Great balls of FIRE – I. The formation of star clusters across cosmic time in a Milky Way-mass galaxy

Michael Grudić, Zachary Hafen, Carl Rodriguez, Dávid Guszejnov, Astrid Lamberts, Andrew Wetzel, Michael Boylan-Kolchin, Claude-André Faucher-Giguère

Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **519,** 1366–1380 (2023) https://doi.org/10.1093/mnras/stac3573
Advance Access publication 2022 December 23

Great balls of FIRE – I. The formation of star clusters across cosmic time in a Milky Way-mass galaxy

Michael Y. Grudić [1★†] Zachary Hafen [2] Carl L. Rodriguez [3] Dávid Guszejnov [4] Astrid Lamberts [5,6] Andrew Wetzel [7] Michael Boylan-Kolchin [4] and Claude-André Faucher-Giguère [8]

[1]*Carnegie Observatories, 813 Santa Barbara St, Pasadena, CA 91101, USA*
[2]*Center for Cosmology, Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
[3]*McWilliams Center for Cosmology and Department of Physics, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[4]*Department of Astronomy, The University of Texas at Austin, Austin, TX 78712, USA*
[5]*Laboratoire Lagrange, Observatoire de la Côte d'Azur, Université Côte d'Azur, CNRS, CS 34229, F-06304 NICE Cedex 4, France*
[6]*Laboratoire Artemis, Observatoire de la Côte d'Azur, Université Côte d'Azur, CNRS, CS 34229, F-06304 Nice Cedex 4, France*
[7]*Department of Physics and Astronomy, University of California, Davis, CA 95616, USA*
[8]*CIERA and Department of Physics and Astronomy, Northwestern University, 1800 Sherman Ave, Evanston, IL 60201, USA*

Accepted 2022 November 24. Received 2022 November 4; in original form 2022 March 14

ABSTRACT

The properties of young star clusters formed within a galaxy are thought to vary in different interstellar medium conditions, but the details of this mapping from galactic to cluster scales are poorly understood due to the large dynamic range involved in galaxy and star cluster formation. We introduce a new method for modelling cluster formation in galaxy simulations: mapping giant molecular clouds (GMCs) formed self-consistently in a FIRE-2 magnetohydrodynamic galaxy simulation on to a cluster population according to a GMC-scale cluster formation model calibrated to higher resolution simulations, obtaining detailed properties of the galaxy's star clusters in mass, metallicity, space, and time. We find ~ 10 per cent of all stars formed in the galaxy originate in gravitationally bound clusters overall, and this fraction increases in regions with elevated $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$, because such regions host denser GMCs with higher star formation efficiency. These quantities vary systematically over the history of the galaxy, driving variations in cluster formation. The mass function of bound clusters varies – no single Schechter-like or power-law distribution applies at all times. In the most extreme episodes, clusters as massive as 7×10^6 $M_\odot$ form in massive, dense clouds with high star formation efficiency. The initial mass–radius relation of young star clusters is consistent with an environmentally dependent 3D density that increases with $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$. The model does not reproduce the age and metallicity statistics of old ($>$ 11Gyr) globular clusters found in the Milky Way, possibly because it forms stars more slowly at $z > 3$.

Key words: ISM: clouds – globular clusters: general – open clusters and associations: general – galaxies: star clusters: general – galaxies: star formation.

1 INTRODUCTION

Stars can form either as members of unbound associations of stars that will disperse into the host galaxy, or as members of gravitationally bound star clusters (hereafter 'star clusters' or 'clusters') that can persist for significantly longer (Gouliermis 2018; Ward & Kruijssen 2018; Adamo et al. 2020a). The persistence of bound clusters is interesting because they are lasting, coherent relics of star formation events whose properties are thought to bear some imprint of their natal environment. Their evolution, and eventual demise, are shaped by both stellar dynamics and the galactic gravitational landscape (Fall & Zhang 2001; Gieles et al. 2006a), so a galaxy's star cluster population contains a record of both its formation and its dynamical history.

Deciphering this record requires an understanding of the evolution of the star-forming interstellar medium (ISM) within galaxies, *and* how these changing conditions map on to the properties of young star clusters. Observations point to several such connections. The fraction of star formation in bound clusters (cluster formation efficiency, CFE, commonly denoted Γ, Bastian 2008) has been found to vary, both from one galaxy to another and between different regions of a given galaxy (Goddard, Bastian & Kennicutt 2010; Cook et al. 2012; Adamo et al. 2015; Johnson et al. 2016; Chandar et al. 2017; Ginsburg & Kruijssen 2018; see Krumholz, McKee & Bland-Hawthorn 2019; Adamo et al. 2020a for review). The *initial mass function* (hereafter simply 'mass function') of young clusters may also vary: various works have reported evidence of a high-mass truncation at a certain characteristic cut-off mass scale (Gieles et al.

★ E-mail: mike.grudich@gmail.com
† NASA Hubble Fellow.

2006b; Adamo et al. 2015; Johnson et al. 2017; Messa et al. 2018a; Wainer et al. 2022), with reported values ranging from $\sim 10^4$ to 10^6 $M_\odot$, generally between regions of low and high star formation intensity Σ_{SFR}.[1] And recently, several works have noted possible variations in the *mass–radius* relation of young star clusters, with more intensely star-forming environments hosting more compact clusters for a given mass (Krumholz et al. 2019; Brown & Gnedin 2021; Choksi & Kruijssen 2021; Grudić et al. 2021b). In all, it is clear that there is an intimate connection between star-forming environment and the properties of young star clusters.

To understand the physical processes driving variations in cluster properties across cosmic time, we require a model that couples the full cosmological context of galaxy formation to the formation and evolution of individual star clusters. Resolution requirements make this presently impossible to do in direct calculations that track the formation of individual stars (e.g. Bate, Bonnell & Bromm 2003; Krumholz, Klein & McKee 2011; Haugbølle, Padoan & Nordlund 2018; Grudić et al. 2021a), so a number of approximate frameworks have been devised to model the formation and evolution of star clusters in galaxy simulations, accounting for various subsets of the relevant physics either self-consistently or with sub-resolution models. Simulations using a sub-grid ISM model (e.g. Springel & Hernquist 2003) do not follow the formation of individual giant molecular clouds (GMCs), so the population of GMCs are modelled according to the available bulk ISM properties on $\sim$ kpc scales (e.g. density, pressure, and metallicity), and the mapping from GMCs to stars and bound clusters in turn via semi-analytical models (Kruijssen et al. 2011; Pfeffer et al. 2018). Li et al. (2017) performed cosmological galaxy simulations with sufficient resolution to resolve some individual GMCs, and modelled cluster formation in them as unresolved accreting sink particles with sub-grid feedback injection (Agertz et al. 2013; Semenov, Kravtsov & Gnedin 2016), internal structure, and dynamical evolution (Gnedin, Ostriker & Tremaine 2014). Other galaxy simulations resolving as fine as $\sim$ 1pc scales have modelled clusters as bound collections of softened subcluster particles (Kim et al. 2018a; Lahén et al. 2019; Ma et al. 2020), using sub-grid prescriptions for star formation and feedback coupled on the relevant resolved scale.

Each of the approaches listed above has different advantages and potential pitfalls, but all rely upon somewhat uncertain prescriptions for unresolved star formation and stellar feedback, which have not been explicitly validated due to the difficulty of simulating star-forming GMCs self-consistently. Lacking a definitive numerical model for cosmological cluster formation and evolution, it is worthwhile to consider alternative approaches to treating cosmological star cluster formation in galaxies, especially ones that can be applied to *existing* galaxy simulations without modification.

In this work, we introduce a new post-processing technique for modelling star cluster formation in existing cosmological simulations that resolve the multiphase ISM: we map the properties of GMCs formed in a FIRE-2 cosmological zoom-in simulation (Wetzel et al. 2016; Guszejnov et al. 2020a; Hopkins et al. 2020b) on to a model star cluster population via a statistical model calibrated to high-resolution ($\sim$0.1 pc) cluster formation simulations with stellar feedback (Grudić et al. 2021b, hereafter G21). This produces definite predictions for the detailed formation efficiency, masses, formation times, metallicities, and initial sizes of star clusters, which can be compared with observed young star cluster catalogues and used as the initial conditions for detailed dynamical treatments of star cluster evolution (Rodriguez et al. 2022).

This paper is structured as follows. In Section 2, we describe our GMC and star cluster population modelling technique based upon coupling the results of Guszejnov et al. (2020a) and G21, and describe the Milky Way-mass galaxy model that we use as a case study. In Section 3, we present the results of our model, showing how the efficiency of bound cluster formation, the cluster initial mass function, and cluster size statistics vary across cosmic time according to the evolving ISM conditions in the galaxy. We also examine the properties of the clusters in age-metallicity space, and compare and contrast those statistics with those those of Milky Way globular clusters to comment on the viability of the simulation as a model for globular cluster formation. In Section 4, we discuss various implications of our results and compare and contrast our findings with previous treatments of cosmological star cluster formation. In Section 5, we summarize our key conclusions about the connection between galactic environment and star cluster formation.

2 METHODS

Our model of galactic star cluster formation has three steps: the FIRE-2 cosmological zoom-in galaxy simulation itself, the extraction of cloud properties from the simulation data, and the mapping of cloud properties on to star cluster properties via the model derived from the G21 GMC simulations. We visualize the procedure in Fig. 1 and describe each step in turn below.

2.1 FIRE-2 simulation

Here, we study the formation of a Milky Way-mass disc galaxy formed in a cosmological zoom-in simulation of the halo model m12i simulated as part of the FIRE-2 simulation suite (Hopkins et al. 2018b) with the GIZMO code (Hopkins 2015). This galaxy simulation accounts for a wide range of relevant cooling mechanisms down to 10 K via detailed fits and tables (Hopkins et al. 2018b), stellar radiative feedback including radiation pressure, photoionization and photoelectric heating (Hopkins et al. 2020a), OB/AGB stellar winds, and type Ia and II supernovae (Hopkins et al. 2018a), with rates derived from a standard simple stellar population model (Leitherer et al. 1999) assuming a Kroupa (2001) stellar initial mass function. The simulation also accounts for magnetic fields using the quasi-Lagrangian, Meshless Finite Mass (MFM) magnetohydrodynamics solver (Hopkins & Raives 2016), anisotropic Spitzer–Braginskii conduction and viscosity, and sub-grid metal diffusion from unresolved turbulence (Hopkins 2017; Su et al. 2017; Hopkins et al. 2020b).

At $z = 0$, the simulated galaxy has a stellar mass of $M_\star = 6.7 \times 10^{10}$ $M_\odot$ and a halo mass of $M_{200} = 1.2 \times 10^{12}$ $M_\odot$, similar to inferred present-day mass measurements of the Milky Way (Bland-Hawthorn & Gerhard 2016). See Sanderson et al. (2020) for various detailed comparisons of non-magnetohydrodynamic (MHD) version of this simulation with the Milky Way, and Gurvich et al. (2020) for a detailed analysis of the phase structure and dynamics of its ISM. We selected the version of the simulation with MHD, conduction, and viscosity as a more physically complete model, note that the incremental effects of such processes upon star formation and galaxy evolution in this simulation have been shown to be modest (Su et al. 2017; Hopkins et al. 2020b).

[1] See however, Mok, Chandar & Fall (2019), who found the significance of various reported mass function cut-off masses in the literature to be marginal, and Wainer et al. (2022) who explored how uncertainties in the few greatest cluster masses propagate into the uncertainty of the cut-off mass.

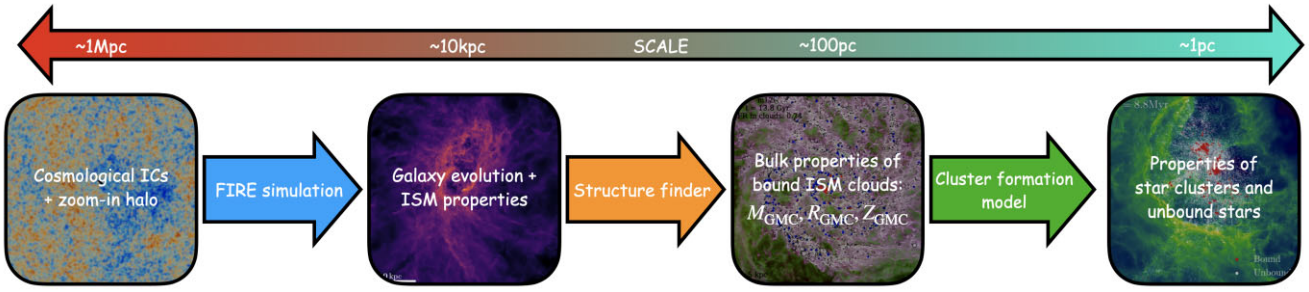

Figure 1. Diagram of our procedure for modelling the star cluster population of a simulated galaxy across cosmic time, described in full in Section 2. Starting with cosmological initial conditions and a choice of zoom-in halo, we simulate the cosmological evolution of the halo to $z = 0$ with FIRE (Hopkins et al. 2014; Wetzel et al. 2016; Hopkins et al. 2018b), run a structure finder to determine the bulk properties of bound clouds in the ISM (Guszejnov et al. 2020a), and plug these cloud properties into a model that predicted detailed star cluster properties, calibrated to high-resolution GMC simulations (Grudić et al. 2021b).

2.2 Cloud catalogue

The galaxy simulation has a baryonic mass resolution of 7070 $M_\odot$ and accounts for detailed multiphase ISM physics, allowing it to resolve the bulk properties of massive ($\gtrsim 10^5$ $M_\odot$) GMCs. The GMCs in the simulation assemble, form stars, and disperse in the simulation self-consistently over a typical time-scale of the order of 10Myr (Hopkins, Quataert & Murray 2012; Benincasa et al. 2020). In Guszejnov et al. (2020a), we used CloudPhinder,[2] an algorithm similar to SUBFIND (Springel et al. 2001), to identify the population of *self-gravitating* gas structures in this simulation: specifically, our algorithm identifies the population of 3D isodensity contours that enclose material with virial parameter $\alpha_{vir} < 2$ (Bertoldi & McKee 1992).

Guszejnov et al. (2020a) found these objects to have surface densities, size–linewidth relations, and maximum masses that resemble GMCs found in nearby galaxies (e.g. Larson 1981; Bolatto et al. 2008; Colombo et al. 2014; Freeman et al. 2017; Faesi, Lada & Forbrich 2018). However, the internal structure and dynamics of the clouds remain largely unresolved at our mass resolution, so we do not generally expect the star clusters that form in clouds to have reliable properties uncontaminated by numerical effects. Hence, we synthesize the cluster population in post-processing, using the properties of the self-gravitating clouds catalogued in Guszejnov et al. (2020a) as inputs to our star cluster formation model. We adopt a minimum mass cut of 2×10^5 $M_\odot$, or ∼30 times the mass resolution.

2.3 Cluster formation model

To determine the properties of clusters formed in the GMCs, we adopt the cluster formation model introduced in G21. In that study we performed a large suite of MHD star cluster-forming GMC simulations including stellar feedback, finding that quantities such as star formation efficiency (SFE), the fraction of star formation in bound clusters, and individual star cluster masses do depend sensitively upon the macroscopic properties of the parent GMC such as mass M_{GMC} and size R_{GMC}, but may also vary strongly from one GMC to another even if these quantities are held fixed, due to variations in the details of the initial turbulent flow. This led us to develop a *statistical* model that reproduces the statistical results (e.g. cluster mass functions and size distributions) of the ensemble of simulation results over many different initial realizations of turbulence. By modelling star cluster formation in this way, we arrived at a model that could reproduce the CFE and young star cluster mass functions observed in M83 (Adamo et al. 2015) fairly well, if the observed properties of GMCs in those respective galactic regions were taken as inputs (Freeman et al. 2017).

We briefly summarize the procedure of the model here. Given the mass M_{GMC}, size R_{GMC}, and metallicity Z_{GMC} of a cloud, the calculation proceeds as follows. First, we determine the total stellar mass formed in the cloud:

$$M_\star = \epsilon_{int} M_{GMC}, \tag{1}$$

where ϵ_{int} is the integrated star formation efficiency, which depends upon the GMC surface density $\Sigma_{GMC} = M_{GMC}/\pi R_{GMC}^2$ as

$$\epsilon_{int} = \left(\epsilon_{max}^{-1} + \left(\frac{\Sigma_{GMC}}{3200\,M_\odot pc^{-2}}\right)^{-1}\right)^{-1} \approx \frac{\Sigma_{GMC}}{3200\,M_\odot\,pc^{-2}}, \tag{2}$$

where $\epsilon_{max} = 0.7$ and the latter approximation holds when $\Sigma_{GMC} \lesssim 3000\,M_\odot pc^{-2}$.

With the total stellar mass known, we then determine the fraction of stars locked into bound clusters f_{bound}. G21 found f_{bound} to vary as a function of Σ_{GMC} and Z_{GMC}, but the significant scatter from one realization to another requires a probablistic model. Specifically, we let

$$f_{bound} = \left(1 + \left(\frac{\Sigma_{GMC}}{e^\delta \Sigma_{bound}}\right)^n\right)^{-1}, \tag{3}$$

where the random variable δ is sampled from a lognormal distribution with mean $\mu = 0$ and width $\sigma = 0.3$dex, and the metallicity-dependent parameters Σ_{bound} and n are

$$\Sigma_{bound} = (30 \log Z_{GMC} + 390)\ M_\odot\ pc^{-2} \tag{4}$$

and

$$n = -0.3 \log Z_{GMC} - 2, \tag{5}$$

where Z_{GMC} is the GMC metallicity in solar units. We found this prescription to reproduce the scaling of f_{bound} with GMC properties, and its intrinsic scatter across different realizations of a given set of bulk properties.

The total bound stellar mass is then $M_{bound} = f_{bound}\epsilon_{int} M_{GMC}$. The simulations generally found significant *multiplicity* of clusters formed in a single parent GMC, so we distribute the bound mass

[2] http://www.github.com/mikegrudic/CloudPhinder

among the individual clusters by sampling from a GMC-level mass distribution, given in G21, equation (11). Then, given the list of cluster masses, we determine their half-mass radii $r_{\rm h}$ by sampling a GMC-level size–mass relation:

$$r_{\rm h} = 3\,{\rm pc}\left(\frac{M_{\rm GMC}}{10^6\,{\rm M}_\odot}\right)^{\frac{1}{5}}\left(\frac{\Sigma_{\rm GMC}}{100\,{\rm M}_\odot\,{\rm pc}^{-2}}\right)^{-1}\left(\frac{Z_{\rm GMC}}{Z_\odot}\right)^{\frac{1}{10}}\left(\frac{M_{\rm cl}}{10^4\,{\rm M}_\odot}\right)^{\frac{1}{3}}, \quad (6)$$

with an intrinsic lognormal scatter of ± 0.4 dex in radius. Lastly, although we do not require the detailed density profile for this work, it is eventually required to model the dynamical evolution and observational characteristics of the clusters. We assume the clusters initially have a Elson, Fall & Freeman (1987) density profile and sample the density profile slope γ from a universal distribution consistent with observations (Grudić et al. 2018b).

For the purposes of the present analysis, we apply a lower mass cut of 10^3 M$_\odot$ to the cluster catalogue, similar to the completeness limits of extragalactic cluster catalogues (Adamo et al. 2015; Johnson et al. 2017; Messa et al. 2018a). Note that our model will have its own incompleteness function due to our lower GMC mass cut-off of 2×10^5 M$_\odot$ – how this maps on to a cluster mass scale will depend upon the detailed SFE and CFE statistics of the cloud sample.

2.3.1 *Sampling procedure*

The clouds selected by the cloud-finding algorithm are not a complete census of all clouds to ever form stars within the model galaxy. The simulation has 601 snapshots, which can be spaced as far apart as ~ 26 Myr, likely significantly longer than the lifetime of all but the most massive GMCs, which, at least in FIRE and similar simulations (Hopkins et al. 2012; Benincasa et al. 2020; Li et al. 2020), and observations (Chevance et al. 2020), is generally of the order of the cloud freefall time, $\sim 3-10$ Myr. Therefore, clouds in the simulation typically form and disperse between snapshots (which are typically ~ 22Myr apart), preventing them from being found by our structure finder. Furthermore, we have found in previous high-resolution GMC simulations (Grudić et al. 2019) that a significant fraction of star formation within a cloud can happen when it is already in a supervirial state due to feedback from the first massive stars that formed in it – under such conditions, the cloud would not be identified by our algorithm, even if it is present in the snapshot. Therefore, a simple 1-to-1 mapping of catalogued clouds to stellar populations will tend to underestimate the total stellar mass in our setup.

To address this issue, we adopt the following sampling procedure to synthesize the cluster population while matching the simulated star formation history, from each snapshot:

(i) Measure the total galactic stellar mass $\Delta M_\star^{\rm gal}$ *actually formed* in the simulation in the time between snapshots i and $i+1$.

(ii) Sample from the catalogue of clouds found in snapshot i randomly until the total stellar mass formed by the cloud sample according to the G21 model exceeds $\Delta M_\star^{\rm gal}$.

In this way, we use the bound clouds as statistical tracers of the full population of progenitor clouds, and recover a model that accounts for the entire stellar mass of the galaxy. Note that while we are requiring 100 per cent of the stellar mass to be formed in the bound clouds, only a fraction of that mass will be in bound star clusters according to our cluster formation model.

One caveat of this model is that bound clouds are not strictly expected to be the sole contributors to star formation: a GMC with essentially *any* virial parameter could form some number of stars, in a collapsing subregion. However, we do expect the overall stellar population to be heavily weighted towards those formed in a bound progenitor cloud, because star formation efficiency is expected to fall off rapidly as a function of virial parameter (Padoan, Haugbølle & Nordlund 2012; Dale 2017; Kim, Ostriker & Filippova 2021). In effect, we model the expected continuous-but-steep transition between starless and star-forming clouds with decreasing virial parameter as a step-function at $\alpha_{\rm vir} = 2$.

3 RESULTS

In Fig. 2, panel 1, we plot the total stellar mass of the galaxy as a function of time. At $z = 0$, the galaxy has a total stellar mass of 6.7×10^{10} M$_\odot$. At $z = 0$, this galaxy has some noted differences from the Milky Way. It is not part of a 'Local Group' that contains another comparably massive galaxy within 1 Mpc. Its gas fraction is ~ 20 per cent, versus ~ 10 per cent for the Milky Way, and its star formation rate at the present epoch is significantly higher, $\sim 10\,{\rm M}_\odot{\rm yr}^{-1}$, compared to the observed $\sim 2\,{\rm M}_\odot{\rm yr}^{-1}$ (Licquia & Newman 2015; Bland-Hawthorn & Gerhard 2016). As such, this galaxy is later-forming than the Milky Way, having a higher SFR at late times and a lower SFR at early times, giving it roughly equal stellar mass at $z = 0$. This fact will prove important when we interpret the age and metallicity statistics of the massive clusters formed in the model, vis-a-vis those found in the Milky Way (Section 3.4).

3.1 Cluster formation efficiency

The galactic CFE Γ varies in time and space according to local GMC properties in the simulation according to the scalings given in Section 3 and the sampling procedure described in Section 2. Overall, this approach finds that 13 per cent of all stars formed in this galaxy are in bound clusters following the dispersal of their natal cloud – in other words, the vast majority of stars are never members of clusters that remain bound after gas expulsion.

In Fig. 2, panels 2 and 3, we break down the star formation rate and CFE Γ for different mass ranges of bound clusters. On average, $\Gamma \sim 10$ per cent, with no clear systematic trend with cosmic time. However, over $\lesssim 100$Myr time-scales, Γ can undergo significant swings, between ~ 1 and 100 per cent. From comparison of panels 2 and 3, it is evident that these swings follow modulations in the star formation rate of the galaxy, indicating that variations in star formation activity are driving variations in GMC properties (and hence Γ in turn). However, Γ is clearly not a one-to-one function of SFR, as the most intense starbursts do not necessarily have the most efficient cluster formation – rather, we will show that Γ depends more sensitively on the *intensity* of star formation $\Sigma_{\rm SFR}$ than the total SFR, as has been inferred from observations (Hollyhead et al. 2016).

To illustrate how GMC properties can vary dramatically from those typically observed in present-day nearby disc galaxies (e.g. Bolatto et al. 2008), giving rise to high cluster formation efficiencies, Fig. 3 plots the surface density of gas in the galaxy surrounding the most prodigious cluster-forming cloud in our catalogue, during the large spike in Γ at $z \sim 0.8$ evident in Fig. 2. This cloud is found at the edge of a 'superbubble', a large cavity evacuated by a major stellar feedback event, similar to some of the more extreme cases noted in high-redshift galaxies simulated in Ma et al. (2020). The cloud has a mass of 3×10^7 M$_\odot$ and a mean surface density $\Sigma_{\rm GMC} = M_{\rm GMC}/\left(\pi R_{\rm GMC}^2\right) = 1200\,{\rm M}_\odot\,{\rm pc}^{-2}$. According to equations (2)–(3), this surface density gives the cloud a star formation efficiency of

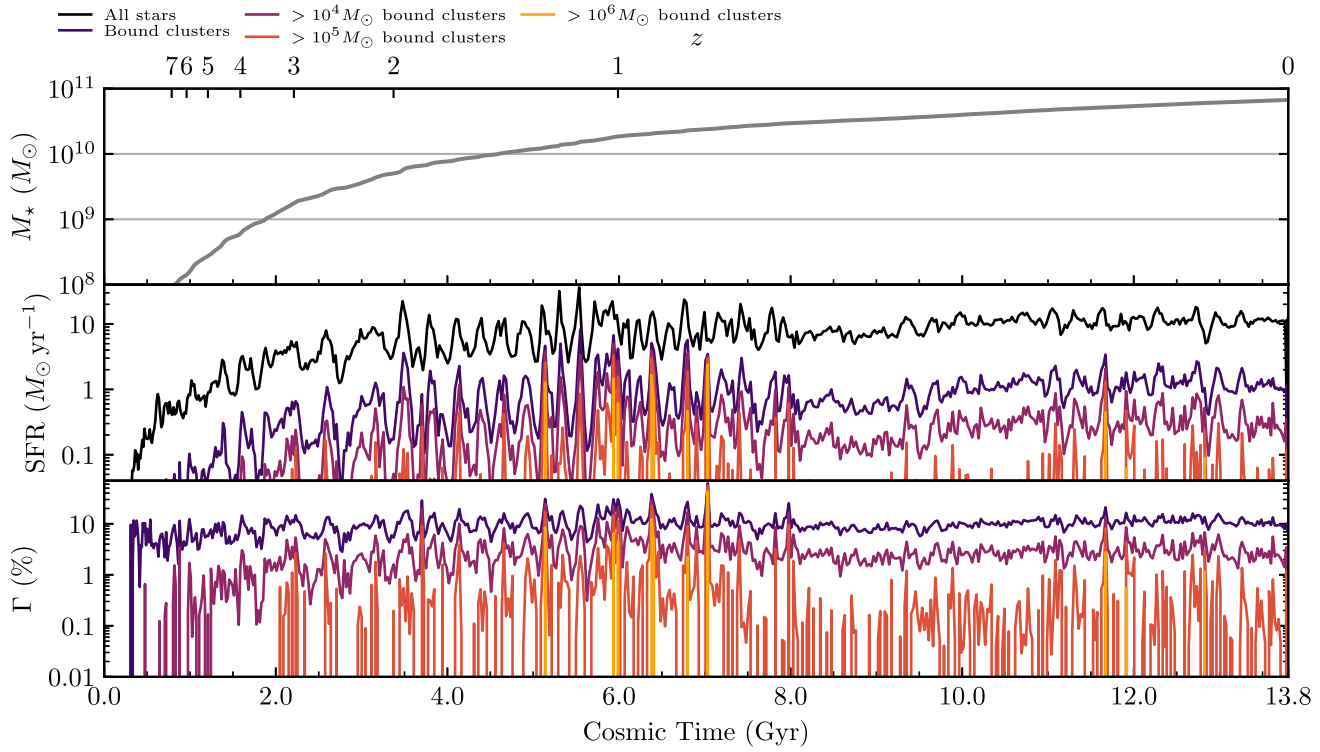

Figure 2. Star and star cluster formation history of the simulated galaxy. *Top:* Total stellar mass of the host galaxy as a function of cosmic time. *Middle*: Star formation rate in each simulation snapshot, showing the contributions of bound clusters above different mass cuts. *Bottom*: CFE of the simulated galaxy across cosmic time, for all bound clusters and various cluster mass cuts. $<10^4$ M$_\odot$ clusters are produced with an efficiency varying only by a factor of $\sim$3, $>10^4$ M$_\odot$ clusters near-constantly with an efficiency varying by an order of magnitude, and more massive clusters only episodically.

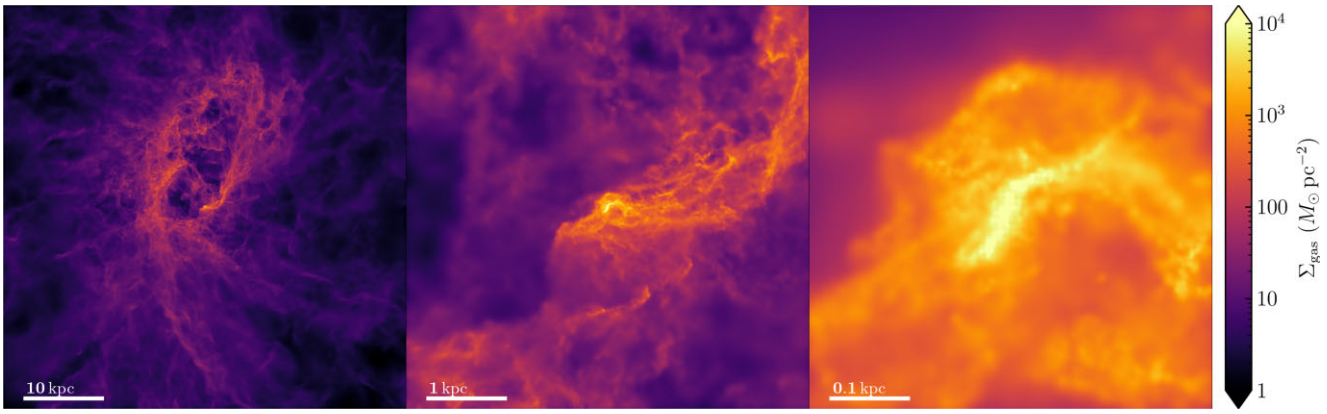

Figure 3. Gas surface density at the formation site of the most massive (7×10^6 M$_\odot$) cluster formed in the history of the simulated galaxy, in a 3×10^7 M$_\odot$ cloud with mean surface density $\Sigma_{\rm GMC} \sim 1200\,{\rm M}_\odot{\rm pc}^{-2}$ at $z \sim 0.8$. The cloud is found at the edge of a large bubble or cavity, and the galaxy still has a highly irregular morphology. 3D animations of this cloud and the two next-most-massive cluster progenitor clouds can be viewed here.

27 per cent and a CFE of almost unity, allowing it to form the most massive cluster in the history of the galaxy, with a mass of 7×10^6 M$_\odot$ and an initial half-mass radius of 5pc.[3] Born near the Galactic Centre, the cluster has a dynamical friction time much less than a Hubble time, so its mostly likely fate is to spiral into the Galactic Centre and merge into the nuclear star cluster (Capuzzo-Dolcetta & Miocchi 2008; Pfeffer et al. 2018; Rodriguez et al. 2022).

[3]This cloud also produced the 'behemoth' cluster originally described and studied in Rodriguez et al. (2020).

3.1.1 *Environmental scaling relations*

To analyse the relation between the local galactic environment and Γ, we break down the cloud catalogue in terms of GMC surface density $\Sigma_{\rm GMC}$, local gas density $\Sigma_{\rm gas}$ measured on 1kpc scales, and local star formation surface density $\Sigma_{\rm SFR}$, also measured on 1kpc scales. Note that only $\Sigma_{\rm GMC}$ is a direct input for our model. To compute $\Sigma_{\rm gas}$ in the vicinity of a cloud, we count the total gas mass within a 1kpc radius of the cloud and take $\Sigma_{\rm gas} = M_{\rm gas}/\pi/\,(1{\rm kpc})^2$. We compute $\Sigma_{\rm SFR}$ similarly, estimating the total SFR within 1kpc of each cloud by counting the total stellar mass $<$ 10Myr old and taking SFR $\approx {\rm M}_{\rm star}/10{\rm Myr}$, and then let $\Sigma_{\rm SFR} = {\rm SFR}/\pi/\,(1{\rm kpc})^2$.

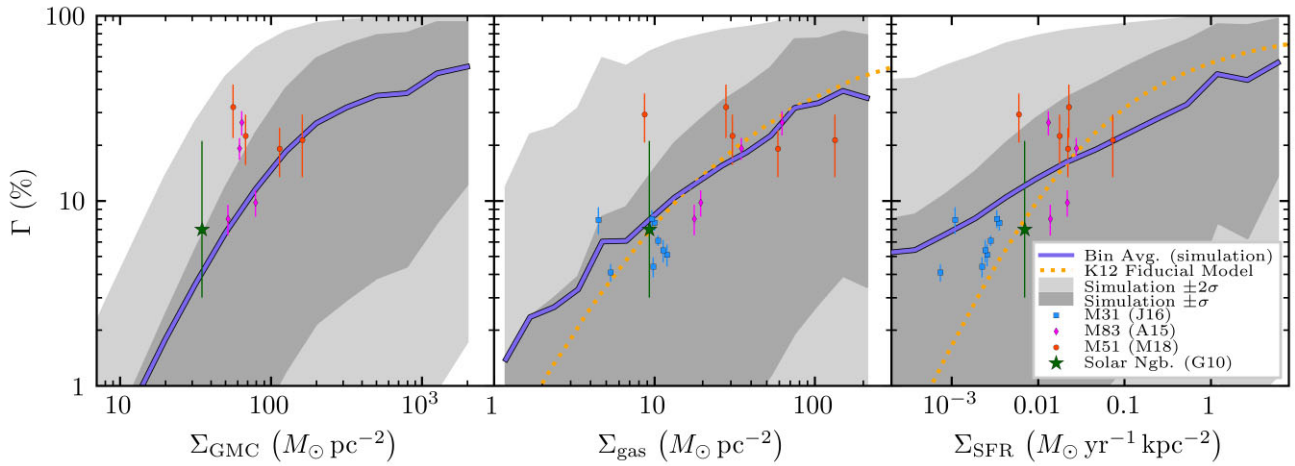

Figure 4. Model-predicted CFE $\Gamma = \sum M_{\rm cl}/\sum M_\star$ as a function of GMC-scale surface density $\Sigma_{\rm GMC}$ (left), kpc-scale galaxy gas surface density $\Sigma_{\rm gas}$ (middle), and kpc-scale star formation surface density $\Sigma_{\rm SFR}$ (right). Solid curves plot the average efficiency ($\sum M_{\rm cl}/\sum M_\star$ in a given bin), shaded regions plot the CFE of the clouds that the $16-84$ per cent and $5-95$ per cent percentiles of stars formed in. We compare with the fiducial model of Kruijssen 2012 (K12) and several local measurements in M31 (Johnson et al. 2016), M83 (Adamo et al. 2015), M51 (Messa et al. 2018b), and the Solar neighbourhood (Goddard et al. 2010) (see Section 3.1.1 for details).

In Fig. 4, we plot the average $\Gamma = \sum M_{\rm cl}/\sum M_\star$ in different bins of $\Sigma_{\rm GMC}$, $\Sigma_{\rm gas}$, and $\Sigma_{\rm SFR}$, and compare these results with various observations and the predictions of the fiducial version of the Kruijssen (2012, hereafter K12) analytic model. We plot Γ measurements in resolved subregions of M83 (Adamo et al. 2015), M31 (Johnson et al. 2016), and M51 (Messa et al. 2018a), using $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ values provided in those respective works. We use Γ values measured with age cuts of >10Myr and $\lesssim 100$Myr from these works, we also plot the measurement for the Solar neighbourhood given in Goddard et al. (2010), and use $\Sigma_{\rm gas} = 10\,{\rm M_\odot pc^{-2}}$ and $\Sigma_{\rm SFR} = 7\times10^{-3}{\rm M_\odot\,yr^{-1}kpc^{-2}}$ (Bovy 2017). For $\Sigma_{\rm GMC}$, we use the mass-weighted median values of clouds in the Colombo et al. (2014) and Freeman et al. (2017) catalogues, in the same respective radial bins as Γ was measured, in M51 and M83, respectively, and for the solar neighbourhood we use the fiducial value of $35\,{\rm M_\odot\,pc^{-2}}$ given in Lada & Dame (2020).

Fig. 4 shows that Γ exhibits a clear scaling with $\Sigma_{\rm GMC}$, $\Sigma_{\rm gas}$, and $\Sigma_{\rm SFR}$. The correlation with $\Sigma_{\rm GMC}$ follows directly from the cluster formation model via the dependence of the cloud-scale $f_{\rm bound}$ in equation (3), which is physically a consequence of the higher star formation efficiency of denser clouds, equation (2). The $\Sigma_{\rm gas}$–Γ relation agrees well with the fiducial Kruijssen (2012) relation, and has a similar level of agreement with the observations. The $\Sigma_{\rm SFR}$–Γ relation also agrees well with the fiducial Kruijssen (2012) model for $\Sigma_{\rm SFR} > 10^{-2}\,{\rm M_\odot\,yr^{-1}\,kpc^{-2}}$, however, it predicts a systematically greater Γ at lower $\Sigma_{\rm SFR}$, and as a result matches the Johnson et al. (2016) M31 measurements better. This discrepancy with the fiducial K12 model was noted in Johnson et al. (2016) and a modification to the model was proposed that reproduces the observations with similar success.

The most glaring discrepancies with *both* our and K12's predictions is M51: taken at face value, none of the measurements provided by Messa et al. (2018b) (red points) substantiate a systematic trend in Γ with *any* environmental property considered here. One possible explanation is that the measurements do not fully capture variations in kpc-scale environmental properties: Messa et al. (2018b) measured Γ in radial bins, but M51 is the prototype for strong spiral structure – within a given radial bin, $\Sigma_{\rm GMC}$, $\Sigma_{\rm gas}$, and $\Sigma_{\rm SFR}$ can vary systematically as a function of azimuth. Within either our or K12's models, cluster formation in a given bin would likely be dominated by the high-density spiral arms, and this could obscure any signal of small Γ values expected in the inter-arm regions.

Another complication of the measurements in M51 is that Messa et al. (2018a) found fairly steep cluster age distributions in certain regions, suggesting that cluster destruction may reduce the measured value of Γ in the $10-100$Myr age window significantly. This would not be as much of an issue in M31 and M83, which have much flatter age distributions (Bastian et al. 2012; Johnson et al. 2016).

The average $\Sigma_{\rm gas}$–Γ relation is well approximated by the fit

$$\Gamma = \min\left(0.063\left(\frac{\Sigma_{\rm gas}}{10\,{\rm M_\odot\,pc^{-2}}}\right)^{0.8}, 1\right), \quad (7)$$

and the dependence on $\Sigma_{\rm SFR}$ is approximated by

$$\Gamma = \min\left(0.12\left(\frac{\Sigma_{\rm SFR}}{10^{-2}\,{\rm M_\odot\,yr^{-1}\,kpc^{-2}}}\right)^{0.3}, 1\right), \quad (8)$$

which is quite similar to the $\Gamma \propto \Sigma_{\rm SFR}^{0.24}$ fit to compiled observational data by Goddard et al. (2010).

3.1.2 *Relating Σ_{GMC} to Σ_{gas} and Σ_{SFR}*

Recalling that $\Sigma_{\rm GMC}$ is the quantity that determines Γ within our model, the trends in Γ with $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ require that $\Sigma_{\rm GMC}$ have some systematic scaling with these quantities. We plot these relations in Figs 5 and 6 – both quantities are similarly predictive of $\Sigma_{\rm GMC}$, with a residual scatter in either relation of 0.6dex across nearly the entire dynamic range. These (mass-weighted) relations and their scatter can be modelled by the fits

$$\Sigma_{\rm GMC} = \left(8 + 1.8\left(\frac{\Sigma_{\rm gas}}{1\,{\rm M_\odot pc^{-2}}}\right)^{0.9}\right){\rm M_\odot pc^{-2}} \pm 0.6{\rm dex}, \quad (9)$$

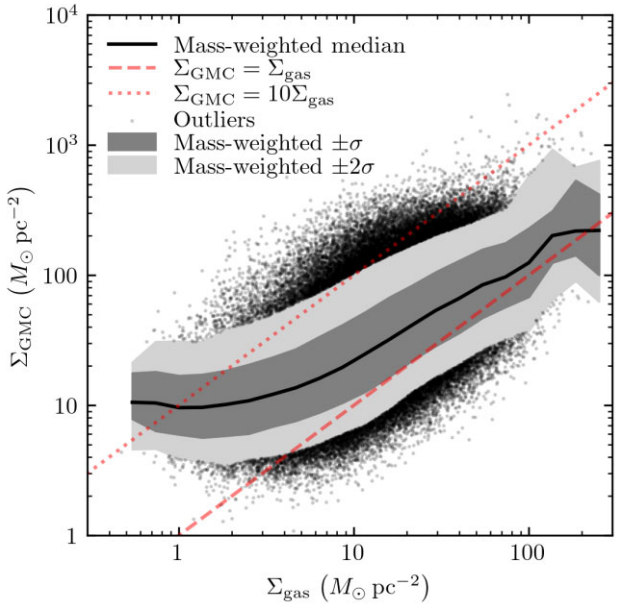

Figure 5. Relation between the surface density $\Sigma_{\rm GMC}$ of individual bound GMCs in our catalogue, and the average gas surface density $\Sigma_{\rm gas}$ in a 1kpc sphere surrounding each cloud. We plot mass-weighted quantiles binned by $\Sigma_{\rm gas}$, and $>2\sigma$ outliers are plotted as points.

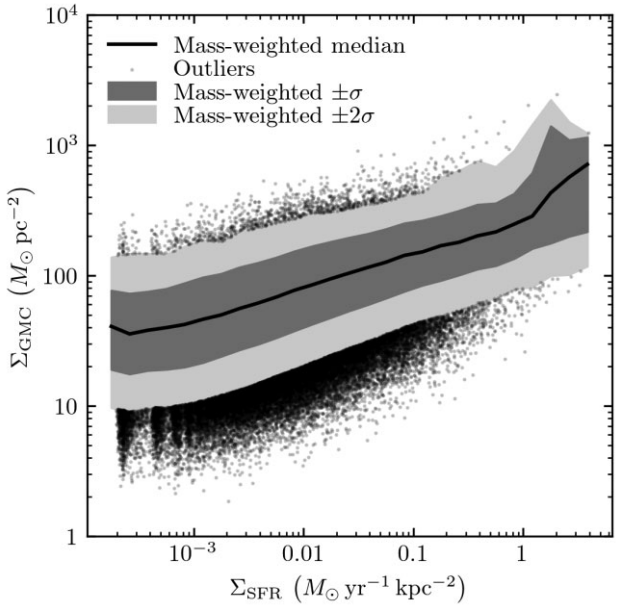

Figure 6. Relation between the surface density $\Sigma_{\rm GMC}$ of GMCs in our catalogue, and the average star formation surface density $\Sigma_{\rm SFR}$ in a 1kpc sphere surrounding them.

and

$$\Sigma_{\rm GMC} = \left(14 + \left(\frac{\Sigma_{\rm SFR}}{10^{-2}\,{\rm M}_\odot\,{\rm yr}^{-1}\,{\rm kpc}^{-2}}\right)^{0.33}\right) {\rm M}_\odot{\rm pc}^{-2} \pm 0.6{\rm dex}. \quad (10)$$

It should be noted that this fit to the $\Sigma_{\rm gas}$–$\Sigma_{\rm GMC}$ relation is not expected to extrapolate to arbitrarily high $\Sigma_{\rm gas}$, as the asymptotic scaling is $\Sigma_{\rm GMC} \propto \Sigma_{\rm gas}^{0.9}$, implying a crossover point where $\Sigma_{\rm gas} \sim \Sigma_{\rm GMC}$ – above this point, an average scaling at least as steep as $\propto \Sigma_{\rm gas}$ is necessary, as otherwise the 'clouds' would be voids against the denser environment.

That $\Sigma_{\rm GMC}$ (and the resulting Γ) should correlate with *both* $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ is unsurprising, as these quantities tend to be highly correlated across a large dynamic range of scales (Schmidt 1959; Kennicutt 1998; Heiderman et al. 2010; Elmegreen 2018; Pokhrel et al. 2021).

3.2 Cluster initial mass function

We now examine the initial mass function of the bound clusters formed in our model, recalling that our model samples cluster masses from a local mass function within each GMC, so the integrated galactic mass function will be the result of stacking samples from the variable mass functions of each cloud within a certain age bin. Fig. 7 plots the mass functions of clusters formed in different 100Myr windows across cosmic time, and the total mass function. We compare these with a variety of mass functions observed in nearby galaxies, generally for clusters in the age range $10-100$Myr where possible. These data include catatlogues from the LMC and SMC (H03; Hunter et al. 2003), M83 (A15; Adamo et al. 2015), M51 (M18; Messa et al. 2018a), the M31 PHAT field (J17; Johnson et al. 2017), the Antennae (W10; Whitmore et al. 2010), NGC1566 (H16; Hollyhead et al. 2016), NGC3256 (M16; Mulia, Chandar & Whitmore 2016), and NGC628 (A17; Adamo et al. 2017).

The clusters formed in the simulation span essentially the entire observed mass range of young star clusters found in nearby galaxies (up to 7×10^6 M$_\odot$), with the exception of NGC 7252, which hosts the most massive known young cluster (Maraston et al. 2004; Bastian et al. 2013). The integrated mass function over cosmic time is fairly bottom-heavy, resembling a power-law $dN/dM_{\rm cl} \propto M_{\rm cl}^{-2.5}$. The explanation for this bottom-heavy mass function can be discerned from the diverse mass functions seen at different periods in the galaxy's history: the galaxy form clusters as massive as $\sim 10^7$ M$_\odot$ only during a couple exceptional episodes, and spends most of its time forming clusters significantly less massive, putting most of the overall bound cluster mass in lower mass clusters.

The sequence of mass functions exhibits a a discernible evolution over cosmic time. At early times (<3Gyr), fewer, lower mass clusters generally form, but as we reach ~ 6Gyr ($z \sim 1$) the galaxy experiences its most intense epsiodes of cluster formation, forming clusters as massive as 7×10^6 M$_\odot$ (as illustrated in Fig. 3). And finally, as we approach $z \sim 0$ the formation of clusters $>10^6$ M$_\odot$ becomes rarer, and the maximum young cluster mass is typically of the order of 10^5 M$_\odot$, as found in various nearby disc galaxies (e.g. Adamo et al. 2015; Messa et al. 2018a). When plotting the mass function in equal time windows, a large portion of the variation is simply driven by variations in the overall star and star cluster formation rate – to control for this, Fig. 7, panels 2 and 3, plot the mass functions controlling for the total stellar mass and total cluster mass formed in the respective time windows. This collapses most of the variation, but even when controlling for the total formation rate, true variations in the shape of the mass function exist – the different mass functions tend to vary in slope at the high-mass end, being steeper (or having lower 'truncation' mass) when lower mass clusters form and shallowest when the highest-mass clusters form.

When analysing the shape of cluster mass functions, it is important to control for the total mass of clusters in the sample, as a poorly sampled mass function with a large or non-existent cut-off can be difficult to distinguish from a mass function with a genuine cut-off (Mok et al. 2019). In Fig. 8, we distinguish between different

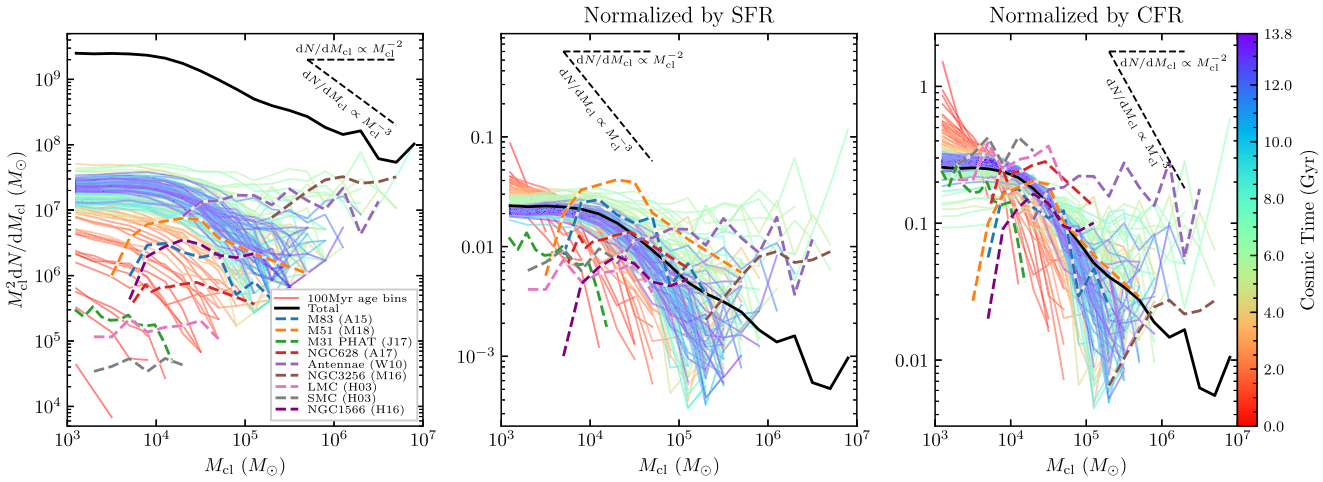

Figure 7. Initial mass functions of bound clusters, plotted as $M_{\rm cl}^2 {\rm d}N/{\rm d}M_{\rm cl}$, i.e. compensating for the expected dominant $\propto M_{\rm cl}^2$ scaling. We compare the mass function of clusters formed in 100 Myr windows across cosmic time (colour coded by time) with observed mass functions in nearby galaxies (see Section 3.2 for data compilation references). Also plotted is the *total* mass function integrated across cosmic time (black). *Left*: $M_{\rm cl}^2 {\rm d}N/{\rm d}M_{\rm cl}$ with no additional normalization. *Centre*: Like panel 1, but dividing out the total stellar mass formed (bound and unbound) in the respective simulation and galactic age bins (i.e. normalizing by the star formation rate). *Right*: Like panel 1, but dividing out the total *bound* cluster mass formed in each age bin (i.e. normalizing by the cluster formation rate).

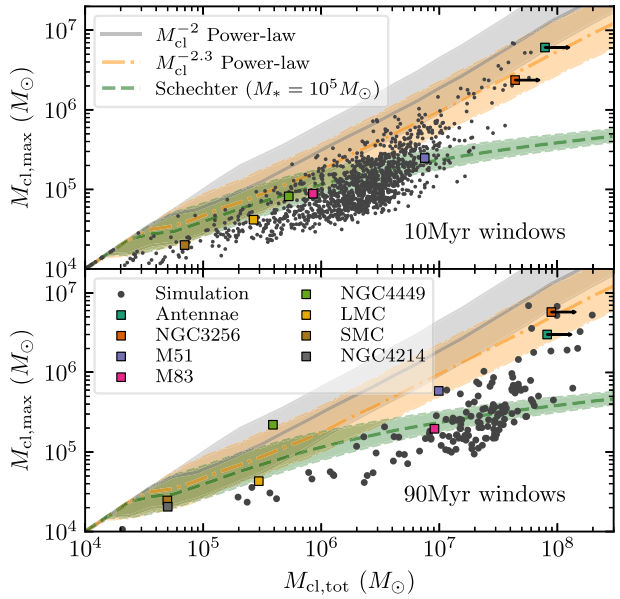

Figure 8. Relation between the total mass in clusters more massive than 10^4 $M_\odot$, and their maximum mass, in 10 Myr (top) and 90 Myr (bottom) age bins, compared with observations of 1–10 Myr old (top) and 10–100 Myr old (bottom) clusters in different galaxies, following Mok et al. (2019). Points show values across for each 10Myr windows in the simulation, squares are the sample of different galaxies compiled in Mok et al. (2019). For comparison we plot the median and $\pm\sigma$ ranges according to three hypotheses for the mass function: a $\propto M_{\rm cl}^{-2}$ or $\propto M_{\rm cl}^{-2.3}$ power-law (here truncated at 10^8 $M_\odot$) and a 'Schechter-like' form $\propto M_{\rm cl}^{-2} \exp\left(-M_{\rm cl}/10^5\,{\rm M}_\odot\right)$.

hypotheses for the mass function by plotting how the mass of the most massive cluster varies as a function of the total cluster mass above 10^4 $M_\odot$, for both 10 and 90 Myr time windows in the simulation, compared to observed 1–10 Myr and 10–100 Myr old cluster populations, respectively. For comparison we plot the expected scalings assuming various different forms for the overall mass function – a pure power-law ${\rm d}N/{\rm d}M_{\rm cl} \propto M_{\rm cl}^{-2}$, a slightly steeper ${\rm d}N/{\rm d}M_{\rm cl} \propto M_{\rm cl}^{-2.3}$, and a Schechter-like form with a cut-off of 10^5 $M_\odot$, $\propto M_{\rm cl}^{-2} \exp\left(-M_{\rm cl}/10^5 M_\odot\right)$. The data – in both the simulations and observations – do not conform perfectly to any one assumed form of the mass function. Rather, they appear to span a sequence that agrees well with the Schechter-like form when the total mass is lower, and then break from this pattern towards a regime that agrees better with the $M_{\rm cl}^{-2.3}$ power-law form. From this is clear that our mass functions defy a description in terms of any one simple, time-invariant power-law or Schechter-like form. The cluster mass function varies intrinsically across environment and cosmic time.

3.2.1 *Environmental dependence of the mass function*

In Section 3.1.1, we found that variations in CFE can be traced to environmental variations in GMC properties, so naturally this is also the case for the mass function, explaining the variations along the sequence of points plotted in Fig. 8. In Fig. 9, we plot the mass-weighted quantiles of the cluster mass function as a function of $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$: within each bin: the cluster mass below which a certain per cent of the total cluster mass in each bin lies. We find that the mass scale of clusters increases monotonically with both environmental properties considered. This trend is primarily driven by the increase in star and star CFE in the denser GMCs found in denser environments (cf. Figs 5 and 6), rather than an increase in the mass scale of GMCs. For example, the most massive 7×10^6 $M_\odot$ cluster formed in a 3×10^7 $M_\odot$ cloud (Fig. 3) with high efficiency due it its high $\gtrsim 10^3\,{\rm M}_\odot\,{\rm pc}^{-2}$ mean surface density, whereas the most massive cloud in the cloud catalogue is 2×10^8 $M_\odot$ but had a mean surface density of $\sim 100\,{\rm M}_\odot\,{\rm pc}^{-2}$, so its most massive cluster was only 10^5 $M_\odot$.

Johnson et al. (2017) proposed a similar correlation between the cut-off of the mass function and the average value of $\Sigma_{\rm SFR}$ in a galaxy, fitting a power-law relation $M_* \propto \langle\Sigma_{\rm SFR}\rangle^{1.1}$ to mass function fits from M31, M83, M51, and the Antennae. Direct comparison to this result is complicated by the fact that not all of our mass functions *are* well fit by a Schechter-like model with a constrained cut-off, but in Fig. 10 we plot the mass below which 90 per cent of the total cluster mass exists, in time windows containing equal formed stellar mass, as a

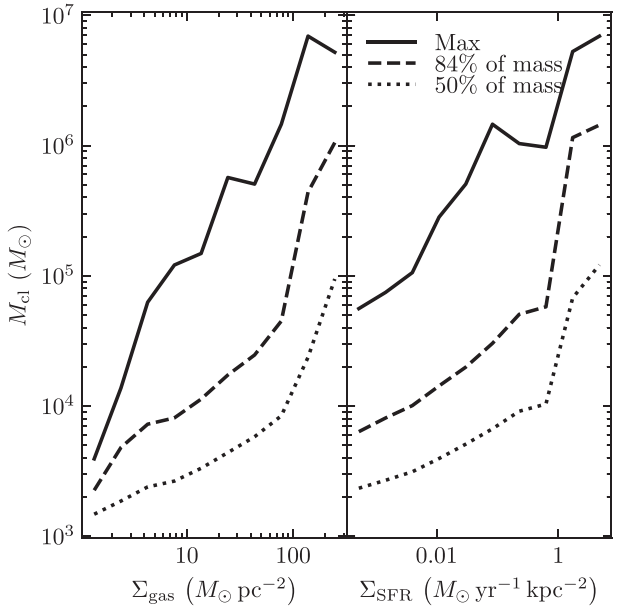

Figure 9. Mass-weighted percentiles of the cluster mass function (cluster mass below which a given percentage of the total cluster mass lies) binned by $\Sigma_{\rm gas}$ (left) and $\Sigma_{\rm SFR}$ (right), taken over all of cosmic time in the simulation.

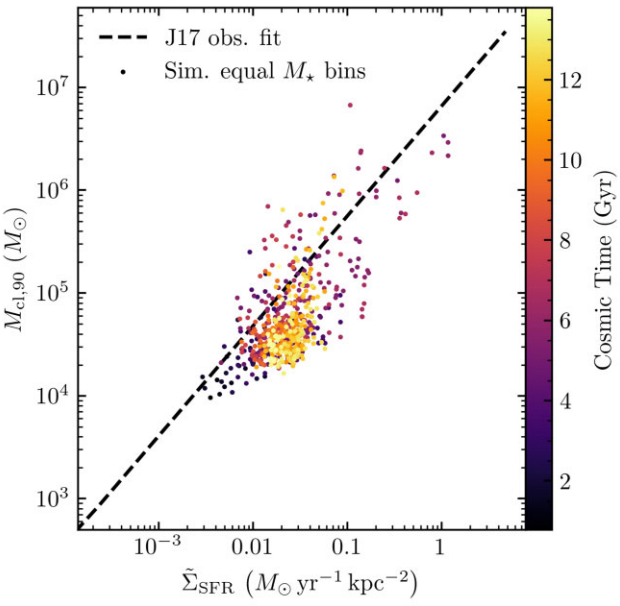

Figure 10. Relation between the 90th mass percentile of the cluster mass function and the median $\Sigma_{\rm SFR}$ that a star formed in, in time windows from the simulation during which equal stellar mass forms. For comparison we plot the fit to measurements of the Schechter cut-off of cluster mass funtions proposed in Johnson et al. (2017) (dashed).

function of $\tilde{\Sigma}_{\rm SFR}$, the median $\Sigma_{\rm SFR}$ that a star formed in each time window. This has a similar relation to that found in Johnson et al. (2017).

3.3 Initial mass–radius relation

In Fig. 11, we plot the 2D projected half-mass radii $R_{\rm eff}$ of the simulated star clusters as a function of their initial mass $M_{\rm cl}$. The

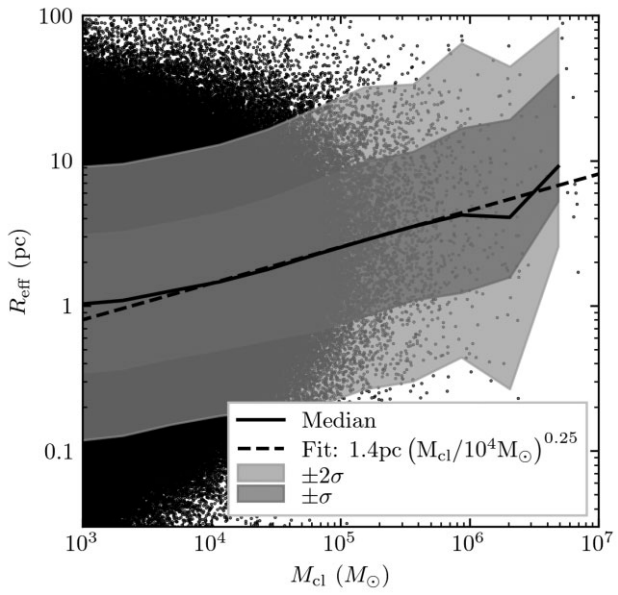

Figure 11. Mass–radius relation of the entire model cluster population, plotting the projected half-mass radius $R_{\rm eff}$. Overlaid are number-weighted median, $\pm\sigma$, and $\pm 2\sigma$ quantiles in different mass bins, and an unweighted least-squares power-law fit giving $R_{\rm eff} \propto M_{\rm cl}^{0.25}$.

initial mass–radius relation has large scatter (~ 0.5dex), which is nearly independent of cluster mass. This scatter is the result of convolving the *intrinsic* scatter of ± 0.4dex set by cluster formation physics resolved in the G21 simulations with the properties of the GMC catalogue, which introduce additional scatter because the median cluster size scales $\propto \Sigma_{\rm GMC}^{-1}$. Fitting the entire data set to a power law gives

$$R_{\rm eff} = 1.4{\rm pc}\left(\frac{{\rm M_{cl}}}{10^4{\rm M_\odot}}\right)^{0.25}, \tag{11}$$

i.e. slightly shallower than a constant–density relation $R_{\rm eff} \propto M_{\rm cl}^{1/3}$, and in agreement with the slope measured from the aggregated LEGUS catalogue of star clusters in nearby star-forming galaxies (Brown & Gnedin 2021).

The normalization of equation (11) is ~ 40 per cent smaller than the relation fitted in Brown & Gnedin (2021), and our scatter is roughly twice as great. Similar discrepancies were noted in Grudić et al. (2021b) when comparing the cluster radii predicted by this model with the measurements by Ryon et al. (2015), and we discussed several possible explanations. First, the numerical simulations from which the cluster formation model is derived are subject to significant uncertainties because they are sensitive to uncertain assumptions about the unresolved conversion of gas to stars (Ma et al. 2020; Hislop et al. 2022). Some discrepancy in the predicted stellar phase-space density is therefore expected. However, it should also be noted that we are predicting only *initial* cluster radii, and stellar and dynamical evolution will tend to increase the size of the cluster over time. And the scatter is also expected to decrease as the cluster population evolves, because while dense clusters expand to fill their tidal Roche lobe, underdense clusters will be stripped of their outer parts through tidal shocks, or destroyed entirely (Gieles & Renaud 2016).

We have also examined the cluster size–mass relations when controlling for their natal environmental conditions $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ – when binning the data by these quantities, we generally find a best-fitting mass–radius relation consistent with $R_{\rm eff} \propto M_{\rm cl}^{1/3}$ (e.g.

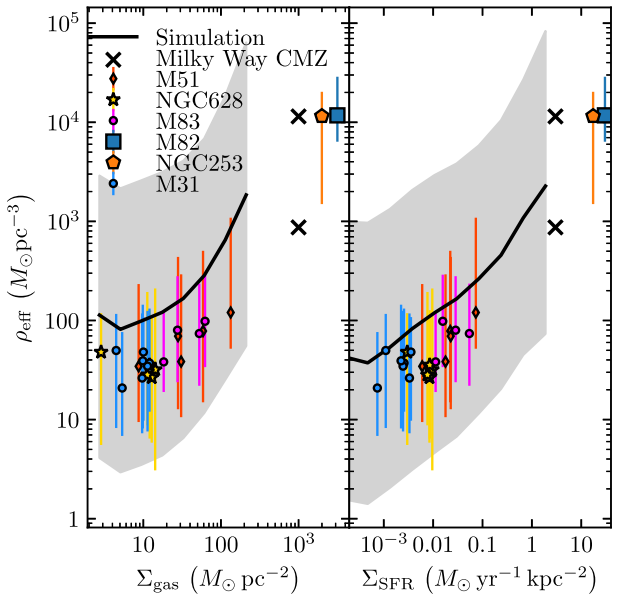

Figure 12. Scaling of 3D cluster density with $\Sigma_{\rm gas}$ (left) and $\Sigma_{\rm SFR}$ (right). The line and shaded interval plot the number-weighted median and $\pm\sigma$ quantiles in different bins of $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$, which we compare with observations in different regions of various galaxies (see Section 3.3 for details).

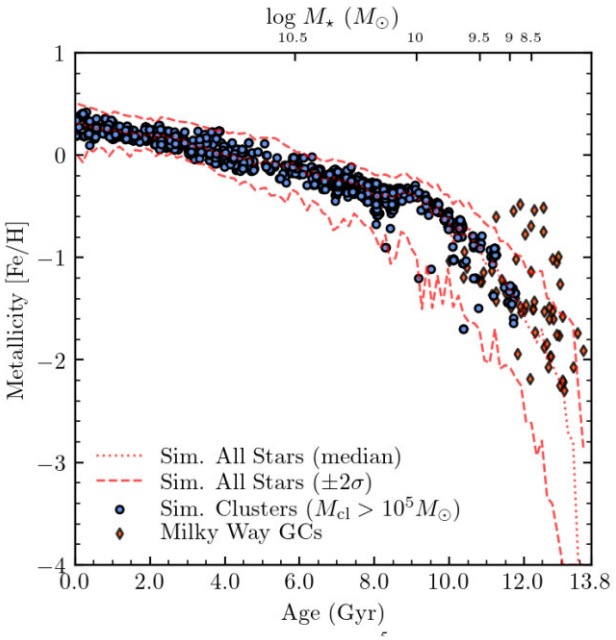

Figure 13. Age–metallicity relation of $>10^5$ $M_\odot$ clusters formed in our model (circles), the entire stellar population of the simulation (red lines), and Milky Way globular clusters (diamonds, data compiled by Kruijssen et al. 2019a). On top we mark the simulated main galactic host stellar mass at various times.

Fall & Chandar 2012), but the proportionality factor varies with environment. Hence, the mass–radius relation can be described by an environmentally varying 3D density, with intrinsic scatter, and the shallower relation of the aggregate sample emerging because more massive clusters tend to form in denser GMCs, and hence tend to be smaller. In Fig. 12, we plot the number-weighted median and $16-84$ per cent range of the 3D half-mass density $\rho_{\rm eff} = 3M_{\rm cl}/\left(8\pi r_{\rm eff}^3\right)$, where $r_{\rm eff}$ is the 3D half-mass radius, as a function of $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$, compared with various observations. Cluster sizes and masses in different subregions of M83 are taken from Ryon et al. (2015), and $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ from Adamo et al. (2015). Sizes, masses, and environmental properties in different M31 PHAT fields are taken from Johnson et al. (2015, 2016, 2017), respectively. Cluster masses in M51 and NGC628 are taken from the data compilation and density profile fits performed by Brown & Gnedin (2021) on data from the LEGUS survey (Calzetti et al. 2015; Adamo et al. 2017; Cook et al. 2019), and radially binned $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ from Messa et al. (2018b) and Chevance et al. (2020) for M51 and NGC628, respectively. Lastly, we use data for M82, NGC253, and the Milky Way central molecular zone (CMZ) compiled by Choksi & Kruijssen (2021).

Fig. 12 shows that the median cluster density scales systematically with both $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ in our model, and we find ± 1.1dex of residual scatter about the median. A similar trend in cluster density is also found in the observational data, as was noted by Choksi & Kruijssen (2021). Our model consistently overpredicts the median density of clusters compared to the observed clusters, but we note that our model predicts *initial* cluster densities, while the observations are of $\sim 1-100$Myr old clusters. These clusters have had time to lose mass and expand under the influence of stellar evolution and dynamical evolution, so we expect observed evolved clusters to be less dense than the predicted initial density. The *scatter* in initial density is considerably greater than the scatter in density of observed clusters, but again, we expect that evolutionary processes will tend to reduce the scatter in the densities of a cluster population: clusters that are initially 'too small' will tend to puff up due to internal evolution, while clusters that are initially 'too large' are more susceptible to stripping, shocking, and destruction in the galactic environment. It will be possible to examine this hypothesis by modelling the dynamical evolution of each cluster (Rodriguez et al. 2022).

3.4 Age–metallicity relation

Finally, we analyse the age–metallicity relation of candidate globular clusters, which we take to be clusters more massive than 10^5 $M_\odot$ for our present purposes. The relation between the ages and metallicities of ancient star clusters in a galaxy contains information about the galaxy's formation history: each cluster surviving to the present day provides a snapshot of the metallicity of the environment in which it formed, with an associated timestamp. The cluster metallicity can be related to a certain stellar mass of the host galaxy, via the redshift-dependent mass–metallicity relation (e.g. Tremonti et al. 2004; Mannucci et al. 2009; Kirby et al. 2013; Ma et al. 2016), so provided the metallicities of old globular clusters reflect those of their host galaxy as whole, they place constraints upon the stellar mass of the progenitor galaxy at a certain time.

Within the model, clusters inherit their abundances from their progenitor cloud in the simulation, and the simulation itself includes a model for turbulent mixing in the ISM that gives realistic metallicity variations in the galaxy (Escala et al. 2018; Bellardini et al. 2021). The age–metallicity relation of massive clusters in our model, and of stars in the galaxy as a whole, are plotted in Fig. 13, which we compare with data for Milky Way globular clusters compiled in Kruijssen et al. (2019b). Our main result is that *massive clusters do not form with a metallicity substantially different from other stars forming within the galaxy*, i.e. massive cluster formation is not strongly biased toward more or less metal-rich regions. Hence, if globular clusters formed as a result of the normal star formation

process at high redshift,[4] their age–metallicity statistics would trace the overall properties of the progenitor galaxies faithfully.

Fig. 13 also suggests that our simulated cluster population is not a good model for the globular cluster population of the Milky Way: the first $>10^5$ $M_\odot$ cluster forms at ~ 2 Gyr ($z \sim 3$), and at this time a significant number of globular clusters should have already formed, in the Milky Way and in other galaxies (Beasley et al. 2000; Woodley & Gómez 2010; VandenBerg et al. 2013; Usher et al. 2019). Moreover, even if massive clusters formed sooner, the age–metallicity relation in the model cannot reproduce the sequence of old, red (metal-rich) globular clusters, which are believed to be the population that formed *in situ* in the Milky Way (Forbes & Bridges 2010; Kruijssen et al. 2020). The galaxy would have to be significantly more enriched at early times to host an old, red population.

4 DISCUSSION

4.1 Does cluster formation efficiency vary with environment?

Many observational works have argued that CFE Γ does vary with galactic environment (Bastian 2008; Goddard et al. 2010; Adamo et al. 2015; Johnson et al. 2016), and many ensuing theoretical works have found that this is to be expected from the physics of star formation (Kruijssen 2012; Li et al. 2017; Pfeffer et al. 2018; Lahén et al. 2019). On the other hand, Chandar et al. (2017) argued that some or all of the scaling in Γ that other works inferred could be explained by contamination of the cluster sample by young ($\lesssim 10$Myr), unbound systems, calling the scaling of Γ into some question.

In Section 3.1.1, we found that denser (higher $\Sigma_{\rm gas}$) and more actively star-forming (higher $\Sigma_{\rm SFR}$) regions host systematically denser self-gravitating GMCs (with higher $\Sigma_{\rm GMC}$, see Figs 5 and 6). In turn, our GMC-scale model predicts that the denser GMCs in these regions form stars more efficiently, resulting in higher Γ in that region. Hence, we concur with the growing consensus of theoretical predictions of variable Γ, and have put it on firmer footing using simulations with a self-consistent GMC population formed from cosmological initial conditions. With that said, we do concur with Chandar et al. (2017) that reliable estimates of Γ without stellar kinematic information are only possible for cluster age ranges that (1) are too old to not be gravitationally bound and (2) are too young to have experienced significant mass-loss and disruption, and caution against the overinterpretation of Γ measurements from cluster populations that may not satisfy these criteria (e.g. Adamo et al. 2020b).

More generally, given the modern understanding of the star cluster formation process, it is increasingly difficult to imagine a scenario wherein Γ does *not* vary with environment: Γ has been extensively shown to correlate with the local SFE of the host GMC, in analytic theory (Hills 1980; Mathieu 1983), idealized stellar dynamics calculations modelling gas removal (Tutukov 1978; Lada, Margulis & Dearborn 1984; Baumgardt & Kroupa 2007; Smith et al. 2011, 2013), and hydrodynamics simulations with spatially resolved star and star cluster formation and gas removal by stellar feedback (Lahén et al. 2019; Li et al. 2019; Grudić et al. 2021b). Star formation efficiency, in turn, has been predicted to vary with GMC properties, a prediction that follows from a very general considerations of limiting cases of momentum- and energy-conserving feedback (Fall, Krumholz & Matzner 2010; Krumholz et al. 2019), which has been almost unanimously supported by GMC simulations that treat stellar feedback and simulate a range of GMC properties (Hopkins et al. 2012; Dale et al. 2014; Geen, Soler & Hennebelle 2017; Howard, Pudritz & Harris 2017; Grudić et al. 2018a; Kim, Kim & Ostriker 2018b; Fukushima & Yajima 2021; Kim et al. 2021).[5] The missing link up to this point has been the relation between GMC properties and the $\sim$ kpc-scale quantities $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$, which has not been possible to study in nearby galaxies in a homogeneous fashion. However, in this work we have shown that GMC properties are coupled to environmental properties, so Γ follows in turn.

4.2 Comparison with previous cosmological star cluster formation studies

4.2.1 E-MOSAICS

E-MOSAICS (Pfeffer et al. 2018) is a suite of simulations coupling semi-analytical cluster formation and evolution prescriptions to the EAGLE cosmological hydrodynamics simulations (Schaye et al. 2015). Unlike the FIRE-2 simulation used in this work, E-MOSAICS simulations do not explicitly resolve GMCs and the multiphase ISM, relying instead on a sub-grid 'effective equation of state' prescription to model the dynamics of the ISM (Springel & Hernquist 2003), and using the Reina-Campos & Kruijssen (2017) prescription to model the GMC population according to coarse-grained (kpc-scale) ISM properties. The GMC mass function is then mapped on to the cluster mass function assuming a constant SFE of 10 per cent and a CFE derived from a local formulation of the (Kruijssen 2012) model. These simulations also model the ongoing evolution of clusters on-the-fly in the simulation, accounting for stellar evolution and a variety internal and external dynamical processes, which we have not attempted here (see however Rodriguez et al. 2022, in which we model cluster evolution in post-processing). This makes it possible to comment on the age distribution of young clusters (which is affected by mass-loss and disruption), as well as the population of clusters surviving to $z = 0$, in a large sample of simulated galaxies.

Our model and the Kruijssen (2012) model agree fairly well on the environmental dependence of Γ (Fig. 4), so the prescription used in E-MOSAICS appears to be a reasonably good approximation of our findings derived from explicitly resolved ISM structures. However, the assumption of constant SFE does not agree with the consensus of numerical simulations with stellar feedback (see references in Section 4.1), including the G21 cluster formation model we have used here, in which SFE scales as a function of $\Sigma_{\rm GMC}$. The assumed constant value of 10 per cent may be a reasonable average value weighted by stellar mass formed, but we expect it to vary with environment, given the environmental variations in $\Sigma_{\rm GMC}$ we find here. For example, the cloud shown in Fig. 3 has an overall SFE of 27 per cent according to our model. This may weight massive cluster formation more heavily toward regions of denser ISM.

Inspection of the GMC population modelled in E-MOSAICS according to the prescription of Reina-Campos & Kruijssen (2017)

[4]Normal galactic star formation need not account for *all* globular clusters: more-exotic mechanisms of extragalactic globular cluster formation have been proposed, e.g. Peebles & Dicke (1968) and more recently Naoz & Narayan (2014).

[5]Note that the agreement of different simulations on this issue is only *qualitative* at present – the SFE predicted for a given GMC model still varies widely between simulation suites, in part due to the variety of prescriptions in use for unresolved star formation and feedback, their chief uncertainty (Grudić & Hopkins 2019).

also reveals some discrepancies with the GMC population found in FIRE simulations by Guszejnov et al. (2020a). Their model hypothesizes that the largest possible collapsing gas mass is of the order of the Toomre mass, and when this is applied to the E-MOSAICS simulations it predicts the existence of self-gravitating clouds in excess of 10^{11} $M_\odot$ (see Pfeffer et al. 2018, fig. 5). In comparison, the most massive self-gravitating gas structure formed self-consistently in the Guszejnov et al. (2020a) catalogue we have used is 2×10^8 $M_\odot$. Even if this mass is to be identified only with the 'collapsed fraction' that is identified as the cluster progenitor cloud in the model, E-MOSAICS simulations host numerous clouds with $M_{\rm GMC} > 10^{10}$ $M_\odot$. Li et al. (2020) pointed out that the properties of GMCs formed in FIRE-like simulations with resolved ISM structure can be still somewhat sensitive to adopted sub-grid feedback and/or star formation prescriptions, so we do not necessarily consider the Guszejnov et al. (2020a) cloud properties definitive, but the mass function variation seen in Li et al. (2020) was not at the level needed to explain a cloud mass discrepancy of 2 orders of magnitude. Thus, at present there appears to be a disconnect between the semi-analytical theory of GMC mass functions applied to the EAGLE simulations, and what is found in numerical simulations with explicit ISM structure.

E-MOSAICS simulations assumed a constant initial cluster radius, surveying various values $r_{\rm eff} = 1.5 - 6$pc and adopting a fidicial value of 4pc. As noted in Choksi & Kruijssen (2021) and this work (Section 3.3), more recently available data show evidence of a variable mass–radius relation, taking the form $R_{\rm eff} \propto M_{\rm cl}^{1/3}$, with a varying proportionality factor (e.g. Fig. 11). Adopting such a relation would make low-mass clusters smaller and high-mass clusters larger, which would affect their susceptibility to the tidal environment in turn. However, because the relation is shallow we expect the scaling relation itself to have modest effects, as shown by Pfeffer et al. (2018). Likely more important is the significant *scatter* found in simulations and observations: this could significantly broaden the range of cluster sizes, and the resulting range of possible dynamical histories.

Lastly, the E-MOSAICS simulations have been used to predict and interpret the age–metallicity relation of globular clusters, with a large sample size of Milky Way-mass galaxies (Kruijssen et al. 2019a, b). These works do find galaxies that fill the region of age–metallicity space occupied by the Milky Way's globular clusters (cf. Fig. 13), but this appears to lie at the upper envelope of the range spanned by the different simulations – simulations that form massive clusters relatively late like ours appear to be common in their sample as well.

4.2.2 *Li et al. ART simulations*

In a series of studies, Li et al. (2017), Li, Gnedin & Gnedin (2018), and Li & Gnedin (2019) performed a suite of cosmological zoom-in simulations of Milky Way-mass galaxy progenitors, run with the ART adaptive mesh refinement code (Kravtsov, Klypin & Khokhlov 1997; Agertz et al. 2013; Semenov et al. 2016). Like ours, their simulations did marginally resolve the multiphase ISM, with a spatial resolution of 6pc, so they were able to model the formation of individual GMCs, and cluster formation in turn, modelling cluster formation as a process of accretion and feedback with various subgrid physics prescriptions.

Qualitatively, all of the conclusions reached in these works concerning CFE and the initial mass function of star clusters agree with ours: denser galactic environments produce more top-heavy mass distributions of clusters, with higher efficiency. In particular, Li et al. (2017) correlate the mass function and cluster efficiency with *merger* activity specifically, with mergers leading to more efficient cluster formation. Quantitatively, the predictions of the Li et al. simulations depend very sensitively upon the assumed sub-grid star formation efficiency (Li et al. 2018), with lower subgrid efficiencies resulting in lower cluster masses. No $\Sigma_{\rm SFR}$–Γ relation presented in that work matches ours especially well in all environments, as the relation is generally shallow compared to ours.

Although the cluster initial mass function found in Li et al. (2017) tended to be quite Schechter-like with a typical slope of ~ -2, the nominally improved Li et al. (2018) suite found a relatively steep (slope between -2 and -3) mass function, similar to the mass function typically found in our model (Section 3.2). Observed mass function slopes have are typically around ~ -2 (e.g. Krumholz et al. 2019, fig. 5), but these can be affected by resolution and completeness effects, and in Fig. 7 we do find fair agreement with the shapes of cluster mass functions derived from catalogues in nearby galaxies (e.g. Adamo et al. 2015; Messa et al. 2018a).

4.2.3 *FIRE simulations*

Kim et al. (2018a) and Ma et al. (2020) used the FIRE and FIRE-2 frameworks, respectively, to model the formation of bound star clusters *on-the-fly* in the simulations at high redshift, in contrast to the post-processed approach explored here. Those simulations arrived at similar conclusions to us regarding the formation mechanism of the most massive clusters: the sites of massive bound cluster formation were found to be very high pressure and/or surface density ($\gtrsim 10^4\,{\rm M_\odot pc^{-2}}$) (similar to e.g. the scenario shown in Fig. 3), achieving high star formation efficiency. Notably, these simulations directly demonstrated that it is possible to achieve such conditions at $z \gtrsim 5$, despite the lack of massive clusters forming at that time in this work.

In Ma et al. (2020) in particular we emphasized that the results of this type of simulation were sensitive to the choice of star formation prescription. To further extend the predictive power of cosmological simulations to detailed predictions of cluster properties on-the-fly, an SF prescription that resolves inherent uncertainties about star formation on small scales is needed. Progress on this front may now be possible by comparing with GMC simulations with individual self-consistent star formation simulations that overlap with the GMC masses that are marginally resolvable in galaxy simulations (Guszejnov et al. 2020b, 2021; Grudić et al. 2021a).

4.3 Differences from the Milky Way's globular cluster population

In Section 3.4, we noted important differences between the age and metallicity statistics of the simulated cluster population and the Milky Way globular cluster population: our model produces no $>10^5$ $M_\odot$ bound clusters in the first 2 Gyr ($z > 3$), and does not reproduce the 'red' population of old, metal-enriched globular clusters. The most obvious explanation for this is that the simulated galaxy's star formation history is so different from that of the Milky Way: as mentioned in Section 3, the simulated galaxy has similar $z = 0$ stellar mass but $\sim 5\times$ higher $z \sim 0$ SFR than the Milky Way, in part due to its relatively high gas fraction of 20 per cent (a common feature in FIRE-2 Milky Way-like galaxies, see Gurvich et al. 2020). To attain similar $z = 0$ mass this way, its SFR had to be lower than the MW at early times. The mean SFR of the Milky Way progenitor in the first 2 Gyr has been inferred to be $\sim 5\,{\rm M_\odot yr^{-1}}$ (Snaith et al. 2014), much greater than the mean $1\,{\rm M_\odot\,yr^{-1}}$ in the first 2 Gyr of

our simulation (Fig. 2). If the SFR was as high as the Milky Way, but concentrated in the same area, the average value of $\Sigma_{\rm SFR}$ would be $\sim 5\times$ greater, increasing the CFE by a factor of 1.6 (equation 8) and the upper cut-off of the cluster mass function by a factor of ~ 6 (e.g. Fig. 10), allowing massive clusters to form much sooner. Shifting star formation from late to early times would also make the model more Milky Way-like by suppressing the mass scale of the cluster mass function at late times, which typically has a truncation of $\sim 10^5$ $M_\odot$ at $z \sim 0$ (Fig. 7), more massive than the most massive young clusters in the Milky Way (several 10^4 $M_\odot$ at most, Portegies Zwart, McMillan & Gieles 2010).

Santistevan et al. (2020) surveyed the star formation histories of Milky Way-mass galaxies in wider FIRE-2 simulation suite, and found that galaxies in Local Group analogues with two Milky Way-mass galaxies in close proximity form preferentially earlier than isolated Milky Way-mass galaxies like the one we have considered here. Thus, environment may be a factor that differentiates the star formation history of the present model from that of the Milky Way. However, the maximum mass achieved by any galaxy at the 2 Gyr mark in Santistevan et al. (2020) was $\sim 4 \times 10^9$ $M_\odot$, so these other galaxies would still have difficulty achieving the SFR intensity and metallicity needed to reproduce the old, red GC population. However, they note how various constraints suggest that a star formation history more like the one simulated here – reaching 50 per cent of the $z = 0$ stellar mass at $z \sim 1$ – may be typical among galaxies with $M_{200} \sim 10^{12}$ $M_\odot$ (e.g. Behroozi et al. 2019). If so, the galaxy simulated here – and its cluster population – may be more representative of a typical galaxy of this mass, and we would expect the Milky Way's GC population to be systematically ($\sim 2-3$Gyr) older than a typical galaxy of this mass.

Even if old, massive globular clusters are produced, old *red* ([Fe/H] ~ -0.5) globular clusters may be difficult to obtain within our framework, even if the early star formation was more rapid. Let us assume the redshift-dependent relation between galactic stellar mass and gas metallicity found in Ma et al. (2016):

$$\log\left(Z_{\rm gas}/Z_\odot\right) = 0.35\log\left(\frac{M_\star}{10^{10}\,{\rm M}_\odot}\right) + 0.93\exp\left(-0.43z\right) - 1.05, \quad (12)$$

and assume that newborn stars and clusters inherent this gas metallicity at a given redshift. Then a hypothetical galaxy that averaged $5\,M_\odot{\rm yr}^{-1}$ in the first 2Gyr would only form clusters with [Fe/H] ~ -1, still half a dex less than the red population. However, the redshift dependence predicted for the mass–metallicity relation does depend crucially upon uncertain feedback and stellar physics (Agertz et al. 2020), so it is possible that the FIRE simulations used to fit the Ma et al. (2016) relation err in the direction of underestimating metal retention.

Lastly, it is also worth emphasizing here that *forming* the clusters is a necessary, but not sufficient condition for obtaining the population at $z = 0$ – once formed, the clusters are subject to mass-loss and disruption in the galactic environment. Detailed predictions of the surviving $z = 0$ GC population require a treatment of the dynamical evolution of clusters in the galactic environment, which has been performed in other simulation setups (Li et al. 2017; Pfeffer et al. 2018), and which we defer to future work for the present model (Rodriguez et al. 2022).

5 CONCLUSIONS

In this work, we have modelled the population of young star clusters forming in a simulated Milky Way-mass galaxy, extending the predictions of Grudić et al. (2021b) for cluster formation in individual GMCs to the population of cluster progenitor clouds that form self-consistently across cosmic time in the simulation (Guszejnov et al. 2020a). We used this model to study various aspects of the star cluster formation:

(i) The efficiency of bound cluster formation Γ is 13 per cent in the simulated galaxy. The efficiency does not exhibit a clear systematic trend with cosmic time, but can vary over a wide range at different periods of the galaxy's history (Fig. 2). Much of this variation is explained by variations in galactic ISM conditions: there is clear relation between Γ and the local $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ (Fig. 4), as measured on $\sim$ 1kpc scales in the galaxy. This is because these quantities correlate with the surface density of self-gravitating GMCs (Figs 5 and 6), which determines star and star CFE in turn, according to the G21 model. The environmental scalings we found appear to reproduce the successes (and possible failures, e.g. Messa et al. 2018a) of the Kruijssen (2012) model.

(ii) The initial mass function of bound star clusters shows significant diversity over different periods of the galaxy's evolution, similar to the range of diversity seen in observations of nearby galaxies (Fig. 7). Both the shape and the normalization of the mass function vary intrinsically, and overall the mass function is not described well by any one simple power-law or Schechter-like form (Fig. 8). This sequence of mass function shapes is similar to what is observed in nearby galaxies, and is driven at least in part by environment: denser environments host denser GMCs, which can form stars more efficiently and produce more massive clusters (Figs 9,10).

(iii) We find a global, time-integrated size–mass relation for star clusters of $R_{\rm eff} \propto M_{\rm cl}^{0.25}$, similar to the relation inferred from recent star cluster catalogues (Brown & Gnedin 2021; Choksi & Kruijssen 2021). Within a given environment of fixed $\Sigma_{\rm gas}$ or $\Sigma_{\rm SFR}$, the relation is best described by $R_{\rm eff} \propto M_{\rm cl}^{1/3}$, i.e. constant 3D density, but this density varies with environment (Fig. 12), leading to a global relation shallower than $\propto M_{\rm cl}^{1/3}$. Within a given environment we also predict a significant initial *scatter* in initial cluster density of $\sim$ 1.1dex. This is larger than what is observed in $\sim$ 100Myr old cluster populations, suggesting that star formation physics alone cannot explain the size-mass relation: evolutionary processes must be invoked to reduce the scatter.

(iv) The age–metallicity relation of massive ($>10^5$ $M_\odot$) bound star clusters formed in the galaxy is very similar to that of the stellar population as a whole. Our age–metallicity statistics were incompatible of those of Milky Way globular clusters (Fig. 13), but it seems plausible that this difference is driven by a difference in star formation histories (Section 4.3), which affect cluster formation via the local scaling relations with e.g. $\Sigma_{\rm SFR}$ we have found.

Thus we have been able to study the properties of *young* star clusters as they vary across cosmic time and galactic environment. To model populations of *evolved* clusters, and old globular clusters in particular, we must extend our model, accounting for stellar evolution, dynamical evolution, and the influence of the surrounding galactic environment (e.g. Pfeffer et al. 2018). This will be the subject of our follow-up work (Rodriguez et al. 2022).

The major caveat of this work is that both steps of our model – predicting galactic ISM structure and mapping those structures on to star clusters – are not yet fully solved problems. Attempts to do either in a systematic fashion are still relatively new, and invariably rely upon ad hoc models for unresolved star formation, turbulence, and feedback. As such, we anticipate that detailed predictions of ISM structure and star cluster formation will continue to evolve as the unresolved microphysics of star formation become better understood.

ACKNOWLEDGEMENTS

We thank S. Michael Fall, J. M. Diederik Kruijssen, Angela Adamo, Marta Reina-Campos, Hui Li, Andrey Kravtsov, Nick Gnedin, Gillen Brown, Omid Sameie, and Anna Schauer for enlightening discussions that informed and motivated this work. MYG was supported by a CIERA Postdoctoral Fellowship and a NASA Hubble Fellowship (award HST-HF2-51479). This work was supported by NSF Grant AST-2009916 at Carnegie Mellon University and a New Investigator Research Grant to CR from the Charles E. Kaufman Foundation. AW received support from: NSF via CAREER award AST-2045928 and grant AST-2107772; NASA ATP grant 80NSSC20K0513; HST grants AR-15809, GO-15902, GO-16273 from STScI. MBK acknowledges support from NSF CAREER award AST-1752913, NSF grants AST-1910346 and AST-2108962, NASA grant NNX17AG29G, and HST-AR-15006, HST-AR-15809, HST-GO-15658, HST-GO-15901, HST-GO-15902, HST-AR-16159, and HST-GO-16226 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555. AL is supported by the Programme National des Hautes Energies and ANR COSMERGE project, grant ANR-20-CE31-001 of the French Agence Nationale de la Recherche. This work used computational resources provided by XSEDE allocation AST-190018 and TACC Frontera allocations AST-20019 and AST-21002. CAFG was supported by NSF through grants AST-1715216, AST-2108230, and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant HST-AR-16124.001-A; and by the Research Corporation for Science Advancement through a Cottrell Scholar Award. Fig. 3 was generated with the help of FIRE studio (Gurvich 2021), an open source PYTHON visualization package designed with the FIRE simulations in mind.

DATA AVAILABILITY

The data supporting the plots within this article are available upon request to the corresponding author. A public version of the GIZMO code is available at http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html. The CloudPhinder code used to catalogue self-gravitating clouds in the simulation (2.2) is available at http://www.github.com/mikegrudic/CloudPhinder.

REFERENCES

Adamo A. et al., 2017, ApJ, 841, 131
Adamo A. et al., 2020a, Space Sci. Rev., 216, 69
Adamo A. et al., 2020b, MNRAS, 499, 3267
Adamo A., Kruijssen J. M. D., Bastian N., Silva-Villa E., Ryon J., 2015, MNRAS, 452, 246
Agertz O. et al., 2020, MNRAS, 491, 1656
Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
Bastian N. et al., 2012, MNRAS, 419, 2606
Bastian N., 2008, MNRAS, 390, 759
Bastian N., Schweizer F., Goudfrooij P., Larsen S. S., Kissler-Patig M., 2013, MNRAS, 431, 1252
Bate M. R., Bonnell I. A., Bromm V., 2003, MNRAS, 339, 577
Baumgardt H., Kroupa P., 2007, MNRAS, 380, 1589
Beasley M. A., Sharples R. M., Bridges T. J., Hanes D. A., Zepf S. E., Ashman K. M., Geisler D., 2000, MNRAS, 318, 1249
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Bellardini M. A., Wetzel A., Loebman S. R., Faucher-Giguère C.-A., Ma X., Feldmann R., 2021, MNRAS, 505, 4586
Benincasa S. M. et al., 2020, MNRAS, 497, 3993
Bertoldi F., McKee C. F., 1992, ApJ, 395, 140
Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529
Bolatto A. D., Leroy A. K., Rosolowsky E., Walter F., Blitz L., 2008, ApJ, 686, 948
Bovy J., 2017, MNRAS, 470, 1360
Brown G., Gnedin O. Y., 2021, MNRAS, 508, 5935
Calzetti D. et al., 2015, AJ, 149, 51
Capuzzo-Dolcetta R., Miocchi P., 2008, MNRAS, 388, L69
Chandar R., Fall S. M., Whitmore B. C., Mulia A. J., 2017, ApJ, 849, 128
Chevance M. et al., 2020, MNRAS, 493, 2872
Choksi N., Kruijssen J. M. D., 2021, MNRAS, 507, 5492
Colombo D. et al., 2014, ApJ, 784, 3
Cook D. O. et al., 2012, ApJ, 751, 100
Cook D. O. et al., 2019, MNRAS, 484, 4897
Dale J. E., 2017, MNRAS, 467, 1067
Dale J. E., Ngoumou J., Ercolano B., Bonnell I. A., 2014, MNRAS, 442, 694
Elmegreen B. G., 2018, ApJ, 854, 16
Elson R. A. W., Fall S. M., Freeman K. C., 1987, ApJ, 323, 54
Escala I. et al., 2018, MNRAS, 474, 2194
Faesi C. M., Lada C. J., Forbrich J., 2018, ApJ, 857, 19
Fall S. M., Chandar R., 2012, ApJ, 752, 96
Fall S. M., Krumholz M. R., Matzner C. D., 2010, ApJ, 710, L142
Fall S. M., Zhang Q., 2001, ApJ, 561, 751
Forbes D. A., Bridges T., 2010, MNRAS, 404, 1203
Freeman P., Rosolowsky E., Kruijssen J. M. D., Bastian N., Adamo A., 2017, MNRAS, 468, 1769
Fukushima H., Yajima H., 2021, MNRAS, 506, 5512
Geen S., Soler J. D., Hennebelle P., 2017, MNRAS, 471, 4844
Gieles M., Larsen S. S., Scheepmaker R. A., Bastian N., Haas M. R., Lamers H. J. G. L. M., 2006b, A&A, 446, L9
Gieles M., Portegies Zwart S. F., Baumgardt H., Athanassoula E., Lamers H. J. G. L. M., Sipior M., Leenaarts J., 2006a, MNRAS, 371, 793
Gieles M., Renaud F., 2016, MNRAS, 463, L103
Ginsburg A., Kruijssen J. M. D., 2018, ApJ, 864, L17
Gnedin O. Y., Ostriker J. P., Tremaine S., 2014, ApJ, 785, 71
Goddard Q. E., Bastian N., Kennicutt R. C., 2010, MNRAS, 405, 857
Gouliermis D. A., 2018, PASP, 130, 072001
Grudić M. Y., Guszejnov D., Hopkins P. F., Lamberts A., Boylan-Kolchin M., Murray N., Schmitz D., 2018b, MNRAS, 481, 688
Grudić M. Y., Guszejnov D., Hopkins P. F., Offner S. S. R., Faucher-Giguère C.-A., 2021a, MNRAS, 506, 2199
Grudić M. Y., Hopkins P. F., 2019, MNRAS, 488, 2970
Grudić M. Y., Hopkins P. F., Faucher-Giguère C.-A., Quataert E., Murray N., Kereš D., 2018a, MNRAS, 475, 3511
Grudić M. Y., Hopkins P. F., Lee E. J., Murray N., Faucher-Giguère C.-A., Johnson L. C., 2019, MNRAS, 488, 1501
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 (G21)
Gurvich A. B. et al., 2020, MNRAS, 498, 3664
Gurvich A. B., 2022, Astrophysics Source Code Library, record ascl:2202.006
Guszejnov D., Grudić M. Y., Hopkins P. F., Offner S. S. R., Faucher-Giguère C.-A., 2020b, MNRAS, 496, 5072
Guszejnov D., Grudić M. Y., Hopkins P. F., Offner S. S. R., Faucher-Giguère C.-A., 2021, MNRAS, 502, 3646
Guszejnov D., Grudić M. Y., Offner S. S. R., Boylan-Kolchin M., Faucher-Gigère C.-A., Wetzel A., Benincasa S. M., Loebman S., 2020a, MNRAS, 492, 488
Haugbølle T., Padoan P., Nordlund Å., 2018, ApJ, 854, 35
Heiderman A., Evans II N. J., Allen L. E., Huard T., Heyer M., 2010, ApJ, 723, 1019
Hills J. G., 1980, ApJ, 235, 986
Hislop J. M., Naab T., Steinwandel U. P., Lahén N., Irodotou D., Johansson P. H., Walch S., 2022, MNRAS, 509, 5938
Hollyhead K., Adamo A., Bastian N., Gieles M., Ryon J. E., 2016, MNRAS, 460, 2087
Hopkins P. F. et al., 2018a, MNRAS, 477, 1578
Hopkins P. F. et al., 2018b, MNRAS, 480, 800
Hopkins P. F. et al., 2020b, MNRAS, 492, 3465
Hopkins P. F., 2015, MNRAS, 450, 53

Hopkins P. F., 2017, MNRAS, 466, 3387
Hopkins P. F., Grudić M. Y., Wetzel A., Kereš D., Faucher-Giguère C.-A., Ma X., Murray N., Butcher N., 2020a, MNRAS, 491, 3702
Hopkins P. F., Keres D., Onorbe J., Faucher-Giguere C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Hopkins P. F., Quataert E., Murray N., 2012, MNRAS, 421, 3488
Hopkins P. F., Raives M. J., 2016, MNRAS, 455, 51
Howard C. S., Pudritz R. E., Harris W. E., 2017, MNRAS, 470, 3346
Hunter D. A., Elmegreen B. G., Dupuy T. J., Mortonson M., 2003, AJ, 126, 1836
Johnson L. C. et al., 2015, ApJ, 802, 127
Johnson L. C. et al., 2016, ApJ, 827, 33
Johnson L. C. et al., 2017, ApJ, 839, 78
Kennicutt R. C. Jr, 1998, ARA&A, 36, 189
Kim J.-h. et al., 2018a, MNRAS, 474, 4232
Kim J.-G., Kim W.-T., Ostriker E. C., 2018b, ApJ, 859, 68
Kim J.-G., Ostriker E. C., Filippova N., 2021, ApJ, 911, 128
Kirby E. N., Cohen J. G., Guhathakurta P., Cheng L., Bullock J. S., Gallazzi A., 2013, ApJ, 779, 102
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Kroupa P., 2001, MNRAS, 322, 231
Kruijssen J. M. D. et al., 2020, MNRAS, 498, 2472
Kruijssen J. M. D., 2012, MNRAS, 426, 3008 (K12)
Kruijssen J. M. D., Pelupessy F. I., Lamers H. J. G. L. M., Portegies Zwart S. F., Icke V., 2011, MNRAS, 414, 1339
Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019a, MNRAS, 486, 3134
Kruijssen J. M. D., Pfeffer J. L., Reina-Campos M., Crain R. A., Bastian N., 2019b, MNRAS, 486, 3180
Krumholz M. R., Klein R. I., McKee C. F., 2011, ApJ, 740, 74
Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227
Lada C. J., Dame T. M., 2020, ApJ, 898, 3
Lada C. J., Margulis M., Dearborn D., 1984, ApJ, 285, 141
Lahén N., Naab T., Johansson P. H., Elmegreen B., Hu C.-Y., Walch S., 2019, ApJ, 879, L18
Larson R. B., 1981, MNRAS, 194, 809
Leitherer C. et al., 1999, ApJS, 123, 3
Li H., Gnedin O. Y., 2019, MNRAS, 486, 4030
Li H., Gnedin O. Y., Gnedin N. Y., 2018, ApJ, 861, 107
Li H., Gnedin O. Y., Gnedin N. Y., Meng X., Semenov V. A., Kravtsov A. V., 2017, ApJ, 834, 69
Li H., Vogelsberger M., Marinacci F., Gnedin O. Y., 2019, MNRAS, 487, 364
Li H., Vogelsberger M., Marinacci F., Sales L. V., Torrey P., 2020, MNRAS, 499, 5862
Licquia T. C., Newman J. A., 2015, ApJ, 806, 96
Ma X. et al., 2020, MNRAS, 493, 4315
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
Mannucci F. et al., 2009, MNRAS, 398, 1915
Maraston C., Bastian N., Saglia R. P., Kissler-Patig M., Schweizer F., Goudfrooij P., 2004, A&A, 416, 467
Mathieu R. D., 1983, ApJ, 267, L97
Messa M. et al., 2018a, MNRAS, 473, 996
Messa M. et al., 2018b, MNRAS, 477, 1683
Mok A., Chandar R., Fall S. M., 2019, ApJ, 872, 93
Mulia A. J., Chandar R., Whitmore B. C., 2016, ApJ, 826, 32
Naoz S., Narayan R., 2014, ApJ, 791, L8
Padoan P., Haugbølle T., Nordlund Å., 2012, ApJ, 759, L27
Peebles P. J. E., Dicke R. H., 1968, ApJ, 154, 891
Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, MNRAS, 475, 4309
Pokhrel R. et al., 2021, ApJ, 912, L19
Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, ARA&A, 48, 431
Reina-Campos M., Kruijssen J. M. D., 2017, MNRAS, 469, 1282
Rodriguez C. L. et al., 2020, ApJ, 896, L10
Rodriguez C. L., Hafen Z., Grudić M. Y., Lamberts A., Sharma K., Faucher-Giguère C. A., Wetzel A., 2022, preprint (arXiv:2203.16547)
Ryon J. E. et al., 2015, MNRAS, 452, 525
Sanderson R. E. et al., 2020, ApJS, 246, 6
Santistevan I. B., Wetzel A., El-Badry K., Bland-Hawthorn J., Boylan-Kolchin M., Bailin J., Faucher-Giguère C.-A., Benincasa S., 2020, MNRAS, 497, 747
Schaye J. et al., 2015, MNRAS, 446, 521
Schmidt M., 1959, ApJ, 129, 243
Semenov V. A., Kravtsov A. V., Gnedin N. Y., 2016, ApJ, 826, 200
Smith R., Goodwin S., Fellhauer M., Assmann P., 2013, MNRAS, 428, 1303
Smith R., Slater R., Fellhauer M., Goodwin S., Assmann P., 2011, MNRAS, 416, 383
Snaith O. N., Haywood M., Matteo P. D., Lehnert M. D., Combes F., Katz D., Gómez A., 2014, ApJ, 781, L31
Springel V., Hernquist L., 2003, MNRAS, 339, 289
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Su K.-Y., Hopkins P. F., Hayward C. C., Faucher-Giguère C.-A., Kereš D., Ma X., Robles V. H., 2017, MNRAS, 471, 144
Tremonti C. A. et al., 2004, ApJ, 613, 898
Tutukov A. V., 1978, A&A, 70, 57
Usher C., Brodie J. P., Forbes D. A., Romanowsky A. J., Strader J., Pfeffer J., Bastian N., 2019, MNRAS, 490, 491
VandenBerg D. A., Brogaard K., Leaman R., Casagrande L., 2013, ApJ, 775, 134
Wainer T. M. et al., 2022, ApJ, 928, 15
Ward J. L., Kruijssen J. M. D., 2018, MNRAS, 475, 5659
Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016, ApJ, 827, L23
Whitmore B. C. et al., 2010, AJ, 140, 75
Woodley K. A., Gómez M., 2010, Publ. Astron. Soc. Aust., 27, 379

This paper has been typeset from a TEX/LATEX file prepared by the author.