



The Supersonic Project: SIGOs, A Proposed Progenitor to Globular Clusters, and Their Connections to Gravitational-wave Anisotropies

William Lake^{1,2} , Smadar Naoz^{1,2} , Yeou S. Chiou^{1,2} , Blakesley Burkhart^{3,4} , Federico Marinacci⁵ ,
Mark Vogelsberger⁶ , and Kyle Kremer^{7,8}

¹ Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095, USA; wlake@astro.ucla.edu

² Mani L. Bhaumik Institute for Theoretical Physics, Department of Physics and Astronomy, UCLA, Los Angeles, CA 90095, USA

³ Department of Physics and Astronomy, Rutgers, The State University of New Jersey, 136 Frelinghuysen Rd., Piscataway, NJ 08854, USA

⁴ Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

⁵ Department of Physics & Astronomy “Augusto Righi”, University of Bologna, via Gobetti 93/2, I-40129 Bologna, Italy

⁶ Department of Physics and Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁷ TAPIR, California Institute of Technology, Pasadena, CA 91125, USA

⁸ The Observatories of the Carnegie Institution for Science, Pasadena, CA 91101, USA

Received 2021 May 3; revised 2021 August 19; accepted 2021 August 23; published 2021 November 23

Abstract

Supersonically induced gas objects (SIGOs), are structures with little to no dark-matter component predicted to exist in regions of the universe with large relative velocities between baryons and dark matter at the time of recombination. They have been suggested to be the progenitors of present-day globular clusters. Using simulations, SIGOs have been studied on small scales (around 2 Mpc) where these relative velocities are coherent. However, it is challenging to study SIGOs using simulations on large scales due to the varying relative velocities at scales larger than a few Mpc. Here, we study SIGO abundances semi-analytically: using perturbation theory, we predict the number density of SIGOs analytically, and compare these results to small-box numerical simulations. We use the agreement between the numerical and analytic calculations to extrapolate the large-scale variation of SIGO abundances over different stream velocities. As a result, we predict similar large-scale variations of objects with high gas densities before reionization that could possibly be observed by JWST. If indeed SIGOs are progenitors of globular clusters, then we expect a similar variation of globular cluster abundances over large scales. Significantly, we find that the expected number density of SIGOs is consistent with observed globular cluster number densities. As a proof-of-concept, and because globular clusters were proposed to be natural formation sites for gravitational wave sources from binary black-hole mergers, we show that SIGOs should imprint an anisotropy on the gravitational wave signal on the sky, consistent with their distribution.

Unified Astronomy Thesaurus concepts: Globular star clusters (656); High-redshift galaxies (734); Gravitational waves (678); Cosmology (343); Galactic and extragalactic astronomy (563); Galaxy formation (595)

1. Introduction

Globular clusters (GCs) are very old (~ 13 Gyr, e.g., Trenti et al. 2015) structures with masses between $\sim 10^5$ and $10^6 M_\odot$ (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 disks (e.g., Elmegreen 2010; Shapiro et al. 2010; Kruijssen 2015), for example, as a result of strong shocks when gas was 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., Ashman & Zepf 1992; Harris & Pudritz 1994;

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 2005; 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 2005).

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 (Tselikhovich & Hirata 2010; Tselikhovich et al. 2011). In the standard model of structure formation, due to 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

five orders of magnitude smaller than dark-matter overdensities (e.g., Naoz & Barkana 2005). 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 $\sim 2\text{--}3 \text{ Mpc}$, allowing it to be modeled 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; McQuinn & O’Leary 2012; Visbal et al. 2012; Cain et al. 2020), enhancing primordial black-hole formation (e.g., Tanaka et al. 2013; Latif et al. 2014; Tanaka & Li 2014; Hirano et al. 2017; 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 (Naoz et al. 2012; Bovy & Dvorkin 2013; Tanaka et al. 2013; O’Leary & McQuinn 2012; Tanaka & Li 2014; Asaba et al. 2016; Fialkov et al. 2012; Tseliakhovich & Hirata 2010; Maio et al. 2011), as well as the overall gas fraction in halos (Dalal et al. 2010; Greif et al. 2011; Maio et al. 2011; Tseliakhovich et al. 2011; Fialkov et al. 2012; Naoz et al. 2012; O’Leary & McQuinn 2012; Richardson et al. 2013; Naoz et al. 2013; Asaba et al. 2016) 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 (Greif et al. 2011; Liu & Wang 2011; Maio et al. 2011; Fialkov et al. 2012; O’Leary & McQuinn 2012; Richardson et al. 2013; Druschke et al. 2020), and the halo mass threshold at which star formation occurs (Greif et al. 2011; Liu & Wang 2011; Maio et al. 2011; Fialkov et al. 2012; O’Leary & McQuinn 2012; Bovy & Dvorkin 2013; 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 yet been studied. 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, 2017a). 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; Stone & van Velzen 2016; Chatterjee et al. 2017; 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 for the following reasons:

1. 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.
2. Even in the event of successfully implementing initial conditions that allow for the stream velocity to change coherently over large scales (\gg 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 Λ CDM cosmology with $\Omega_{\Lambda} = 0.73$, $\Omega_M = 0.27$, $\Omega_B = 0.044$, $\sigma_8 = 1.7$, and $h = 0.71$.

This paper is organized as follows: we first provide an overview of our simulations in Section 2.1. We then discuss our analytic model in Section 2.2. We provide a comparison between the simulation and model results in Section 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 4. We discuss our model results in Section 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. Methods

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

2.1. Simulations

We present three simulations with the moving-mesh code AREPO (Springel 2010) in a 2 Mpc box⁹ with 512^3 DM particles of mass $M_{\text{DM}} = 1.9 \times 10^3 M_{\odot}$ and 512^3 Voronoi mesh cells with $M_b = 360 M_{\odot}$, evolved from $z = 200$ to $z = 20$. These runs had stream velocities of $v_{\text{bc}} = 1\sigma_{\text{vbc}}$, $2\sigma_{\text{vbc}}$, and $3\sigma_{\text{vbc}}$ where σ_{vbc} is the rms value of the stream velocity—the relative velocity of the gas component with respect to the dark matter component. $\sigma_{\text{vbc}} = 5.9 \text{ km sec}^{-1}$ 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 & Barkana 2005). 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,¹⁰ 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 must 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 40%¹¹ and are outside the virial radius of the nearest dark-matter halo.

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_{bc} is approximately constant. We assigned a value of v_{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 (2005).

Stream velocities follow a Maxwell distribution with scale parameter $\sigma = \sigma_{\text{vbc}}/\sqrt{3}$, 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_{vbc} as the input power spectrum P_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_F of that field.
3. If this power spectrum output was consistent with the target output, we accepted this Maxwell-distributed field as our velocity grid. Otherwise, we set our input power spectrum

$$P_{I,\text{new}} = \frac{P_F(k)}{P_I(k)} \times P_{\text{vbc}}(k), \quad (1)$$

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:

$$\begin{aligned} \ddot{\delta}_{\text{dm}} + 2H\dot{\delta}_{\text{dm}} - f_{\text{dm}} \frac{2i}{a} \mathbf{v}_{\text{bc}} \cdot \mathbf{k} \dot{\delta}_{\text{dm}} \\ = \frac{3}{2} H_0^2 \frac{\Omega_m}{a^3} (f_b \delta_b + f_{\text{dm}} \delta_{\text{dm}}) + \left(\frac{\mathbf{v}_{\text{bc}} \cdot \mathbf{k}}{a} \right)^2 \delta_{\text{dm}}, \end{aligned} \quad (2)$$

⁹ 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).

¹⁰ This linking length was shown to give converging values of object abundances by Naoz et al. (2011).

¹¹ 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.

$$\ddot{\delta}_b + 2H\dot{\delta}_b = \frac{3}{2}H_0^2 \frac{\Omega_m}{a^3} (f_b \delta_b + f_{dm} \delta_{dm}) - \frac{k^2 k_b \bar{T}}{a^2 \mu} (\delta_b + \delta_T), \quad (3)$$

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 2005) in our calculations. These evolve according to

$$\frac{\delta_T}{dt} = \frac{2}{3} \frac{\delta_b}{dt} + \frac{x_e(t)}{t_\gamma} a^{-4} \left\{ \delta_\gamma \left(\frac{\bar{T}_\gamma}{\bar{T}} - 1 \right) + \frac{\bar{T}_\gamma}{\bar{T}} (\delta_{T_\gamma} - \delta_T) \right\}. \quad (4)$$

As was discussed in Naoz et al. (2012) and Naoz & Narayan (2014), the stream velocity introduces 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.¹² The comoving number density of halos of mass M at redshift z in this model is given by

$$\frac{dn}{dM} = \frac{\rho_0}{M} f_{ST} \left| \frac{dS}{dM} \right|, \quad (5)$$

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 : $\text{mathop="5pt"} \text{mathbottom="5pt"}$

$$f_{ST}(\delta_c, S) = A' \nu \frac{\sqrt{a'}}{S} \sqrt{\frac{a'}{2\pi}} \left[1 + \frac{1}{(a'\nu^2)^{q'}} \right] \exp\left(\frac{-a'\nu^2}{2}\right) \quad \text{and } \nu = \frac{\delta_c}{\sigma} = \frac{\delta_c}{\sqrt{S}}. \quad (6)$$

We use best-fit parameters $A' = 0.75$ and $q' = 0.3$ (Sheth & Tormen 2002). By evolving the power spectrum analytically,

¹² 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.

we can use this model to effectively predict halo abundances, as shown in Section 3.

It is 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 because of 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 formation of SIGOs. Thus, we postulate that the SIGOs' overdensity may be related to the DM underdensity and could obey a simple relation such as

$$\frac{dN(v_{bc})}{dM} \Big|_{\text{SIGO}} \propto \left(\frac{dN(v_{bc}=0)}{dM} \Big|_{\text{Halo}} - \frac{dN(v_{bc})}{dM} \Big|_{\text{Halo}} \right), \quad (7)$$

where the proportionality here is aimed to emphasize that the nonlinear 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 v_{bc} (Equation (A1)).

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.

3. Comparison between Analytical and Numerical Calculations

3.1. Dark-matter Halos

In Figure 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 1, present only analytical calculations. We show the fraction of DM number density $N(>M)_{\text{Halo}}$ with the stream velocity compared to the number density without the stream velocity. In other words,

$$f_{\text{DM}} = \frac{N(>M)_{\text{Halo}, v_{bc}}}{N(>M)_{\text{Halo}, 0}}, \quad (8)$$

where the subscript “0” means $v_{bc} = 0$ and the cumulative comoving number density of DM halos $N(>M)$ is given by

$$N(>M)_{\text{Halo}} = \int_M^\infty \frac{dN_{\text{Halo}}}{dM} dM. \quad (9)$$

The bottom left panel of Figure 1 shows Equation (8) as a function of the DM halo mass for different stream velocity effects. The bottom right panel shows Equation (8) as a function of v_{bc} , for different DM halo masses. As depicted, the higher stream velocities reduce the abundance of dark matter

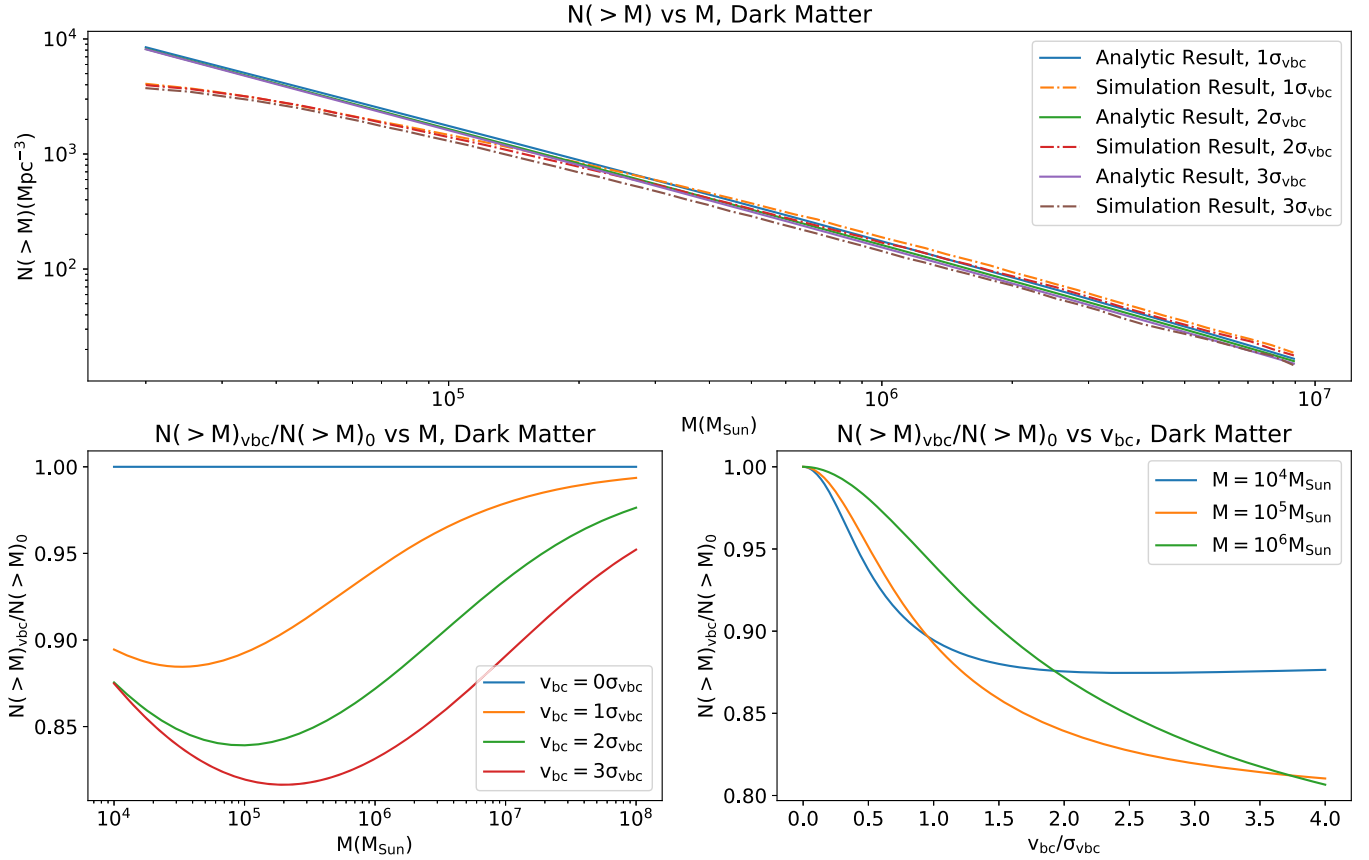


Figure 1. The effect of the stream velocity on “classical” dark-matter halos. In the top panel, we consider the number abundance $N(>M)_{\text{Halo},v_{\text{bc}}}$ as a function of the DM halo mass, and compare the analytical calculation (solid lines) and simulation results (dotted–dashed lines). In the bottom panels, we consider the analytical reduction fraction of the number of halos due to the stream velocity (i.e., $N(>M)_{\text{Halo},v_{\text{bc}}}/N(>M)_{\text{Halo},0}$, where subscript “0” indicates $v_{\text{bc}} = 0$). 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 $z = 20$.

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 σ_{vbc} being on the order of tens of percent.

In Figure 2 we depict the DM halo fluctuations due to the stream velocity effects on a large scale. In particular, we show

$$\delta_{\text{Halo}} = \frac{N(>M)_{\text{Halo},v_{\text{bc}}} - \langle N(>M)_{\text{Halo}} \rangle}{\langle N(>M)_{\text{Halo}} \rangle}, \quad (10)$$

with $N(>M)_{\text{Halo}}$ as defined in Equation (9), and using the velocity field generated with the method described in Section 2.2.

3.2. SIGOs

Using the tightly fitted ellipsoid method to find the SIGOs (see Section 2) we find 8, 75, and 188 SIGOs for the 1, 2, and $3\sigma_{\text{vbc}}$ simulation runs. From this simulation data, we construct $N(>M)_{\text{SIGO},\text{sim}}$. As discussed in Section 2.2, we also calculate the abundance of SIGOs as a function σ_{vbc} analytically using Equation (7). 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 a function of mass and stream velocity, i.e.,

$$\left. \frac{dN(u_{\text{vbc}})}{dM} \right|_{\text{SIGO}} \approx \left(\frac{u_{\text{vbc}}}{20.76} \right)^{2.43} \times \left(\left. \frac{dN(u_{\text{vbc}}=0)}{dM} \right|_{\text{Halo}} - \left. \frac{dN(u_{\text{vbc}})}{dM} \right|_{\text{Halo}} \right), \quad (11)$$

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

Figure 3, top panel, shows the comparison between the SIGOs’ model and simulation results (as in Figure 1). In the analytical model, we have integrated the number density dN/dM (see Equation [11]) only from a defined cutoff mass ($M_{\text{cutoff}} = M_{\text{SIGO},\text{max}} = 1.1 \times 10^6 M_{\odot}$) 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 $\times 10^6 M_{\odot}$, above which they are incapable of escaping their parent DM halo. Because of this, we select M_{cutoff} as the mass of the largest SIGO observed in any of our simulations. In particular, the cumulative number density can be expressed as

$$N(>M_{\text{min}}) = \int_{M_{\text{min}}}^{M_{\text{SIGO},\text{max}}} \left(\frac{dN}{dM} \right)_{\text{SIGO}} dM, \quad (12)$$

where dN/dM is defined in Equation (11). We show the analytical calculations based on this equation in Figure 3, solid lines, in all panels. From top to bottom, we consider 1, 2, and $3\sigma_{\text{vbc}}$ 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 SIGO abundances at the range of a few $\times 10^4$ –few $\times 10^5 M_{\odot}$.

At small masses ($M \approx 10^4 M_{\odot}$), the limited resolution of the simulation yields lower abundances. At large masses, Poisson

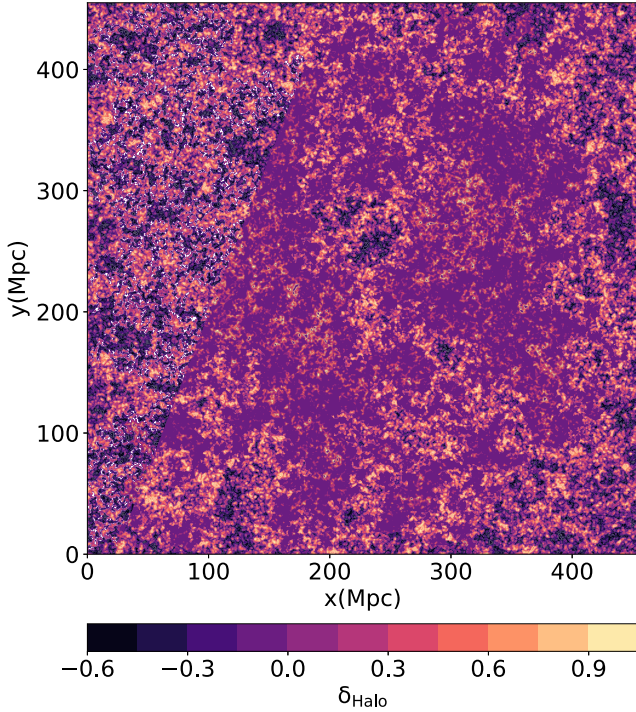


Figure 2. Map of DM halo density contrast at $z=20$, showing a $456 \times 456 \times 1$ Mpc box (a subset of our model box, to enhance the visibility of structures).

fluctuations increase the uncertainty of our simulation results, again creating an apparent disagreement between our simulation results and analytic model.¹³ 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 3 show analytical calculations only, analogous to those of Figure 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 (12) 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_{SIGO} and σ_{vbc} that fits the analytical model.

$$\log_{10}\left(\frac{N(>M_{\text{SIGO}})}{\text{Mpc}^3}\right) = a(v_{\text{bc}}) \left\{ \log_{10}\left(\frac{M_{\text{SIGO}}}{M_{\odot}}\right) \right\}^{b(v_{\text{bc}})} + c(v_{\text{bc}}). \quad (13)$$

Fits for the prefactors $a(v_{\text{bc}})$, $b(v_{\text{bc}})$, and $c(v_{\text{bc}})$ 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

$$f_{\text{SIGO}} = \frac{N(>M)_{\text{SIGO},\text{vbc}}}{\langle N(>M)_{\text{SIGO}} \rangle}, \quad (14)$$

¹³ For example, the largest such fluctuation, for $M \sim \text{few} \times 10^5 M_{\odot}$, is about a 3σ deviation in the $2\sigma_{\text{vbc}}$ data. This represents a 1% fluctuation, which is consistent with the largest fluctuations due to Poisson statistics we would expect in our data set.

with

$$\langle N(>M) \rangle = \int_0^{\infty} p(v_{\text{bc}}) N(>M)_{\text{SIGO},\text{vbc}} dv_{\text{bc}}, \quad (15)$$

where $p(v_{\text{bc}})$ is the likelihood of a given stream velocity given by a Maxwell distribution, and $\langle N(>M)_{\text{SIGO}} \rangle$ 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 1 because SIGOs are only found when the stream velocity effect is present. As shown in Figure 3, left panel, large values of the stream velocity ($\sim 3.5\sigma_{\text{vbc}}$) result in an enhancement in SIGO abundances on the order of 20–30 \times 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 $3.5\sigma_{\text{vbc}}$. 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 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

$$\delta_{\text{SIGO}} = \frac{N(>M)_{\text{vbc}} - \langle N(>M) \rangle}{\langle N(>M) \rangle}, \quad (16)$$

with $N(>M)$ as defined in Equation (12). The overdensities in the plot were generated using the method for generating appropriately distributed density fields outlined in Section 2.2.

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

$$\langle N_{k,\text{SIGO}} N_{k',\text{SIGO}}^* \rangle = (2\pi)^3 P(k) \delta^{(3)}(\mathbf{k} - \mathbf{k}'). \quad (17)$$

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 $k > 0.5 \text{ Mpc}^{-1}$, implying that number densities are coherent on scales of a few comoving Mpc.

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

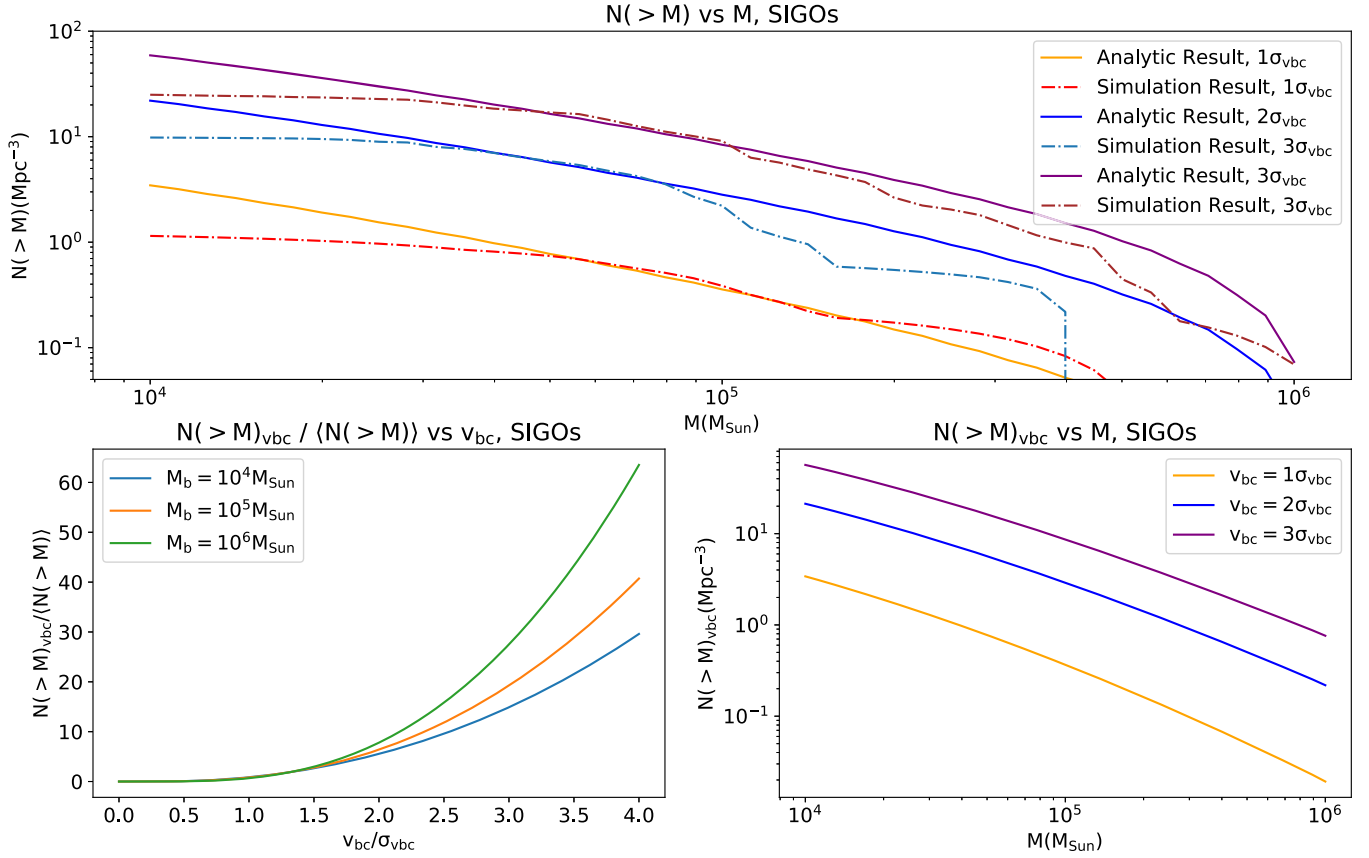


Figure 3. The effect of the stream velocity on SIGO abundances. In the top panel, we consider the number abundance $N(>M)_{\text{SIGO},v_{\text{bc}}}$ as a function of SIGO mass and compare the analytical calculation (solid lines) and simulation results (dotted–dashed lines). See the text for an explanation of how SIGO abundances are estimated in the analytical calculations and how they are defined in the simulations, and see Equation (15) for a definition of $\langle N(>M)_{\text{SIGO}} \rangle$. 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 $z = 20$.

probability density function:

$$\log_{10}[p(N_{\text{SIGO}})] \approx -0.933 \times \exp[1.734 \times \log_{10}(N_{\text{SIGO}})] + 0.261 \times \exp[-0.507 \times \log_{10}(N_{\text{SIGO}})] \quad (18)$$

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

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 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 $\sim 0.5 \text{ Mpc}^{-3}$ 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} (Rodríguez et al. 2015) to a few $\times \text{Mpc}^{-3}$ (e.g., Portegies Zwart & McMillan 2000; Harris et al. 2013). Since SIGOs are early structures, they are likely to have low metallicities. Rodríguez 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

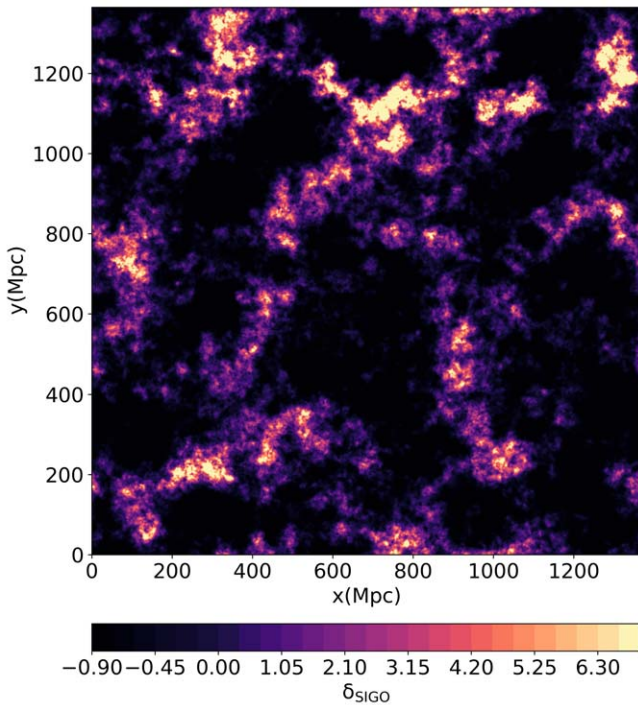


Figure 4. Map of SIGO density contrast at $z = 20$. Specifically, we plot Equation (16) for a minimum mass of $10^3 M_\odot$. Here the color scale was capped at $\delta_{\text{SIGO}} = 7.0$ to enhance visibility of smaller fluctuations; the true maximum in this plot is $\delta_{\text{SIGO}} \approx 30$ (which represents a rare fluctuation).

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, modeled as a puffy disk, has a Toomre stability criterion above unity (Toomre 1964), and therefore will not fragment (except for 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. 2021), 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.

4. Implications for 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 (Willems et al. 2005; Belczynski et al. 2006; Wong et al. 2012). On a timescale of $\lesssim 1$ 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 mergers¹⁴ (Portegies Zwart & McMillan 2000; O’Leary et al. 2006; Rodriguez et al. 2015; Rodriguez et al. 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., Fragione & Kocsis 2018; Rodriguez et al. 2018; Antonini & Gieles 2020; Kremer et al. 2020). Assuming that these mergers are roughly uniformly distributed in time at zeroth order, this implies a BBH merger rate of roughly 10^{-8} 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 anisotropy in BBH mergers from GCs.

As a proof-of-concept, in Figure 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). 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 $\sim 10^\circ$. To a distance of 675 Mpc, we also estimate an approximate rate of BBH mergers of $0.5 \text{ mergers sr}^{-1} \text{ yr}^{-1}$, with a standard deviation of approximately $0.3 \text{ mergers sr}^{-1} \text{ yr}^{-1}$ across the sky. Payne et al. (2020) found that 10 black-hole mergers from the first LIGO/Virgo catalog are consistent with an isotropic distribution over large scales (> 100 Mpc). Further, recent endeavors by The LIGO Scientific Collaboration et al. (2021) 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 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. 2021) 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. 2017b), this anisotropy may potentially be observable and if observed, would further constrain the connection between GCs and SIGOs.

We stress, however, that Figure 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

¹⁴ However, other processes have been suggested to be comparable, from isolating binaries (e.g., Belczynski et al. 2016; de Mink & Mandel 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).

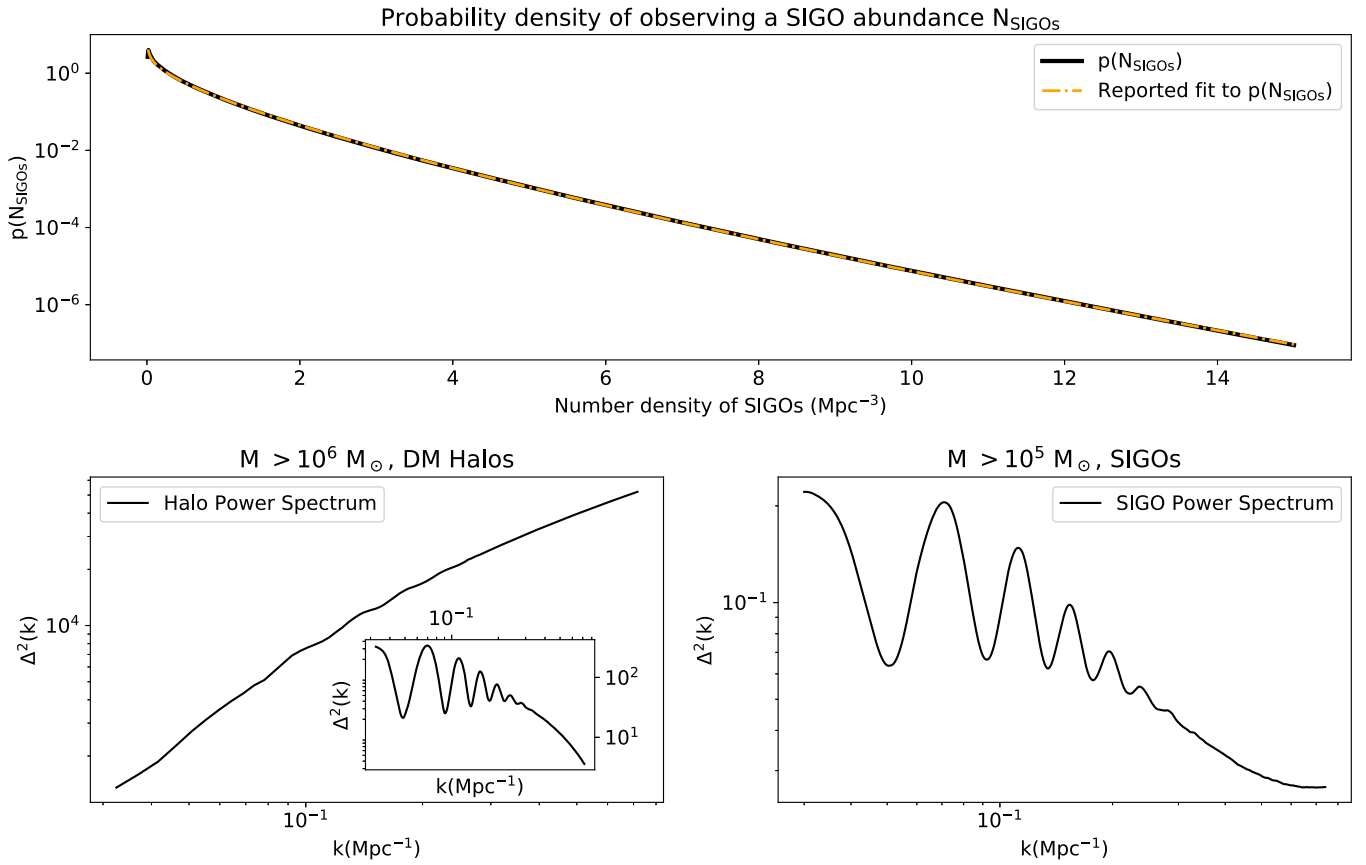


Figure 5. The distribution of SIGOs in the universe. In the top panel, we show the probability density $p(N_{\text{SIGO}})$ of observing a given abundance of SIGOs with $M > 10^5 M_{\odot}$ in a randomly selected region of the universe with an unknown (but constant) stream velocity. We also show a fit to this probability density function given by Equation (18). In the bottom left panel, we present the power spectrum of the distribution of $M_{\text{Halo}} > 10^6 M_{\odot}$ 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 $N(>M)_{\text{Halo}}$ 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.

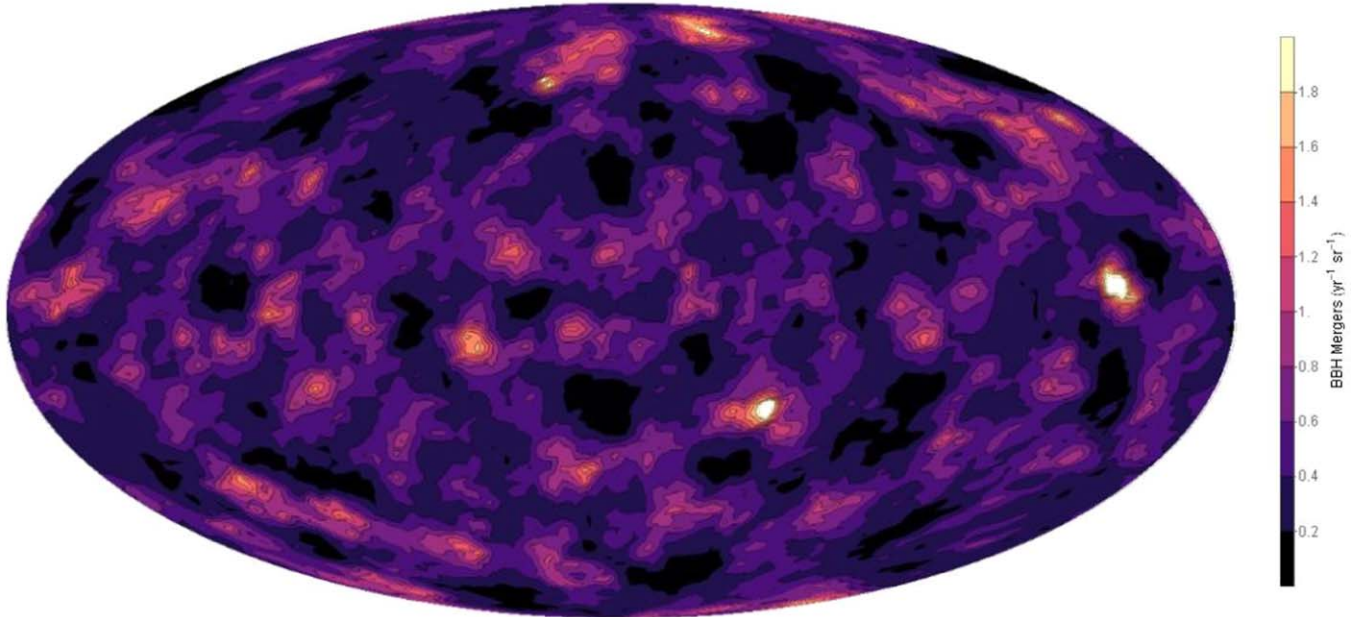


Figure 6. Sky map of integrated BBH merger abundances in globular clusters to a distance of 675 Mpc. Numbers are given in mergers per steradian per year, assuming a merger rate of 10^{-8} per year per cluster.

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.¹⁵

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 $\lesssim \text{few} \times 10^6 M_\odot$ (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; Tselikhovich & Hirata 2010). This stream velocity is coherent only on small scales ($\sim \text{few Mpc}$), 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. In addition, the relatively small mass of SIGOs ($\sim \text{few} \times 10^5 - 10^6 M_\odot$) 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 1). We additionally show a comparison between our model results and simulations of SIGO formation and abundance (Figure 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 $M < \text{few} \times 10^7 M_\odot$ in both our analytic model and simulations, consistent with previous studies (e.g., Tselikhovich & Hirata 2010; Naoz et al. 2012; Popa et al. 2016) and with each other (Figure 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). For a comparison more directly related to our universe, we also show a power spectrum of dark-matter halo abundances across the sky (Figure 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 stream velocity can be linked directly to the increase in SIGOs and we 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 3). The strong agreement between our analytical model and the small-box simulation results, as depicted in the top panel of Figure 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), and to create a power spectrum of SIGO abundances (Figure 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 $\sigma_{\text{SIGO}} = 0.6 \text{ Mpc}^{-3}$. 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 (18) and the top panel of Figure 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 5. The distribution of observed SIGOs should be qualitatively similar to the map presented in Figure 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 $\sim 0.5 \text{ Mpc}^{-3}$ 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 $\times \text{Mpc}^{-3}$ (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 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 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° , to a distance of 675 Mpc, in an idealized case (though we caution that that variation would decrease on longer sightlines). Future observations

¹⁵ 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.

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 $\sim 0.5 \text{ Mpc}^{-3}$, 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.

We thank the anonymous referee for useful comments. We also thank Naoki Yoshida and Carl Rodriguez for useful discussions. W.L., S.N., Y.S.C, B.B., F.M., and M.V. are grateful for the support of NASA grant No. 80NSSC20K0500 and the XSEDE AST180056 allocation, as well as the Simons Foundation Center for Computational Astrophysics and the UCLA cluster *Hoffman2* for computational resources. F.M. acknowledges support from the Program ‘‘Rita Levi Montalcini’’ of the Italian MUR. W.L. thanks Ryan Carlson for helpful conversations. S.N. thanks Howard and Astrid Preston for their generous support. Y.S.C. is thankful for partial support from a UCLA dissertation year fellowship. B.B. also thanks the Alfred P. Sloan Foundation and the Packard Foundation for support. M.V. acknowledges support through NASA ATP grants 16-ATP16-0167, 19-ATP19-0019, 19-ATP19-0020, 19-ATP19-0167, and NSF grants AST-1814053, AST-1814259, AST-1909831 and AST-2007355. K.K. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-2001751.

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_{bc} using Equation (7). For completeness, we show the equation again here, with terms labeled as coming from simulations or analytic models:

$$\left. \frac{dN(u_{vbc})}{dM} \right|_{\text{SIGO, Sim}} = A(u_{vbc}) \left(\left. \frac{dN(u_{vbc} = 0)}{dM} \right|_{\text{DM, Analytic}} - \left. \frac{dN(u_{vbc})}{dM} \right|_{\text{DM, Analytic}} \right) \quad (\text{A1})$$

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

$$A(u_{vbc}) = C \times u_{vbc}^n, \quad (\text{A2})$$

We use the Levenburg–Marquardt method to find the optimal value of A for each simulated stream velocity:

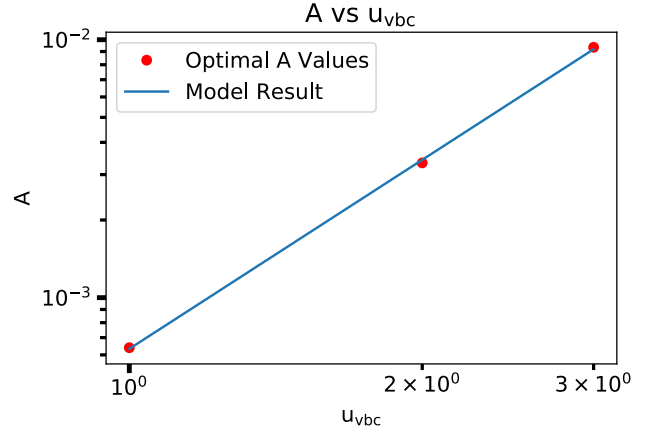


Figure 7. Plot of the agreement between our model normalization parameter and our optimal normalization parameters obtained from comparison with simulations, as a function of stream velocity.

Table 1
Optimal Values of Our Normalization Factor A with Error Estimates for Each of Our Simulated Stream Velocities

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}

$u_{vbc} = 1, 2,$ 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 7 as red points. They are also listed in Table 1, with error estimates δA :

Second, we fit the points for all three simulations for the different v_{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 7 as the blue line.

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 (12).

We found a form for this relation given by Equation (13), repeated here for convenience:

$$\begin{aligned} & \log_{10} \left(\frac{N(>M_{\text{SIGO}})}{\text{Mpc}^3} \right) \\ &= a(u_{vbc}) \left\{ \log_{10} \left(\frac{M_{\text{SIGO}}}{M_{\odot}} \right) \right\}^{b(u_{vbc})} + c(u_{vbc}), \end{aligned}$$

where

$$\begin{aligned} a(u_{vbc}) &= a_1 u_{vbc}^5 + a_2 u_{vbc}^4 + a_3 u_{vbc}^3 \\ &+ a_4 u_{vbc}^2 + a_5 u_{vbc} + a_6, \end{aligned} \quad (\text{B3})$$

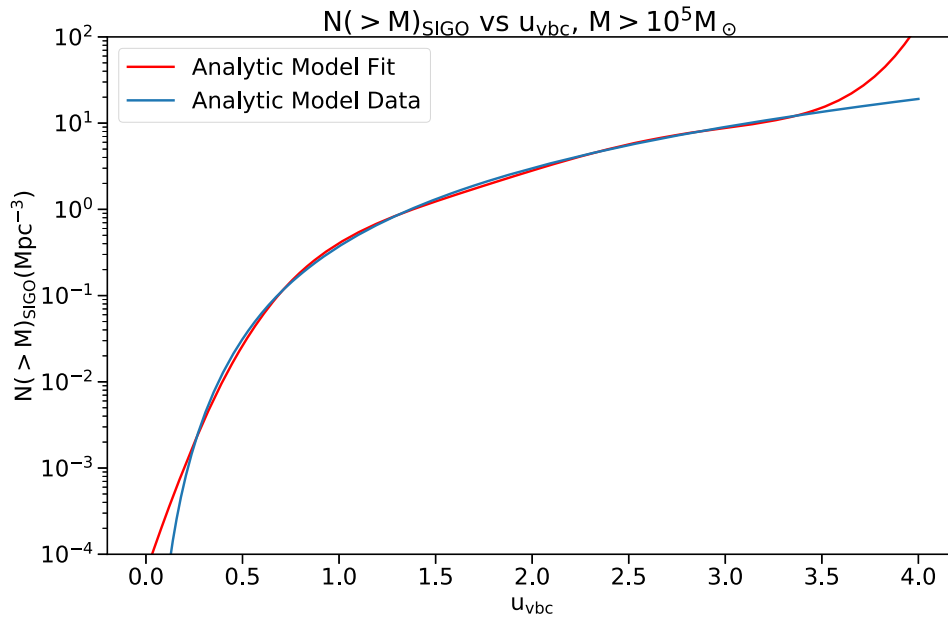


Figure 8. Plot illustrating the agreement between our analytic model for SIGO abundances (Equation (12)) and our reported fit to the model output (Equation (13)) at $M > 10^5 M_{\odot}$. The two equations agree closely for $\text{few} \times 10^4 M_{\odot} < M < 8 \times 10^5 M_{\odot}$, and for $u_{\text{vbc}} < 3.4$.

Table 2

Best-Fit Parameters Used in Equation (13) That Match Analytic Model Results from Equation (12), Obtained Using a Nonlinear Least-Squares Regression

	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

Note. Coefficients are given to high precision due to the sensitivity of the model to the coefficients.

$$b(u_{\text{vbc}}) = b_1 u_{\text{vbc}}^5 + b_2 u_{\text{vbc}}^4 + b_3 u_{\text{vbc}}^3 + b_4 u_{\text{vbc}}^2 + b_5 u_{\text{vbc}} + b_6, \quad (\text{B4})$$

and

$$c(u_{\text{vbc}}) = c_1 u_{\text{vbc}}^5 + c_2 u_{\text{vbc}}^4 + c_3 u_{\text{vbc}}^3 + c_4 u_{\text{vbc}}^2 + c_5 u_{\text{vbc}} + c_6. \quad (\text{B5})$$

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 2. An example of the agreement between our analytic model and the fit to the model is presented in Figure 8. The results of Equation (12) are shown in blue, giving $N(>M)_{\text{SIGO}}$ for $M > 10^5 M_{\odot}$, with the result of Equation (13) shown in red. Notably, this reported fit holds for $2.5 \times 10^4 M_{\odot} \lesssim M \lesssim 8 \times 10^5 M_{\odot}$, and for $v_{\text{bc}} < 3.4 \sigma_{\text{vbc}}$.

ORCID iDs

William Lake <https://orcid.org/0000-0002-4227-7919>
 Smadar Naoz <https://orcid.org/0000-0002-9802-9279>
 Yeou S. Chiou <https://orcid.org/0000-0003-4962-5768>
 Blakesley Burkhart <https://orcid.org/0000-0001-5817-5944>
 Federico Marinacci <https://orcid.org/0000-0003-3816-7028>
 Mark Vogelsberger <https://orcid.org/0000-0001-8593-7692>
 Kyle Kremer <https://orcid.org/0000-0002-4086-3180>

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *PhRvL*, **116**, 061102
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017a, *PhRvL*, **119**, 161101
 Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017b, *CQGrA*, **34**, 044001
 Abbott, R., Abbott, T. D., Abraham, S., et al. 2021, *PhRvX*, **11**, 021053
 Antonini, F., Barausse, E., & Silk, J. 2015, *ApJL*, **806**, L8
 Antonini, F., & Gieles, M. 2020, *PhRvD*, **102**, 123016
 Asaba, S., Ichiki, K., & Tashiro, H. 2016, *PhRvD*, **93**, 023518
 Ashman, K. M., & Zepf, S. E. 1992, *ApJ*, **384**, 50
 Barkana, R., & Loeb, A. 2004, *ApJ*, **609**, 474
 Baumgardt, H., & Kroupa, P. 2007, *MNRAS*, **380**, 1589
 Bekki, K., & Yong, D. 2012, *MNRAS*, **419**, 2063
 Belczynski, K., Holz, D. E., Bulik, T., & O’Shaughnessy, R. 2016, *Natur*, **534**, 512
 Belczynski, K., Sadowski, A., Rasio, F. A., & Bulik, T. 2006, *ApJ*, **650**, 303
 Böker, T., Sarzi, M., McLaughlin, D. E., et al. 2004, *AJ*, **127**, 105
 Bovy, J., & Dvorkin, C. 2013, *ApJ*, **768**, 70
 Bradford, J. D., Geha, M., Muñoz, R. R., et al. 2011, *ApJ*, **743**, 167
 Breivik, K., Chatterjee, S., & Andrews, J. J. 2019, *ApJL*, **878**, L4
 Breivik, K., Coughlin, S., Zevin, M., et al. 2020, *ApJ*, **898**, 71
 Bromm, V., & Clarke, C. J. 2002, *ApJL*, **566**, L1
 Brown, I. A. 2013, *JCAP*, **2013**, 003
 Brown, T. M., Tumlinson, J., Geha, M., et al. 2014, *ApJ*, **796**, 91
 Cain, C., D’Aloisio, A., Iršič, V., McQuinn, M., & Trac, H. 2020, *ApJ*, **898**, 168
 Chatterjee, S., Rodriguez, C. L., & Rasio, F. A. 2017, *ApJ*, **834**, 68
 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
 Chiou, Y. S., Naoz, S., Marinacci, F., & Vogelsberger, M. 2018, *MNRAS*, **481**, 3108
 Conroy, C., Loeb, A., & Spergel, D. N. 2011, *ApJ*, **741**, 72
 Dalal, N., Pen, U.-L., & Seljak, U. 2010, *JCAP*, **11**, 007
 de Mink, S. E., & Mandel, I. 2016, *MNRAS*, **460**, 3545
 Druschke, M., Schauer, A. T. P., Glover, S. C. O., & Klessen, R. S. 2020, *MNRAS*, **498**, 4839
 Elmegreen, B. G. 2010, *ApJL*, **712**, L184
 Elmegreen, B. G., & Efremov, Y. N. 1997, *ApJ*, **480**, 235
 Fall, S. M., & Zhang, Q. 2001, *ApJ*, **561**, 751
 Fialkov, A., Barkana, R., Tselikhovich, D., & Hirata, C. M. 2012, *MNRAS*, **424**, 1335
 Fragione, G., & Kocsis, B. 2018, *PhRvL*, **121**, 161103
 Greif, T. H., Springel, V., White, S. D. M., et al. 2011, *ApJ*, **737**, 75

- Grillmair, C. J., Freeman, K. C., Irwin, M., & Quinn, P. J. 1995, *AJ*, **109**, 2553
- Grudić, M. Y., Kruijssen, J. M. D., Faucher-Giguère, C.-A., et al. 2021, *MNRAS*, **506**, 3239
- Gunn, J. E. 1980, in *Globular Clusters*, ed. D. Hanes & B. Madore (Cambridge: Cambridge Univ. Press), 301
- Harris, W. E. 1996, *AJ*, **112**, 1487
- Harris, W. E., Harris, G. L. H., & Alessi, M. 2013, *ApJ*, **772**, 82
- Harris, W. E., & Pudritz, R. E. 1994, *ApJ*, **429**, 177
- Heggie, D. C., & Hut, P. 1996, in *IAU Symp.*, 174, *Dynamical Evolution of Star Clusters: Confrontation of Theory and Observations*, ed. P. Hut & J. Makino (Dordrecht: Kluwer Academic), 303
- Hirano, S., Hosokawa, T., Yoshida, N., & Kuiper, R. 2017, *Sci*, **357**, 1375
- Hoang, B.-M., Naoz, S., Kocsis, B., Rasio, F. A., & Dosopoulou, F. 2018, *ApJ*, **856**, 140
- Ibata, R., Nipoti, C., Sollima, A., et al. 2013, *MNRAS*, **428**, 3648
- Kaiser, N. 1984, *ApJL*, **284**, L9
- Kocsis, B., & Levin, J. 2012, *PhRvD*, **85**, 123005
- Kravtsov, A. V., & Gnedin, O. Y. 2005, *ApJ*, **623**, 650
- Kremer, K., Ye, C. S., Rui, N. Z., 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., & Bland-Hawthorn, J. 2019, *ARA&A*, **57**, 227
- Latif, M. A., Niemeyer, J. C., & Schleicher, D. R. G. 2014, *MNRAS*, **440**, 2969
- 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
- Maio, U., Koopmans, L. V. E., & Ciardi, B. 2011, *MNRAS*, **412**, L40
- Marchant, P., Langer, N., Podsiadlowski, P., Tauris, T. M., & Moriya, T. J. 2016, *A&A*, **588**, A50
- Mashchenko, S., & Sills, A. 2005, *ApJ*, **619**, 258
- McLaughlin, D. E., & Fall, S. M. 2008, *ApJ*, **679**, 1272
- McQuinn, M., & O’Leary, R. M. 2012, *ApJ*, **760**, 3
- Moore, B. 1996, *ApJL*, **461**, L13
- Muratov, A. L., & Gnedin, O. Y. 2010, *ApJ*, **718**, 1266
- Naoz, S., & Barkana, R. 2005, *MNRAS*, **362**, 1047
- Naoz, S., & Barkana, R. 2007, *MNRAS*, **377**, 667
- Naoz, S., Barkana, R., & Mesinger, A. 2009, *MNRAS*, **399**, 369
- Naoz, S., & Narayan, R. 2013, *PhRvL*, **111**, 051303
- Naoz, S., & Narayan, R. 2014, *ApJL*, **791**, L8
- Naoz, S., Noter, S., & Barkana, R. 2006, *MNRAS*, **373**, L98
- 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
- Odenkirchen, M., Grebel, E. K., Dehnen, W., et al. 2003, *AJ*, **126**, 2385
- O’Leary, R. M., Kocsis, B., & Loeb, A. 2009, *MNRAS*, **395**, 2127
- O’Leary, R. M., & McQuinn, M. 2012, *ApJ*, **760**, 4
- O’Leary, R. M., Meiron, Y., & Kocsis, B. 2016, *ApJL*, **824**, L12
- O’Leary, R. M., Rasio, F. A., Fregeau, J. M., Ivanova, N., & O’Shaughnessy, R. 2006, *ApJ*, **637**, 937
- O’Leary, R. M., Rasio, F. A., Fregeau, J. M., Ivanova, N., & O’Shaughnessy, R. 2006, *ApJ*, **637**, 937
- Park, H., Ahn, K., Yoshida, N., & Hirano, S. 2020, *ApJ*, **900**, 30
- Payne, E., Banagiri, S., Lasky, P. D., & Thrane, E. 2020, *PhRvD*, **102**, 102004
- Peebles, P. J. E. 1969, *ApJ*, **155**, 393
- Peebles, P. J. E. 1984, *ApJ*, **277**, 470
- Popa, C., Naoz, S., Marinacci, F., & Vogelsberger, M. 2016, *MNRAS*, **460**, 1625
- Portegies Zwart, S. F., & McMillan, S. L. W. 2000, *ApJ*, **528**, L17
- Press, W. H., & Schechter, P. 1974, *ApJ*, **187**, 425
- Punturo, M., Abernathy, M., Acernese, F., et al. 2010, *CQGra*, **27**, 194002
- Richardson, M. L. A., Scannapieco, E., & Thacker, R. J. 2013, *ApJ*, **771**, 81
- Rodriguez, C. L., Amaro-Seoane, P., Chatterjee, S., et al. 2018, *PhRvD*, **98**, 123005
- Rodriguez, C. L., Chatterjee, S., & Rasio, F. A. 2016a, *PhRvD*, **93**, 084029
- Rodriguez, C. L., Kremer, K., Chatterjee, S., et al. 2021, *RNAAS*, **5**, 19
- Rodriguez, C. L., Morscher, M., Pattabiraman, B., et al. 2015, *PhRvL*, **115**, 051101
- Rodriguez, C. L., Morscher, M., Wang, L., et al. 2016b, *MNRAS*, **463**, 2109
- Saitoh, T. R., Koda, J., Okamoto, T., Wada, K., & Habe, A. 2006, *ApJ*, **640**, 22
- Schauer, A. T. P., Glover, S. C. O., Klessen, R. S., & Ceverino, D. 2019, *MNRAS*, **484**, 3510
- Schauer, A. T. P., Regan, J., Glover, S. C. O., & Klessen, R. S. 2017, *MNRAS*, **471**, 4878
- Schneider, A., Smith, R. E., Macciò, A. V., & Moore, B. 2012, *MNRAS*, **424**, 684
- Seljak, U., & Zaldarriaga, M. 1996, *ApJ*, **469**, 437
- Shapiro, K. L., Genzel, R., & Förster Schreiber, N. M. 2010, *MNRAS*, **403**, L36
- Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *MNRAS*, **323**, 1
- Sheth, R. K., & Tormen, G. 1999, *MNRAS*, **308**, 119
- Sheth, R. K., & Tormen, G. 2002, *MNRAS*, **329**, 61
- Sigurdsson, S., & Hernquist, L. 1993, *Natur*, **364**, 423
- Springel, V. 2010, *MNRAS*, **401**, 791
- Stacy, A., Bromm, V., & Loeb, A. 2011, *ApJL*, **730**, L1
- Stephan, A. P., Naoz, S., Ghez, A. M., et al. 2019, *ApJ*, **878**, 58
- Stone, N. C., & van Velzen, S. 2016, *ApJL*, **825**, L14
- Tanaka, T. L., & Li, M. 2014, *MNRAS*, **439**, 1092
- Tanaka, T. L., Li, M., & Haiman, Z. 2013, *MNRAS*, **435**, 3559
- The LIGO Scientific CollaborationThe Virgo CollaborationThe KAGRA Collaboration et al. 2021, *PhRvD*, **104**, 022005
- 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, *ApJL*, **808**, L35
- Tselikhovich, D., Barkana, R., & Hirata, C. M. 2011, *MNRAS*, **418**, 906
- Tselikhovich, D., & Hirata, C. 2010, *PhRvD*, **82**, 083520
- Visbal, E., Barkana, R., Fialkov, A., Tselikhovich, D., & Hirata, C. M. 2012, *Natur*, **487**, 70
- Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, *NatRP*, **2**, 42
- Vogelsberger, M., & White, S. D. M. 2011, *MNRAS*, **413**, 1419
- Walcher, C. J., Böker, T., Charlot, S., et al. 2006, *ApJ*, **649**, 692
- Walcher, C. J., van der Marel, R. P., McLaughlin, D., et al. 2005, *ApJ*, **618**, 237
- Wang, H., Stephan, A. P., Naoz, S., Hoang, B.-M., & Breivik, K. 2021, *ApJ*, **917**, 76
- Wen, L. 2003, *ApJ*, **598**, 419
- Whitmore, B. C., & Schweizer, F. 1995, *AJ*, **109**, 960
- Whitmore, B. C., Zhang, Q., Leitherer, C., et al. 1999, *AJ*, **118**, 1551
- Willems, B., Henninger, M., Levin, T., et al. 2005, *ApJ*, **625**, 324
- Wong, T.-W., Valsecchi, F., Fragos, T., & Kalogera, V. 2012, *ApJ*, **747**, 111