić M. Y., Guszejnov D., Hopkins P. F., Offner S. S. R., Faucher-Giguère C.-A., 2021a, MNRAS, 506, 2199

Grudić M. Y., Kruijssen J. M. D., Faucher-Giguère C.-A., Hopkins P. F., Ma X., Quataert E., Boylan-Kolchin M., 2021b, MNRAS, 506, 3239

Grudić M. Y., Hafen Z., Rodriguez C. L., Guszejnov D., Lamberts A., Wetzel A., Boylan-Kolchin M., Faucher-Giguère C.-A., 2022, arXiv e-prints, p. arXiv:2203.05732

Gunn J. E., 1980, in Hanes D., Madore B., eds, Globular Clusters. p. 301

Guszejnov D., Grudić M. Y., Offner S. S. R., Faucher-Giguère C.-A., Hopkins P. F., Rosen A. L., 2022, MNRAS, 515, 4929

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

Harikane Y., et al., 2023, ApJS, 265, 5

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

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

Harris W. E., Harris G. L. H., Alessi M., 2013, ApJ, 772, 82

Hassan S., et al., 2023, ApJ, 958, L3

Hegde S., Furlanetto S. R., 2023, arXiv e-prints, p. arXiv:2304.03358

Heggie D. C., Hut P., 1996, in Hut P., Makino J., eds, IAU Symposium Vol. 174, Dynamical Evolution of Star Clusters: Confrontation of Theory and Observations. p. 303 (`arXiv:astro-ph/9511115`)

Hirano S., Hosokawa T., Yoshida N., Umeda H., Omukai K., Chiaki G., Yorke H. W., 2014, ApJ, 781, 60

Hirano S., Hosokawa T., Yoshida N., Omukai K., Yorke H. W., 2015, MNRAS, 448, 568

Hirano S., Hosokawa T., Yoshida N., Kuiper R., 2017a, preprint, (`arXiv:1709.09863`)

Hirano S., Hosokawa T., Yoshida N., Kuiper R., 2017b, Science, 357, 1375

Hirano S., Shen Y., Nishijima S., Sakai Y., Umeda H., 2023, MNRAS, 525, 5737

Hoang B.-M., Naoz S., Kocsis B., Rasio F. A., Dosopoulou F., 2018, ApJ, 856, 140

Hollenbach D., McKee C. F., 1979, ApJS, 41, 555

Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., et al., 2023, MNRAS, 519, 3154

Hummel J. A., Stacy A., Bromm V., 2016, MNRAS, 460, 2432

Ibata R., Nipoti C., Sollima A., Bellazzini M., Chapman S. C., Dalessandro E., 2013, MNRAS, 428, 3648

Jaura O., Glover S. C. O., Wollenberg K. M. J., Klessen R. S., Geen S., Haemmerlé L., 2022, MNRAS, 512, 116

Jog C. J., 2013, MNRAS, 434, L56

Johnson J. L., 2010, Monthly Notices of the Royal Astronomical Society, 404, 1425

Kaiser N., 1984, ApJ, 284, L9

Kamann S., et al., 2018, MNRAS, 473, 5591

Klessen R. S., Glover S. C. O., 2023, ARA&A, 61, 65

Kocsis B., Levin J., 2012, Phys. Rev. D, 85, 123005

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

Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astronomy Letters, 44, 8

Kremer K., et al., 2020, ApJS, 247, 48

Kroupa P., 2001, MNRAS, 322, 231

Kruijssen J. M. D., 2015, MNRAS, 454, 1658

Krumholz M. R., McKee C. F., 2005, ApJ, 630, 250

Krumholz M. R., Tan J. C., 2007, ApJ, 654, 304

Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227

Kulkarni M., Visbal E., Bryan G. L., 2021, ApJ, 917, 40

Lada C. J., Forbrich J., Lombardi M., Alves J. F., 2012, ApJ, 745, 190

Lake W., Naoz S., Chiou Y. S., Burkhart B., Marinacci F., Vogelsberger M., Kremer K., 2021, ApJ, 922, 86

Lake W., et al., 2023a, ApJ, 943, 132

Lake W., et al., 2023b, ApJ, 956, L7

Lake W., et al., 2024a, arXiv e-prints, p. arXiv:2410.02868

Lake W., et al., 2024b, ApJ, 973, 115

Larson R. B., 1981, MNRAS, 194, 809

Latif M. A., Niemeyer J. C., Schleicher D. R. G., 2014, MNRAS, 440, 2969

Lenoble R., Commerçon B., Rosdahl J., 2024, A&A, 685, A7

Li H., Vogelsberger M., Marinacci F., Gnedin O. Y., 2019, MNRAS, 487, 364

Lithwick Y., Dalal N., 2011, ApJ, 734, 100

Liu R.-Y., Wang X.-Y., 2011, ApJ, 730, 1

Mac Low M.-M., Klessen R. S., 2004, Reviews of Modern Physics, 76, 125

Machida M. N., Doi K., 2013, MNRAS, 435, 3283

Machida M. N., Matsumoto T., Inutsuka S.-i., 2008, ApJ, 685, 690

Madau P., Ferrara A., Rees M. J., 2001, ApJ, 555, 92

Magg M., et al., 2022, Monthly Notices of the Royal Astronomical Society, 514, 4433

Maio U., Koopmans L. V. E., Ciardi B., 2011, MNRAS, 412, L40

Mandelker N., van Dokkum P. G., Brodie J. P., van den Bosch F. C., Ceverino D., 2018, ApJ, 861, 148

Marchant P., Langer N., Podsiadlowski P., Tauris T. M., Moriya T. J., 2016, A&A, 588, A50

Marinacci F., Sales L. V., Vogelsberger M., Torrey P., Springel V., 2019, Monthly Notices of the Royal Astronomical Society, 489, 4233

Mashchenko S., Sills A., 2005a, ApJ, 619, 243

Mashchenko S., Sills A., 2005b, ApJ, 619, 258

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

McLaughlin D. E., Fall S. M., 2008, ApJ, 679, 1272

McQuinn M., O'Leary R. M., 2012, ApJ, 760, 3

Mebane R. H., Mirocha J., Furlanetto S. R., 2020, Monthly Notices of the Royal Astronomical Society, 493, 1217

Menon S. H., Lancaster L., Burkhart B., Somerville R. S., Dekel A., Krumholz M. R., 2024, arXiv e-prints, p. arXiv:2405.00813

Messa M., et al., 2024, arXiv e-prints, p. arXiv:2407.20331

Mocz P., Burkhart B., 2018, MNRAS, 480, 3916

Mocz P., Burkhart B., Hernquist L., McKee C. F., Springel V., 2017, ApJ, 838, 40

Moore B., 1996, ApJ, 461, L13

Muniruzzaman A. N. M., 1957, Calcutta Statistical Association Bulletin, 7, 115

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

Nakazato Y., Chiaki G., Yoshida N., Naoz S., Lake W., Chiou Y. S., 2022, ApJ, 927, L12

Naoz S., Barkana R., 2005a, MNRAS, 362, 1047

Naoz S., Barkana R., 2005b, MNRAS, 362, 1047

Naoz S., Barkana R., 2007, MNRAS, 377, 667

Naoz S., Narayan R., 2013, Physical Review Letters, 111, 051303

Naoz S., Narayan R., 2014, ApJ, 791, L8

Naoz S., Noter S., Barkana R., 2006, MNRAS, 373, L98

Naoz S., Barkana R., Mesinger A., 2009, MNRAS, 399, 369

Naoz S., Yoshida N., Barkana R., 2011, MNRAS, 416, 232

Naoz S., Yoshida N., Gnedin N. Y., 2012, ApJ, 747, 128

Naoz S., Yoshida N., Gnedin N. Y., 2013, ApJ, 763, 27

O'Leary R. M., McQuinn M., 2012, ApJ, 760, 4

O'Leary R. M., Rasio F. A., Fregeau J. M., Ivanova N., O'Shaughnessy R., 2006, ApJ, 637, 937

O'Leary R. M., Kocsis B., Loeb A., 2009, MNRAS, 395, 2127

O'Leary R. M., Meiron Y., Kocsis B., 2016, ApJ, 824, L12

Odenkirchen M., et al., 2003, AJ, 126, 2385

Omukai K., Hosokawa T., Yoshida N., 2010, ApJ, 722, 1793

O'Leary R. M., Rasio F. A., Fregeau J. M., Ivanova N., O'Shaughnessy R., 2006, The Astrophysical Journal, 637, 937–951

Padoan P., Nordlund Å., 1999, ApJ, 526, 279

Padoan P., Nordlund Å., 2011, ApJ, 730, 40

Park H., Ahn K., Yoshida N., Hirano S., 2020, ApJ, 900, 30

Pascale M., et al., 2022, arXiv e-prints, p. arXiv:2207.07102

Payne E., Banagiri S., Lasky P. D., Thrane E., 2020, Phys. Rev. D, 102, 102004

Peebles P. J. E., 1969, ApJ, 155, 393

Peebles P. J. E., 1984, ApJ, 277, 470

Pelletier G., Pudritz R. E., 1992, ApJ, 394, 117

Pérez-González P. G., et al., 2023, ApJ, 951, L1

Planck Collaboration et al., 2014, A&A, 571, A16

Pontoppidan K. M., et al., 2016, in Peck A. B., Seaman R. L., Benn C. R., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9910, Observatory Operations: Strategies, Processes, and Systems VI. p. 991016 (`arXiv:1707.02202`), doi:10.1117/12.2231768

Popa C., Naoz S., Marinacci F., Vogelsberger M., 2016, MNRAS, 460, 1625

Portegies Zwart S. F., McMillan S. L. W., 2000, The Astrophysical Journal, 528, L17–L20

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

Prole L. R., Clark P. C., Klessen R. S., Glover S. C. O., Pakmor R., 2022, MNRAS, 516, 2223

Prole L. R., Schauer A. T. P., Clark P. C., Glover S. C. O., Priestley F. D., Klessen R. S., 2023, MNRAS, 520, 2081

Punturo M., et al., 2010, Classical and Quantum Gravity, 27, 194002

Richardson M. L. A., Scannapieco E., Thacker R. J., 2013, ApJ, 771, 81

Robertson B., et al., 2023, arXiv e-prints, p. arXiv:2312.10033

Rodriguez C. L., Morscher M., Pattabiraman B., Chatterjee S., Haster C.-J., Rasio F. A., 2015, Phys. Rev. Lett., 115, 051101

Rodriguez C. L., Chatterjee S., Rasio F. A., 2016a, Phys. Rev. D, 93, 084029

Rodriguez C. L., Morscher M., Wang L., Chatterjee S., Rasio F. A., Spurzem R., 2016b, MNRAS, 463, 2109

Rodriguez C. L., Amaro-Seoane P., Chatterjee S., Kremer K., Rasio F. A., Samsing J., Ye C. S., Zevin M., 2018, Phys. Rev. D, 98, 123005

Rodriguez C. L., Kremer K., Chatterjee S., Fragione G., Loeb A., Rasio F. A., Weatherford N. C., Ye C. S., 2021, Research Notes of the American Astronomical Society, 5, 19

Rosolowsky E., 2005, PASP, 117, 1403

Sadanari K. E., Omukai K., Sugimura K., Matsumoto T., Tomida K., 2021, MNRAS, 505, 4197

Sadanari K. E., Omukai K., Sugimura K., Matsumoto T., Tomida K., 2024, PASJ, 76, 823

Saitoh T. R., Koda J., Okamoto T., Wada K., Habe A., 2006, ApJ, 640, 22

Sameie O., et al., 2022, arXiv e-prints, p. arXiv:2204.00638

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

Schauer A. T. P., Regan J., Glover S. C. O., Klessen R. S., 2017, MNRAS, 471, 4878

Schauer A. T. P., Glover S. C. O., Klessen R. S., Ceverino D., 2019, MNRAS, 484, 3510

Schauer A. T. P., Bromm V., Boylan-Kolchin M., Glover S. C. O., Klessen R. S., 2021, ApJ, 922, 193

Schauer A. T. P., Boylan-Kolchin M., Colston K., Sameie O., Bromm V., Bullock J. S., Wetzel A., 2023, ApJ, 950, 20

Schaye J., et al., 2015, MNRAS, 446, 521

Schmidt M., 1959, ApJ, 129, 243

Schneider A., Smith R. E., Macciò A. V., Moore B., 2012, MNRAS, 424, 684

Seljak U., Zaldarriaga M., 1996, ApJ, 469, 437

Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2016, ApJ, 826, 200

Senchyna P., Plat A., Stark D. P., Rudie G. C., 2023, arXiv e-prints, p. arXiv:2303.04179

Shapiro K. L., Genzel R., Förster Schreiber N. M., 2010, MNRAS, 403, L36

Sharda P., Menon S. H., 2024, arXiv e-prints, p. arXiv:2405.18265

Sharda P., Federrath C., Krumholz M. R., 2020, MNRAS, 497, 336

Sheth R. K., Tormen G., 1999, Monthly Notices of the Royal Astronomical Society, 308, 119–126

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Shu F. H., Lizano S., Ruden S. P., Najita J., 1988, ApJ, 328, L19

Sigurdsson S., Hernquist L., 1993, Nature, 364, 423

Smith B. D., et al., 2017, MNRAS, 466, 2217

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010, MNRAS, 401, 791

Stacy A., Bromm V., 2013, MNRAS, 433, 1094

Stacy A., Bromm V., Loeb A., 2011, ApJ, 730, L1

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

Stephan A. P., et al., 2019, ApJ, 878, 58

Stone N. C., van Velzen S., 2016, ApJ, 825, L14

Sun G., Furlanetto S. R., 2016, MNRAS, 460, 417

Tanaka T. L., Li M., 2014, MNRAS, 439, 1092

Tanaka T. L., Li M., Haiman Z., 2013, MNRAS, 435, 3559

Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1

Toomre A., 1964, ApJ, 139, 1217

Topping M. W., Shapley A. E., Steidel C. C., Naoz S., Primack J. R., 2018, ApJ, 852, 134

Trenti M., Padoan P., Jimenez R., 2015, ApJ, 808, L35

Tseliakhovich D., Hirata C., 2010, Phys. Rev. D, 82, 083520

Tseliakhovich D., Barkana R., Hirata C. M., 2011, MNRAS, 418, 906

Uysal B., Hartwig T., 2022, arXiv e-prints, p. arXiv:2211.12838

Vanzella E., et al., 2024, arXiv e-prints, p. arXiv:2407.20327

Visbal E., Barkana R., Fialkov A., Tseliakhovich D., Hirata C. M., 2012, Nature, 487, 70

Visbal E., Haiman Z., Bryan G. L., 2016, Monthly Notices of the Royal Astronomical Society: Letters, 460, L59

Vogelsberger M., White S. D. M., 2011, MNRAS, 413, 1419

Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42

Walcher C. J., et al., 2005, ApJ, 618, 237

Walcher C. J., Böker T., Charlot S., Ho L. C., Rix H.-W., Rossa J., Shields J. C., van der Marel R. P., 2006, ApJ, 649, 692

Wang H., Stephan A. P., Naoz S., Hoang B.-M., Breivik K., 2021, ApJ, 917, 76

Welch B., et al., 2022, Nature, 603, 815

Wen L., 2003, ApJ, 598, 419

Whalen D. J., Fryer C. L., Holz D. E., Heger A., Woosley S. E., Stiavelli M., Even W., Frey L. H., 2012, The Astrophysical Journal Letters, 762, L6

Whitmore B. C., Schweizer F., 1995, AJ, 109, 960

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

Willems B., Henninger M., Levin T., Ivanova N., Kalogera V., McGhee K., Timmes F. X., Fryer C. L., 2005, ApJ, 625, 324

Williams C. E., et al., 2023, ApJ, 945, 6

Williams C. E., et al., 2024a, ApJ, 960, L16

Williams C. E., et al., 2024b, ApJ, 960, L16

Willott C. J., et al., 2023, arXiv e-prints, p. arXiv:2311.12234

Wollenberg K. M. J., Glover S. C. O., Clark P. C., Klessen R. S., 2020, MNRAS, 494, 1871

Wong T.-W., Valsecchi F., Fragos T., Kalogera V., 2012, ApJ, 747, 111

Xu H., Norman M. L., O'Shea B. W., Wise J. H., 2016, ApJ, 823, 140

Yoshida N., Sugiyama N., Hernquist L., 2003a, Monthly Notices of the Royal Astronomical Society, 344, 481

Yoshida N., Abel T., Hernquist L., Sugiyama N., 2003b, ApJ, 592, 645

Yoshida N., Omukai K., Hernquist L., 2008, Science, 321, 669

Zel'dovich Y. B., 1970, A&A, 5, 84

de Mink S. E., Mandel I., 2016, MNRAS, 460, 3545

van Donkelaar F., Mayer L., Capelo P. R., Tamfal T., Quinn T. R., Madau P., 2023, Monthly Notices of the Royal Astronomical Society, 522, 1726