Cluster assembly and the origin of mass segregation in the STARFORGE simulations

Dávid Guszejnov1*, Carleen Markey2,3, Stella S. R. Offner1, Michael Y. Grudić4, Claude-André Faucher-Giguère4, Anna L. Rosen5, Philip F. Hopkins6,

1Department of Astronomy, University of Texas at Austin, TX 78712, USA
2Department of Physics and Astronomy, 525 Northwestern Avenue, Purdue University, West Lafayette, IN 47907, USA
3Department of Physics | Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA
4CIERA and Department of Physics and Astronomy, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA
5Center for Astrophysics | Harvard & Smithsonian, 60 Garden St, Cambridge, MA 02138, USA
6TAPIR, Mailcode 350-17, California Institute of Technology, Pasadena, CA 91125, USA

ABSTRACT
Stars form in dense, clustered environments, where feedback from newly formed stars eventually ejects the gas, terminating star formation and leaving behind one or more star clusters. Using the STARFORGE simulations, for the first time it is possible to simulate this process in its entirety within a molecular cloud, while explicitly evolving the gas radiation and magnetic fields and following the formation of individual, low-mass stars. We find that individual star-formation sites merge to form ever larger structures, while still accreting gas. Thus clusters are assembled through a series of mergers. During the cluster assembly process a small fraction of stars are ejected from their clusters; we find no significant difference between the mass distribution of the ejected stellar population and that of stars inside clusters. The star-formation sites that are the building blocks of clusters start out mass segregated with one or a few massive stars at their center. As they merge the newly formed clusters maintain this feature, causing them to have mass-segregated substructures without themselves being centrally condensed. The cluster relaxes through dynamical interactions to a centralized configuration, but this process does not finish before feedback expels the remaining gas from the cluster. The gas-free clusters then become unbound and break up. We find that turbulent driving and a periodic cloud geometry can significantly reduce clustering and prevent gas expulsion. Meanwhile, the initial surface density and level of turbulence have little qualitative effect on cluster evolution, despite the significantly different star formation histories.

Keywords: galaxies: star clusters: general – stars: formation – stars: kinematics and dynamics – stars: luminosity function, mass function

1 INTRODUCTION
Stars predominantly form in dense clusters of hundreds to a few $10^5$ stars (Lada & Lada 2003; Bressert et al. 2010), making cluster formation a key part of the star formation process. Newly formed clusters can dissolve due to gas ejection resulting from stellar feedback, internal relaxation, dynamical friction and tidal fields (Krumholz et al. 2019), making the present day observable clusters the surviving members of the original population. Observed bound clusters have historically been categorized as open clusters and globular clusters depending on their location and age, but emerging evidence suggests that these two classes are not different with regards to their formation and internal dynamics but instead experience a different cosmological history (see e.g., Kruisjes 2014 and the review of Krumholz et al. 2019). Unbound clusters are often referred to as stellar associations and are typically found at sites of recent star formation (Gouliermis 2018).

The relatively low number of observed clusters compared to the abundance of star formation sites suggests that most (non-massive) star formation sites create only short-lived clusters (Lada & Lada 2003). Longer-lived bound clusters must require specific star formation histories and initial conditions (see Krumholz & McKee 2020 for details). The exact formation mechanism of clusters within star-forming molecular clouds is not known, despite intense theoretical and observational effort. However, recent observations (e.g., Bressert et al. 2010; Gouliermis 2018) support the idea of hierarchical star formation, where stars form in regions of various densities, prescribed by the underlying hierarchy of ISM structure (e.g., along filaments). Simulations of small star-forming clouds have reproduced this scenario and formed bound star clusters through hierar-
chical assembly, where small sub-clusters merge with their neighbors, eventually forming a massive bound structure (e.g., Bonnell et al. 2003; Grudić et al. 2018; Vázquez-Semadeni et al. 2017).

A key step in the cluster formation process is the onset of stellar feedback that first stops the accretion of individual stars then expels the gas from the cluster. Exactly how this gas expulsion happens has dramatic effects on the future evolution of the cloud (Krause et al. 2020). Violent gas expulsion leads to the quick dissolution of the cluster (i.e., "infant mortality", see Hills 1980; Lada & Lada 2003; Baumgardt & Kroupa 2007; Fall et al. 2010a), however highly substructured clusters may survive even instantaneous gas expulsion (Farias et al. 2018). Recent hydrodynamical simulations have also found indications of gravitational feedback from gas expulsion (Geen et al. 2018; Zamora-Avilés et al. 2019), such that asymmetry in the expelled gas shell produces a net gravitational force on the stars. Gaia measurements have identified several clusters undergoing gas expulsion, which appear to be expanding (Kuhn et al. 2019).

The stellar distribution also provides insights into the initial conditions and past cluster evolution. Many observed star clusters exhibit mass segregation, whereby massive stars are concentrated in the centers of clusters (Hillenbrand & Hartmann 1998; Kirk & Myers 2011). Mass-segregation may be a natural outcome of the star formation process, such that clusters are born segregated (e.g., McKee & Tan 2003; Bonnell & Bate 2006). In this scenario massive stars form at the locations with the highest density gas, such that mass segregation is primordial. Alternatively, star clusters may not be initially non-segregated but become so due to dynamical interactions (Spitzer 1969) that cause massive stars to sink to the bottom of the potential well, i.e., the cluster center. Numerical investigations have been limited by the dynamic range of star formation simulations as the simulation must track the formation of individual stars and model their motions over the cluster relaxation timescale. Thus, works investigating the origin of mass segregation have been constrained to modeling small clusters (e.g., Kirk et al. 2014) and clusters without self-consistent gas treatment (e.g., Parker 2014).

In this paper we present radiation-magneto-hydrodynamic (RMHD) simulations from the STARFORMation in Gaseous Environments (STARFORGE) project1. These simulations follow the evolution of turbulent and magnetized giant molecular clouds (GMCs) from the onset of star formation until it is disrupted by stellar feedback, while also following the formation of individual stars above the H burning limit (for details see Grudić et al. 2021a and Guszejnov et al. 2021, henceforth referred to as Paper I and Paper II). Note that the hydrodynamical simulations previously used to study star cluster formation in the literature had smaller dynamic ranges so they were either restricted to simulating a small clump (e.g., Kirk et al. 2014) or did not follow individual low-mass stars (e.g., Geen et al. 2018; Zamora-Avilés et al. 2019). The STARFORGE simulations follow the assembly of star clusters through gas dispersal, which is modeled self-consistently by including all major feedback processes (i.e., protostellar jets, stellar radiation and winds, supernovae). This allows us to determine whether mass-segregation is primordial and explore the role mergers play in cluster assembly. Note that in this work we focus on the stellar clustering in the simulations. For a detailed analysis on the cloud evolution, star formation history, and stellar mass spectrum see the companion paper Grudić et al. (2022) (henceforth referred to as Paper II).

We briefly summarize the STARFORGE simulations in §2.1.

---

1 http://www.starforge.space
2 http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html
and dust collisional processes. The cooling module self-consistently solves for the internal energy and ionization state of the gas (see Appendix B of Hopkins et al. 2018). The gas adiabatic index is calculated from a fit to density based on the results of Vaidya et al. (2015). The runs in this paper explicitly treat radiation (RHD runs), unlike Paper II. This means co-evolving the gas, dust, and radiation temperature self-consistently as in Hopkins et al. (2020), including the stellar luminosity in various bands accounting for photon transport, absorption and emission using dust opacity. We use a first-moment or M1 (Levermore 1984) RHD solver with a reduced-speed-of-light (RSOL) of 30 km/s and transport photons in 5 distinct bands (IR, optical/NIR, NUV, FUV, and ionizing). Our treatment automatically handles the trapping of cooling radiation in the optically-thick limit. In addition to local sources (i.e., stars) an external heating source is added representing the interstellar radiation field (ISRF) and a zero velocity periodic cubic box with the same temperature prescription as “Sphere” ICs. This periodic box is then “stirred” using the driving algorithm from by Federrath et al. 2010; Bauer & Springel 2012. This involves a spectrum of \( E_k \propto k^{-2} \) for driving modes in Fourier space at wavelengths 1/2 - 1 the box size, with an appropriate decay time for driving mode correlations (\( t_{\text{decay}} \sim t_{\text{cross}} \)). This stirring is initially performed without gravity for five global freefall times \( \left( t_{\text{ff}} = \sqrt{\frac{3 \pi M}{4 \pi G \Sigma}} \right) \), to achieve saturated MHD turbulence. The normalization of the driving spectrum is set so that in equilibrium the gas in the box has a turbulent velocity dispersion that gives the desired \( M \) and \( \sigma_{\text{turb}} \). We use purely solenoidal driving, which remains active throughout the simulation after gravity is switched on. We take the box side length \( L_{\text{box}} \) to give a box of equal volume to the associated Sphere cloud model. An important difference between the Sphere and Box runs is that in the case of driven boxes the magnetic field is enhanced by a turbulent dynamo (Federrath et al. 2014b) and saturates at about \( \alpha_B \sim 0.1 \) (i.e., 10% relative magnetic energy to gravitational, see Guszejnov et al. 2020), so for Box runs the “pre-stirring” magnetic field strength (defined by \( \mu \)) does not directly specify the actual initial magnetic field strength when gravity is turned on (however the “pre-stirring” flux in the box will still affect the large-scale geometry of the magnetic field).

As shown in Paper II, protostellar jets represent a crucial feedback mechanism as they dramatically reduce stellar masses that is achieved not just by launching some of the accreted material, but also by perturbing the accretion flow around the star. We model their effects by having sink particles launch a fixed fraction of the accreted material along their rotational axis with the Keplerian velocity at the protostellar radius. See Paper I for details on the numerical implementation. In addition to their radiative feedback, massive main-sequence stars inject a significant amount of mass, energy, and momentum into their surroundings through stellar winds. We calculate the mass-loss rates based on a prescription given in Grudić et al. (2021a), motivated by Smith (2014), and wind velocities per Lamers et al. (1995). Winds are implemented either through local mass, momentum and energy injection or direct gas cell scattering, depending on whether the free-expansion radius can be resolved. To account for Wolf-Rayet (WR) stars that dominate the wind energy and momentum budget we use a simple prescription where the mass loss rate of \( M > 20 \, M_\odot \) stars is increased at the end of their lifetime using the WR lifetime prescription of Meynet & Maeder (2005).

Finally, massive stars end their life as a supernova (SN). In the simulation all \( > 8 \, M_\odot \) stars are eligible to become a supernova at the end of their lifetime, for which the minimum is set as 3 Myr. SNe lead to an isotropic ejection of all mass with a total energy of \( E_{\text{SN}} = 10^{51} \) erg, which is implemented through direct gas cell scattering.

The simulations in this paper include all of the physical processes detailed above.

### 2.1.2 Initial Conditions & Parameters of Clouds

We generate our initial conditions (ICs) using MakeCloud (Grudić & Guszejnov 2021), identically to Paper II. Unless otherwise specified our runs utilize “Sphere” ICs, meaning that we initialize a spherical cloud (radius \( R_{\text{cloud}} \) and mass \( M_0 \)) with uniform density, surrounded by diffuse gas with a density contrast of 1000. The cloud is placed at the center of a periodic 10\( R_{\text{cloud}} \) box. The initial velocity field is a Gaussian random field with power spectrum \( E_k \propto k^{-2} \) (Ostriker et al. 2001), scaled to the value prescribed by the \( \sigma_{\text{turb}} = 5 \sigma R_{\text{cloud}}/(3G M_0) \) turbulent virial parameter where \( \sigma \) is the 3D gas velocity dispersion. The initial clouds have a uniform \( B_z \) magnetic field whose strength is set by the \( \mu \) normalized mass-to-flux ratio (Mouschovias & Spitzer 1976). There is no external driving in these simulations. Note that the initial temperature is effectively set by the gas-dust mixture quickly reaching equilibrium with the interstellar radiation field (ISRF), for which we assume solar neighborhood conditions (Draine 2011).

We also run simulations using “Box” ICs, similar to the driven boxes used in e.g., Li et al. (2004); Federrath et al. (2014a); Cunningham et al. (2018). These are initialized as a constant density, zero velocity periodic cubic box with the same temperature prescription as “Sphere” ICs. This periodic box is then “stirred” using the driving algorithm from by Federrath et al. 2010; Bauer & Springel 2012. This involves a spectrum of \( E_k \propto k^{-2} \) for driving modes in Fourier space at wavelengths 1/2 - 1 the box size, with an appropriate decay time for driving mode correlations (\( t_{\text{decay}} \sim t_{\text{cross}} \)). This stirring is initially performed without gravity for five global freefall times \( \left( t_{\text{ff}} = \sqrt{\frac{3 \pi M}{4 \pi G \Sigma}} \right) \), to achieve saturated MHD turbulence. The normalization of the driving spectrum is set so that in equilibrium the gas in the box has a turbulent velocity dispersion that gives the desired \( M \) and \( \sigma_{\text{turb}} \). We use purely solenoidal driving, which remains active throughout the simulation after gravity is switched on. We take the box side length \( L_{\text{box}} \) to give a box of equal volume to the associated Sphere cloud model. An important difference between the Sphere and Box runs is that in the case of driven boxes the magnetic field is enhanced by a turbulent dynamo (Federrath et al. 2014b) and saturates at about \( \alpha_B \sim 0.1 \) (i.e., 10% relative magnetic energy to gravitational, see Guszejnov et al. 2020), so for Box runs the “pre-stirring” magnetic field strength (defined by \( \mu \)) does not directly specify the actual initial magnetic field strength when gravity is turned on (however the “pre-stirring” flux in the box will still affect the large-scale geometry of the magnetic field).

### 2.2 Cluster identification

Despite almost a century of study there is no one accepted definition of what a star cluster is, as the “classical” picture of an isolated, bound, centralized group of stars is not applicable to most observations (Krumholz et al. 2019). Previous work in the literature defined star clusters using an absolute density threshold (Lada & Lada 2003), relative density contrast (McKee et al. 2015), boundedness (Portegies Zwart et al. 2010), Bayesian decomposition into ellipsoids (Kuhn et al. 2014) and numerous other techniques (see Schmeja 2011 for examples). Due to the lack of consensus in the literature, we choose a cluster definition that is both simple and
Table 1. Simulations used in this paper described with STARFORGE label conventions. Top: Physics modules included, see §2.1.1 and Paper I for details on the individual physics modules. Bottom: Initial conditions of clouds used in our runs, with $M_0$, $R_{\text{cloud}}$, $\sigma$ and $\mu$ being the initial cloud mass, size, virial parameter, mass to magnetic flux ratio respectively (note that in runs that explicitly evolve RHD the initial gas-dust temperature is set by the ISRF). We also report the initial 3D turbulent velocity dispersion $\sigma$, thermal virial parameter $\alpha$, Alfvén Mach number $M_A$, plasma $\beta$, magnetic virial parameter $\alpha_m$ as well as the relative Jeans, sonic and magnetic mass scales (see §2 in Guszejnov et al. 2020 for definitions) Note that the parameters in this table apply to both Box and Sphere runs as they are set up to have identical initial global parameters, with $L_{\text{box}}$ being the box size for Box runs and $R_{\text{cloud}}$ being the cloud radius for Sphere runs. Note that Box runs have slightly different initial parameters (e.g., Mach number, virial parameter) due to the non-exact scaling of the driving, so the values shown here are the target values.

| Cloud label | $M_0$ [$M_\odot$] | $R_{\text{cloud}}$ [pc] | $L_{\text{box}}$ [pc] | $\sigma$ [km/s] | $\alpha$ | $\alpha_m$ | $\beta$ | $M_A$ | $\lambda$ | $\Delta\lambda$ [AU] |
|-------------|-------------------|-------------------------|----------------------|-----------------|--------|-----------|--------|------|--------|-----------------|
| M2e4        | 2 $\times$ 10$^7$ | 10                      | 16                   | 2               | 4.2    | 3.2       | 0.008  | 2.03 | 10     | 7 $\times$ 10$^{-7}$ | 0.1              | 2 $\times$ 10$^7$ | 30 |
| M2e4_R3     | 2 $\times$ 10$^7$ | 3                       | 4.2                  | 5.8             | 0.008  | 2.02      | 10     | 0.23 | 3 $\times$ 10$^{-7}$ | 4 $\times$ 10$^{-7}$ | 0.1              | 2 $\times$ 10$^7$ | 30 |
| M2e4_a1     | 2 $\times$ 10$^7$ | 10                      | 4.2                  | 2.3             | 0.008  | 1.03      | 10     | 0.78 | 3 $\times$ 10$^{-7}$ | 4 $\times$ 10$^{-7}$ | 0.1              | 2 $\times$ 10$^7$ | 30 |
| M2e4_a4     | 2 $\times$ 10$^7$ | 10                      | 4.2                  | 4.5             | 0.008  | 4.03      | 10     | 0.78 | 3 $\times$ 10$^{-7}$ | 1 $\times$ 10$^{-7}$ | 0.1              | 2 $\times$ 10$^7$ | 30 |

Figure 1. Surface density maps for M2e4_C_M_I_RT_W (our fiducial run), which is an $M_0 = 2 \times 10^7 M_\odot$ mass cloud resolved with $M_0/\Delta m = 2 \times 10^7$, initial gas cells (see Table 1), at different times, from the beginning of the simulation until cloud disruption. The color scale is logarithmic and the circles represent sink particles (stars) that form in high-density regions where fragmentation can no longer be resolved. The size of the circles is increasing with mass as well as their color changing from red ($M > 0.1 M_\odot$), through green ($M \sim 1 M_\odot$) to blue ($M < 0.1 M_\odot$). This simulation resolves a dynamic range from ~20 pc down to ~30 AU and is run until stellar feedback quenches star formation and disrupts the cloud.

(i) Any star above the H burning limit (> 0.08 $M_\odot$) whose $N_{\text{min}}$ closest neighbors are within $\lambda$ distance is considered a “core particle”.

(ii) All connected core particles and any particles within $\lambda$ distance are considered to be part of the same cluster. Particles not assigned to clusters are considered to be “noise.”

We apply DBSCAN to the 3D spatial positions of the stars, and we adopt $N_{\text{min}} = 10$ and $\lambda = 1$ pc, which effectively serves as our cluster definition. Note that we also experiment with other, more advanced clustering methods that do not require a specified spatial scale, e.g., HDBSCAN (McInnes et al. 2017). Algorithms like HDBSCAN identify the clustering scales from the data, thus provid-
ing results that are not biased by the somewhat arbitrary choice of clustering scale in DBSCAN. While HDBSCAN has been successfully applied to observed young clusters (Kerr et al. 2021), we find that it can create confusing cluster assignments if applied to time series data (i.e., splitting up and merging clusters between different snapshots of a simulation). We also experiment with applying the clustering algorithm to the full 6D phase space data instead of only the 3D spatial positions, similar to the procedure applied to observational data. Doing so, however, requires a mapping from velocity to spatial scales (i.e., a phase-space metric, see Behroozi et al. 2013 for an example). After experimenting with several different methods (e.g., assume a linewidth-size relation, “pre-cluster” in 3D and find velocity dispersion within clusters), we ultimately find no clear advantage to using velocity data, as their main role in observations does not take into account the final label (i.e., assume a linewidth-size relation, “pre-cluster” in 3D and find velocity dispersion within clusters), we ultimately find no clear advantage to using velocity data, as their main role in observations does not take into account the final label (i.e., assume a linewidth-size relation, “pre-cluster” in 3D and find velocity dispersion within clusters), we ultimately find no clear advantage to using velocity data, as their main role in observations

2.3 Cluster properties and definitions

To describe the star clusters in our simulations we introduce several physical quantities. We define the cluster radius (also known as “mean-square radius” or “Spitzer radius”, see Spitzer & Harm 1958), as

\[ R^2 = \langle ||\Delta x||^2 \rangle, \]  

where \( \langle \ldots \rangle \) denotes averaging over cluster members and \( \Delta x \) is the distance of a member star from the center of mass of the cluster. We also define the half-mass radius, \( R_{50} \), as the radius around the center of mass that encloses half the cluster mass. We define the 3D cluster velocity dispersion as

\[ \sigma_{3D}^2 = \langle ||\Delta v||^2 \rangle, \]  

where \( \Delta v \) is the relative velocity of a member star to the center of mass of the cluster.

To characterize the cluster boundedness we use the virial parameter

\[ \alpha = \frac{2E_{\text{kin}}}{-E_{\text{grav}}}, \]  

where \( E_{\text{kin}} \) and \( E_{\text{grav}} \) are the total kinetic and gravitational binding energy of the stars within the cluster. Note that hard binaries are common, but their binding energy has little effect on the overall boundedness of the cluster\(^3\). Thus, it is instructive to define the system virial parameter where we merge binary and multiple systems (identified using the same algorithm as Bate 2009 and Guszejnov et al. 2017)

\[ \alpha_{\text{sys}} = \frac{2E_{\text{kin,sys}}}{-E_{\text{grav,sys}}}. \]  

Here, \( E_{\text{kin,sys}} \) and \( E_{\text{grav,sys}} \) are the total kinetic and gravitational binding energy of the cluster after we replaced binary/triple/quadruple systems with their centers of mass. We can similarly define a 3D system velocity dispersion within the cluster

\[ \sigma_{3D,\text{sys}}^2 = \langle ||\Delta v_{\text{sys}}||^2 \rangle_{\text{sys}}, \]  

where \( \langle \ldots \rangle_{\text{sys}} \) is averaging over systems within the cluster. Note that close binaries are often unresolved in observed clusters, making the velocity dispersion inferred by observations closer to \( \sigma_{3D,\text{sys}}^2 \) than \( \sigma_{3D}^2 \).

Note that these definitions take neither gas cells nor sink particles outside the cluster into account. Considering that clusters inevitably form in areas with dense gas, the contribution of gas to the initial boundedness is significant. As a crude estimate we calculate the amount of gas within the spatial extent of the cluster (\( R \) from Eq. 1) and calculate its contribution to the gravitational energy of the cluster members (\( E_{\text{grav,gas}} \)) by assuming that this mass

\(^3\) Observational estimates of the virial parameter likely also suffer from biases introduced by binaries, see Gieles et al. (2010).
is distributed homogeneously within the cluster. This leads to the \( \alpha' \) and \( \alpha'_{\text{sys}} \) virial parameters:

\[
\alpha' = \frac{2E_{\text{kin}}}{-E_{\text{grav}} - E_{\text{grav, gas}}},
\]

\[
\alpha'_{\text{sys}} = \frac{2E_{\text{kin, sys}}}{-E_{\text{grav, sys}} - E_{\text{grav, gas, sys}}}.
\]

Note that by definition \( \alpha' \leq \alpha \), and it only becomes equal at later times when most of the gas has been expelled from the cluster. These estimates are within a factor of few of the values returned by directly calculating the contributions from gas within the cluster.

In our simulations we find that clusters tend to expand after gas expulsion. The clustering algorithm (§2.2) often breaks these expanding clusters into separate smaller clusters. In order to quantify the cluster expansion we introduce the mass-weighted mean radial velocity

\[
\bar{v}_{\text{rad}} = \frac{\sum m_i v_{\text{rad}, i}}{\sum m_i},
\]

where \( v_{\text{rad}} \) is the radial velocity of a star relative to the cluster center of mass and the summation is over all cluster members.

### 2.3.1 Mass segregation

Observed clusters exhibit mass segregation, i.e., massive stars are “distributed differently” than lower mass stars (Krumholz et al. 2019). This often means that they are concentrated at the minimum of the gravitational potential, i.e., the dense center of the cluster (Hillenbrand & Hartmann 1998). Many studies adopt this more specific criterion to define mass segregation. There are several methods in the literature to characterize this phenomenon relying on the cluster density profiles (e.g., Hillenbrand 1997) or characteristic radial distance (Gouliermis et al. 2009) of stars in various mass bins. Alternatively, one can also calculate the slope of the mass function of stars at different radii from the cluster center (de Grijs et al. 2002).

These methods, however, are sensitive to the choice of mass bins and annuli (Gouliermis et al. 2004) and to the precise determination of the cluster center. An alternative metric that is insensitive to these is to construct a Minimum Spanning Tree (MST), the shortest graph connecting all stars without closed loops. Comparing the characteristic MST edge length between massive stars and randomly chosen stars can quantify the level of mass segregation in the cluster (see e.g. Cartwright & Whitworth 2004 and Allison et al. 2009).

In this work we consider two separate mass segregation metrics. The first one is based on the definition of Allison et al. (2009), which quantifies the degree of mass segregation using the mass segregation ratio (MSR):

\[
\Lambda_{\text{MSR}} = \frac{\langle d_{\text{norm}} \rangle_{\text{MC}}}{\langle d_{\text{massive}} \rangle_{\text{massive}}},
\]

where \( \langle d_{\text{norm}} \rangle_{\text{MC}} \) is the mean edge length of the MST between massive stars only. We define massive stars for the remainder of this paper as any star above 5 M\(_{\odot}\). Meanwhile, \( \langle d_{\text{massive}} \rangle_{\text{massive}} \) is the mean edge length for \( N_{\text{massive}} \) randomly chosen stars, where \( N_{\text{massive}} \) is the number of massive stars. The \( \langle d_{\text{norm}} \rangle_{\text{MC}} \) operation denotes constructing \( N_{\text{sets}} = 500 \) random sets and averaging over them, so \( \langle d_{\text{norm}} \rangle_{\text{MC}} \) is the mean of the median MST edge lengths from \( N_{\text{sets}} \) of random realizations. Since this metric is only meaningful if at least several massive stars exist, we require \( N_{\text{massive}} \geq 5 \) for it to be defined.

A significant drawback of the MST based method is that it requires at least several massive stars to be already present in the cluster, while both observations (Kirk & Myers 2011) and simulations (Kirk et al. 2014) find that even small groupings of stars with a single massive star exhibit signs of mass segregation. We find that our clusters, like observed young clusters (Kirk & Myers 2011; Kerr et al. 2021), are initially highly structured, where the stellar distribution follows the hierarchical distribution of the star-forming gas. In this case, the MST method does not detect mass segregation as it only exists within smaller sub-groups of stars, i.e., if the cluster consists of several mass-segregated subclusters.

To address this issue we identify coherent groups of stars, sub-clusters, within each cluster. We define these subclusters as centrally condensed stellar over-densities and divide every cluster into one or more subclusters. We identify subclusters in each of our clusters by applying the Variational Bayesian Gaussian Mixture cluster identifying algorithm (Attias 2000; Bishop 2006) from the scikit-learn library. This Gaussian Mixture method decomposes the cluster into several Gaussian density distributions, which are, by our definition, subclusters (see Figure 2 for a cartoon illustration and Figure 3 for later examples from the simulations). Unlike the DBSCAN algorithm we use to identify the clusters themselves, Gaussian Mixture Models in general require no specific length scale and the specific variational method can infer the appropriate number of Gaussian components, i.e. subclusters. Note that we do not use the Variational Bayesian Gaussian Mixture model to identify the clusters themselves in the simulations, because this method suffers from the same assignment persistence issues as other clustering algorithms without spatial scales (see §2.2).

In order to account for cluster substructure, we introduce a second metric for mass segregation, the mass segregation offset (MSO):

\[
\Lambda_{\text{MSO}} = \left( \frac{d_{\text{subcl}}}{R_{\text{subcl}}} \right)_{\text{all}} \left( \frac{d_{\text{subcl}}}{R_{\text{subcl}}} \right)_{\text{massive}},
\]

where \( d_{\text{subcl}} \) is the distance from a star to the center of the nearest subcluster with \( R_{\text{subcl}} \) being its size (defined following Eq. 1), while the \( \langle \cdot \rangle_{\text{massive}} \) operation denotes averaging over all massive stars in the cluster (see Figure 2). Note that in this definition we introduce the concept of “subcluster”, which makes the definition of Equation 10 in theory different from similar offset measures in the literature (Kirk et al. 2014), although it gives the same answer for small or highly centralized clusters.

Finally we define the mass segregation time scale from for a star of mass \( M \) as (Spitzer 1969; Binney & Tremaine 1987):

\[
t_\text{seg}(M) = \frac{(m)}{M} \frac{N}{8 \ln N} \frac{R}{\sigma_{\text{sys}}},
\]

where \( (m) \) is the average stellar mass in the cluster, while \( N \) is the number of its members, \( R \) is the cluster size (Eq. 1) and \( \sigma_{\text{sys}} \) is its velocity dispersion. Note that using the system velocity dispersion changes the results by a factor of 2.

### 3 CLUSTER FORMATION AND EVOLUTION IN THE FIDUCIAL M2E4 RUN

In this section we detail the formation and evolution of clusters within our fiducial run (M2e4, Sphere) that are identified using the methodology described in §2.2.

#### 3.1 General behavior

We find that star formation begins at several locations in the cloud, which we refer to as star formation sites. These sites produce a few
1–2 massive stars, the mean edge length of the MST (Eq. 9) would yield
respectively, showing also the centers of both the cluster and the subclusters
denoted with cross markers. Due to the subclusters having only
1–2 massive stars, the mean edge length of the MST (Eq. 9) would yield
Msrs ∼ Rs cluster/RC cluster = 1, ignoring mass segregation on the subcluster
level. Meanwhile, the mass segregation offset (see Eq. 10) would be
\Lambda MSO ∼ Rs subc/Δsubc ≫ 1.

massive stars (often just a single one) as well as many lower mass
stars, forming a small cluster with an order of 20 stars (see Figure 3).
These small clusters are still gas-dominated, actively accreting and
star-forming when they encounter nearby clusters and merge with
them, forming larger clusters. This behavior is similar to previous
results claiming hierarchical cluster assembly from similar initial
conditions (e.g., Bonnell et al. 2003; Grudić et al. 2018). The newly
formed clusters continue accreting gas and forming new stars, as
well as merging with other structures until feedback from massive
stars terminates star formation and expels the remaining gas (see
Figure 1).

Figure 4 shows the formation and merger history of newly
formed clusters depicting the hierarchical build-up of larger struc-
tures via mergers of smaller clusters. This leads to the formation of
a “dominant” cluster that ultimately encompasses most of the stellar
mass in the simulation. Once stellar feedback expels the gas from
a cluster, the remaining stars are not gravitationally bound and the
cluster starts breaking into smaller structures. This mainly affects
the largest cluster, which becomes unbound and expands, breaking
up into many smaller clusters.

3.2 Cluster properties
To illustrate the evolution of cluster properties in our simulations
we focus on the “dominant” cluster that eventually encompasses
the majority of stars at the end of the simulation. Figure 5 shows
the cluster properties defined in §2.3. The dominant cluster reaches
about 1000 members and attains roughly 1000 M⊙ by the time the
cloud disrupts (see panels a-b), containing the majority of the total
stellar mass. This run ends with SFE = M* / M0 ∼ 7%, cor-
responding to M* ∼ 1400 M⊙). Clusters form around individual star
formation sites, and these structures merge to form larger objects,
leading to “jumps” in the cluster size. Although gravitational at-
traction between the various substructures and stellar interactions
should shrink the cluster over time and increase the central stellar
density (Krause et al. 2020), the continuous formation of new stars
from infalling gas and mergers with other clusters causes the cluster
to maintain its size until gas expulsion occurs (Figure 5 panel c).

As the star cluster grows rapidly in both mass and size the
velocity dispersion also increases (Figure 5, panel d). Note that the
stellar velocity dispersion, \sigma_{3D,sys} (Eq. 5) is super-virial due to the effect
of close binaries. Meanwhile, the velocity dispersion for systems,
\sigma_{3D,sys} (Eq. 5) is fairly close to the virial value (if the gas potential
is also taken into account). The stellar velocity dispersion peaks as
the cluster reaches its maximum mass, just as gas expulsion starts. It
decreases subsequently as the cluster breaks apart. Thus, the statistics
follow a shrinking fraction of the original cluster.

We expect that most close binaries are unresolved in observed
clusters (e.g., Foster et al. 2015; Kerr et al. 2021), so the observ-
ationally inferred velocity dispersion should be close to \sigma_{3D,sys}.
Thus, simple estimates using the global cluster mass and size scales
(after including the enclosed gas mass) would conclude these clus-
ters are virialized during star formation (as in Foster et al. 2015)
and highly supervirial during the breakup phase after star formation
ceases and/or a significant fraction of the gas mass has been ex-
pelled. Meanwhile, direct calculation of the virial parameter using
Equation 3 indicates the cluster is highly sub-virial with \alpha ∼ 1/2 − 1
(Figure 5, panel e). We find that the gravitational potential energy
is dominated by hard binaries, leading to low \alpha values. After merg-
ing these systems, i.e., using the definition from Eq. 4, \alpha_{sys}
is consistently above the boundedness limit of \alpha = 2. However, ini-
tially the clusters are strongly gas-dominated, so after correcting for
the gas potential (Eq. 7) we find the clusters are initially strongly
bound (\alpha’ < 1, similar to the results of Offner et al. 2009b) and
then become unbound after gas expulsion (\alpha’ ∼ 10). The result-
ing unbound cluster immediately expands and breaks into smaller
structures. Since the simulation stops shortly after gas expulsion, it
is unclear what fraction of the original cluster will remain bound.
This will be investigated in a future STARFORGE project.

3.3 Evolution of Mass Segregation
Figure 5 shows that the dominant cluster develops mass segrega-
tion (panel f) according to both the \Lambda MSR and \Lambda MSO metrics (see
Equations 9-10). However, these measures differ on the initial degree
of mass-segregation, with \Lambda MSO identifying segregation from the
time the first massive stars form, while \Lambda MSR only detecting it at
much later times. In the early stage of cluster evolution, each cluster
is composed of the stars formed in one to a few star formation site
and hosts only a few massive stars in the center (see Figure 3). These
sites continue to accrete, form more stars, and merge with others,
thereby forming ever-larger clusters. The resulting clusters inherit
the centrally condensed substructures, which then interact and sink
toward the center, increasing \Lambda MSR. Eventually the cluster relaxes
to a centrally condensed, “classical” star cluster. This redistribution
occurs on a timescale of tseg ∼ 2 Myr for massive stars (see Eq.
11). Before the cluster can fully dynamically relax, stellar feedback
expels the remaining gas and unbinds the cluster. As gas expulsion
begins, massive stars are preferentially located near the center of the
dominant cluster, leading to high \Lambda MSR and \Lambda MSO values. During
the gas dispersal process massive stars that formerly reside near the
center move outwards with the rest of the cluster, causing both mass
segregation metrics to drop. Note that at later times \Lambda MSO may
increase, but it is due to the cluster identification algorithm splitting
the dominant cluster into smaller clusters, which tend to have a few
massive stars at their centers.
3.4 Mass distribution of stars inside and outside clusters

We find that the majority of stars that form in our simulations end up in clusters, although a significant fraction (~ 10%) are ejected before gas expulsion (see Table 2 and Figure 6). During its lifetime the dominant cluster in the fiducial simulation gains stellar mass from two sources: 1) mergers with other clusters and 2) gas accretion by its stellar members or newly formed stars within the cluster. We find these two mechanisms to have roughly similar weight during most of the cluster lifetime, with accretion becoming more important after most stars have merged with the dominant cluster, leaving no other clusters to merge with. After gas expulsion the cluster becomes unbound and loses an order unity of its mass.

Unsurprisingly the stellar mass distributions of the dominant cluster and the full simulation are also similar.

4 EFFECTS OF INITIAL CONDITION VARIATIONS ON CLUSTERS

In this section we investigate the effects of turbulent driving (§4.1), the initial level of turbulence (§4.2) and surface density (§4.3) on the properties of the dominant cluster. These properties have a significant effect on the star formation history of the system, and here we examine the impact of these properties on clustering.

4.1 Cloud setup and turbulent driving (Box vs. Sphere)

As noted in §2.1.2, the Sphere vs Box configurations have two important differences, which may lead to different clustering properties. First, the periodic boundary conditions of the Box setup leads to both an order-of-magnitude shallower gravitational potential (Federrath & Klessen 2012) and prevents the escape of radiation and gas. Second, the Box setup starts from a self-consistent, prestirred state and this external driving is continuous throughout the run, providing energy for turbulent modes on the box scale that cascade down to smaller scales. To disentangle the effects of these two factors, we compare three M2e4 runs (Table 1): 1) our fiducial Sphere run, 2) a Box run with continuous external driving and 3)
Table 2. General cluster and sink properties of various runs (see Table 1), including the final star formation efficiency ($\text{SFE} = \frac{M_{\ast,\text{final}}}{M_{\text{cloud}}}$), the relative mass of stars in clusters vs outside clusters (calculated as $\max(M_{\text{in clusters}})/\max(M_{\ast,\text{final}})$) and the relative weight of the dominant cluster ($\max(M_{\text{dominant cluster}})/\max(M_{\text{in clusters}})$).

| Label           | SFE ($M_{\ast,\text{final}}/M_{\text{cloud}}$) | $\max(M_{\text{in clusters}})/\max(M_{\ast,\text{final}})$ | $\max(M_{\text{dominant cluster}})/\max(M_{\text{in clusters}})$ |
|-----------------|-----------------------------------------------|---------------------------------------------------------------|---------------------------------------------------------------|
| $\text{M2e4}$   | 8%                                           | 83%                                                          | 94%                                                          |
| $\text{M2e4}$ Box, with driving | 5%                                           | 45%                                                          | 64%                                                          |
| $\text{M2e4}$ Box, no driving  | 9%                                           | 73%                                                          | 46%                                                          |
| $\text{M2e4_a1}$ | 11%                                          | 80%                                                          | 94%                                                          |
| $\text{M2e4_a4}$ | 4%                                           | 67%                                                          | 79%                                                          |
| $\text{M2e4_R3}$| 14%                                          | 87%                                                          | 99%                                                          |

Figure 4. Merger history of clusters in the fiducial M2e4 run. Each line represents a cluster (using its assigned color) with a width that logarithmically increases with mass. Mergers and splits are denoted by connecting the lines at the time of the event. The line representing the cluster ends if the cluster dissolves. The initial behavior is hierarchical assembly where clusters merge to form ever greater structures. This continues until feedback expels gas from the cluster, so it becomes unbound and begins to break into smaller clusters.

Figure 8 shows that while the evolution of the Sphere run is well described as the hierarchical assembly of one dominant cluster, this is not the case in the driven Box run. Turning off turbulent driving restores this behavior, as the gas undergoes global gravitational collapse once the initial turbulence decays (see Figure 9). Note that even without turbulent driving the shallower gravitational potential of the Box run relative to the Sphere run leads to weaker gravitational focusing, delaying mergers (see Figure 10). We also find that continued turbulent driving leads to the formation of a significant number of transient clusters that survive for a few 100 kyr before dissolving.

Turbulent driving dramatically slows down star formation in the cloud ($\text{SFE} \propto r^2$ vs $\text{SFE} \propto r$), while in the non-driven case star formation is only suppressed until the initial turbulent velocity field decays. This is due to the weaker gravitational potential in the Box runs (Federrath & Klessen 2012), which produces weaker gravitational focusing in addition to the external driving that prevents global gravitational collapse. This is apparent in Figure 11 as the dominant cluster in both Box runs grow significantly slower than in the Sphere one. While the initial cluster masses are similar between the driven and decaying runs, the decaying run has (on average) more massive members due to the slightly less top heavy IMF in the Box runs. Unlike the Sphere run, the Box runs experience no cloud disruption, and stellar feedback is unable to permanently expel gas from the cluster before the gas in the simulation volume is heated to unphysical temperatures by radiation trapped by the periodic boundary condition. There is no permanent gas expulsion, the clusters themselves do not suddenly become unbound (like in the Sphere run). Their future evolution, however, is uncertain as we stop the simulation when it reaches the unphysical, radiation-filled regime.

4.2 Initial level of turbulence

In this section we vary the initial velocity dispersion to determine the impact of the cloud turbulence on clustering. We compare Sphere runs with $\alpha_{\text{turb}}$ values of 1.2 and 4, which correspond to bound, marginally bound and unbound clouds, respectively. We find that the final star formation efficiency decreases with increasing $\alpha_{\text{turb}}$ (see Table 2). Figure 12 shows that higher $\alpha_{\text{turb}}$ also leads to less concentrated star formation and weaker global gravitational collapse due to increased turbulent support. However, the hierarchical cluster formation picture that we find for the fiducial run ($\alpha_{\text{turb}} = 2$) still qualitatively applies (see Figures 8 and 12).

All runs produce a single dominant cluster. The final mass of this dominant cluster decreases with increased turbulence, mostly due to the lower final SFE values of the clouds (Table 2). Figure 13 shows that the dominant cluster follows a qualitatively similar evolutionary trend in all runs, with higher masses and consequently higher velocity dispersions for runs with higher SFE (i.e., lower $\alpha_{\text{turb}}$). The dominant cluster becomes unbound once stellar feedback expels the remaining gas, leading to its expansion and breakup into smaller clusters. In Figure 13 it appears as if the $\alpha_{\text{turb}} = 4$ run had
Figure 5. Evolution of the dominant cluster in the fiducial run (M2e4, Sphere), highlighting major events, i.e., mergers with another cluster (diamonds) and the splitting of the cluster (crosses), colored according to Figure 4. Dotted horizontal lines mark the characteristic scale of our cluster size definition (1 pc), the virial equilibrium and marginal boundedness ($\alpha = 1$ and 2) and the boundary between regular and inverse mass segregation. The top row shows the total stellar mass in the cluster (top left), number of cluster members (top middle) and the cluster size scales (top right, Eq. 1) respectively. The bottom row shows the cluster velocity dispersion (left, Eq. 2), virialization state (middle, Eqs. 3-7) and mass segregation (right, Eqs. 9-10). Note that all values are smoothed with a 30 kyr averaging window to make the plots easier to read. For an analysis of the main trends see §3.2.

a much longer cluster lifetime. In fact, feedback causes the cluster to expel its gas content and become unbound in roughly the same time (4 Myr) as in the other cases (see gas mass fraction and $\alpha_{turb}$ panels of Figure 13). At the same time the dominant cluster merges with two neighboring clusters, increasing the effective size of the resulting cluster and the relative gas mass content.

We find that all three runs are mass-segregated from early times, and we see no clear trend in either $\Lambda_{MSR}$ or $\Lambda_{MSO}$ as a function of $\alpha_{turb}$. Therefore, we conclude that mass segregation is not very sensitive to modest changes in the initial cloud virial parameter.

4.3 Surface density

Cloud surface density is thought to be a key parameter of star formation (Krumholz & McKee 2008; Fall et al. 2010b; Grudić et al. 2021b) due to its influence on the dynamics of fragmentation and degree of stellar feedback. Although we present only one run with a different surface density (Sphere run with a factor 10 times increase in $\Sigma$: M2e4_R3), we also ran a calculation with 10 times lower surface density, but it had a final SFE value of only 1% and produced no clusters with more than 20 stars, preventing a meaningful cluster analysis.

As expected, increasing the surface density leads to enhanced star formation and a higher final SFE (Table 2). Higher surface density also means that the cloud is smaller, making it easier for the clustering algorithm (see §2.2) to join star formation sites. Consequently, nearly all stars end up in one massive cluster (Figure 14). Similar to the fiducial run, the dominant cluster is gas-dominated and becomes unbound once stellar feedback expels the gas. The characteristic timescale of cloud evolution is the freefall time, which, due to the higher overall density, is significantly shorter than that of our fiducial run (Figure 15). Note that the cluster assembly phase is mainly determined by this timescale, while the length of the following gas expulsion phase depends on both the freefall time and the timescales for stellar evolution. Apart from this non-trivial rescaling, the time evolution of the dominant cluster is similar in the fiducial and the high surface density runs.

5 DISCUSSION

5.1 Cluster assembly and dynamical effects

Clusters in our simulations form through the mergers of accreting subclusters, which in turn merge to form ever larger structures, corresponding to a hierarchical assembly of clusters. By the time stellar feedback becomes important in the simulation most stars are concentrated in one or a few clusters. Stellar feedback eventually expels the remaining gas and the clusters become unbound, leading to their expansion and breakup (Tutukov 1978; Hills 1980; Mathieu 1983).

This qualitative picture appears to be robust to changes in ini-
Is there primordial mass segregation?

There is no primordial mass segregation. This conclusion is based on the findings of Weisz et al. (2015) and others.

5.2 Mass segregation

One key question of cluster formation is whether clusters form mass segregated or become so through dynamical interactions. In our simulations star formation sites often host a single or several massive stars ($M_\star > 5M_\odot$) at their center, making them mass segregated by all definitions. These small clusters merge to form larger structures, which in turn inherit the centrally condensed, mass-segregated substructures. Due to dynamical interactions these substructures strip stars from each other while merging. Over time dynamical processes cause the dense centers, which also host the massive stars, to sink to the center of the larger merged structure. While these processes are taking place the substructures continue to grow, forming new stars as well as continuing to accrete gas. Thus we find that whether a cluster is considered mass-segregated depends greatly on the definitions of “cluster” and “mass segregation,” neither of which have one accepted definition in the literature (see Krumholz et al. 2019 for discussion). If one defines mass segregation as any stellar configuration where “massive stars are distributed differently” than lower mass stars (as in Krumholz et al. 2019), then star clusters start out mass segregated regardless of how cluster membership is assigned, since clusters contain substructures that host a single or a few massive stars at their respective centers. However, if mass-segregation on the cluster scale occurs when “massive stars are preferentially at the center” of the cluster, as in de Grijs et al. (2002) and Krause et al. (2020), then the cluster definition determines whether the stars are mass segregated. Choosing a method that picks out structures containing several star formation sites (like the one we used) leads to no initial mass segregation (see the scenario in Figure 2 and the evolution in Figure 5). If, however, a cluster definition picks out individual star formation sites (e.g., by defining a smaller characteristic length or by requiring that a cluster be centrally condensed), then mass segregation will appear primordial regardless of metric (see Table 3 for a summary).

Most observers define clusters as pc-sized objects with many massive stars (Kirk & Myers 2011) and use mass segregation metrics that are insensitive to the mass segregation of any substructures within the cluster (e.g., de Grijs et al. 2002). In this framework, clusters in the simulation are initially not mass segregated and become so through dynamical interactions. The process takes several Myr, which is enough time for feedback from massive stars to expel gas but not enough time for the cluster to reach a fully relaxed state.

### 5.3 Caveats

While the simulations presented here represent the current state-of-the-art for simulating star-forming clouds, STARFORGE employs a large number of approximations and assumptions to make the simulations computationally tractable like other simulations in the literature (see Paper I for detailed discussions). In particular, there are significant caveats when applying STARFORGE to model star cluster formation. First, the runs have a ~ 30 AU Jeans-resolution,
Figure 7. **Left:** Mass distribution of sink particles (stars), comparing three populations: those assigned to clusters (blue), those that are not assigned to clusters (orange) and the full stellar population (black). The spectrum is taken at the time when the dominant cluster reaches its maximum mass (i.e., just as gas expulsion starts). Shaded regions denote the 2 sigma Poisson uncertainties. We also show the Kroupa (2002) canonical fitting function for the MW IMF from the literature and mark the completeness limit of the simulation. **Right:** Similar to the left panel, but concentrating on the dominant cluster only. We compare three populations: those assigned to the dominant cluster (blue), those that were ejected from it (orange) and the full stellar population (black).

Figure 8. The surface density maps and cluster assignments of the M2e4 runs with both “Sphere” and “Box” ICs. The surface density maps use the same conventions as Figure 1. On the cluster assignment maps each star is represented with its cluster ID and colored with the color assigned to the cluster, while stars not assigned to clusters are marked with black circles.
Box run with driven turbulence  

Box run with decaying turbulence

Figure 9. Same as Figure 8 for M2e4 Box runs with and without turbulent driving.

Figure 10. Merger histories for the M2e4 Box runs with and without turbulent driving and the fiducial Sphere run, same as Figure 4. Driving prevents the hierarchical merging of smaller clusters and leads to the formation of many transient clusters. The periodic boundary conditions of the Box run also reduce gravitational focusing relative to the Sphere run, suppressing mergers.

6 CONCLUSIONS

In this work we analyze the star cluster assembly process in the STARFORGE radiation-magnetohydrodynamics simulations. These simulations follow the evolution of a mid-sized molecular cloud \((M = 20000 \, M_\odot, \Sigma \sim 60 \, M_\odot/pc^2)\) taking into account gravity, gas thermodynamics, turbulence, magnetic fields, and radiation as well as stellar feedback processes (jets, radiation, winds, SNe).

Star clusters assemble through a series of mergers, whereby accreting star-formation sites come together to form larger structures. This hierarchical assembly continues until most stars end up in one or a few large, gas-dominated clusters. Once stellar feedback expels the gas they become unbound and the stars disperse. During the assembly process clusters eject a small fraction of their members \((\lesssim 10\%)\). We find no significant difference between the mass distribution of the ejected stellar population and that of the overall stellar mass spectrum of the simulation.

We also investigate the effects of different initial cloud geometries and turbulent driving. We find that turbulent driving and a periodic “Box” geometry significantly reduces clustering and suppresses cluster mergers. This is caused by weaker gravitational focusing, as periodic boundaries lead to a shallower gravitational potential, while maintaining turbulence reduces collapse.

We consider two different definitions for mass segregation. In all simulations, small forming groups of stars are initially mass segregated with one or a few massive stars at their center. As these structures merge, they (at first) become mass-segregated substructures within the newly formed cluster. Thus massive stars are not initially in the center of merged clusters. Through dynamical interactions they relax to a centralized configuration, similar to that of accuracy of N-body interactions is significantly lower than in pure N-body simulations. It should be noted, however, that we run our simulations until gas expulsion, corresponding to a relatively short time after star formation starts \((< 5 \, \text{Myr})\). Also, our clusters are gas-dominated for most of their lifetime and do not achieve high stellar densities, making close encounters rarer and lessening the effects of the gravitational softening on stellar interactions. Since we terminate the simulations soon after cloud disruption, we cannot predict the ultimate fate of the clusters, bound mass fraction and cluster mass function. We will investigate the long-term evolution and fate of the STARFORGE clusters in a follow-up paper.
observed clusters. We find that whether clusters are quantitatively considered to be mass segregated depends greatly on how one defines a cluster and mass segregation. If clusters are defined as structures that include many star-formation sites distributed throughout the GMC and mass segregation requires massive stars to be at the center (both of these are true for most definitions used in observations), then there is no primordial mass segregation. Rather, mass segregation results from dynamical interactions. On the other hand, massive stars are usually centrally located within bound sub-groups of stars, such that they are distributed differently with respect to low-mass stars. Thus, a definition of mass segregation that does not require massive stars to be globally centralized, concludes that clusters start out mass segregated (see Table 3 for a summary).

In the simulations, dynamical evolution is still ongoing at the time of gas expulsion. Future work will investigate the evolution of the stellar distribution over 100 Myr timescales and determine the survival rate of the star clusters we identify here.

7 DATA AVAILABILITY

The data supporting the plots within this article are available on reasonable request to the corresponding authors. A public version of the data is expected to be made available through a public repository.
$\alpha_{\text{turb}} = 1$ (M2e4\_a1) \hspace{1cm} \alpha_{\text{turb}} = 4$ (M2e4\_a4)

![Figure 12. Same as Figure 8 but for runs with initial turbulent virial parameters of $\alpha_{\text{turb}} = 1$ and $\alpha_{\text{turb}} = 4$ respectively (M2e4\_a1 and M2e4\_a4).](image)

of the GIZMO code is available at [http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html](http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html).

**ACKNOWLEDGEMENTS**

We would like to thank Sinan Deger for his thoughtful comments.

DG is supported by the Harlan J. Smith McDonald Observatory Postdoctoral Fellowship and the Cottrell Fellowships Award (#27982) from the Research Corporation for Science Advancement. Support for MYG was provided by NASA through the NASA Hubble Fellowship grant #HST-HF2-51479 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. Support for PFH was provided by NSF Collaborative Research Grants 1715847 & 1911233, NSF CAREER grant 1455342, and NASA grants 80NSSC18K0562 & JPL 1589742. SSRO and CM are supported by NSF Career Award AST-1748571 and by a Cottrell Scholar Award from the Research Corporation for Science Advancement. 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. ALR acknowledges support from Harvard University through the ITC Postdoctoral Fellowship. This work used computational resources provided by XSEDE allocation AST-190018, the Frontera allocation AST-20019, and additional resources provided by the University of Texas at Austin and the Texas Advanced Computing Center (TACC; [http://www.tacc.utexas.edu](http://www.tacc.utexas.edu)).

**REFERENCES**

Allison R. J., Goodwin S. P., Parker R. J., Portegies Zwart S. F., de Grijs R., Kouwenhoven M. B. N., 2009, MNRAS, 395, 1449
Attias H., 2000, Advances in neural information processing systems, 12, 209
Bate M. R., 2009, MNRAS, 392, 1363
Bate M. R., Bonnell I. A., Price N. M., 1995, MNRAS, 277, 362
Bauer A., Springel V., 2012, MNRAS, 423, 2558
Baumgardt H., Kroupa P., 2007, MNRAS, 380, 1589
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Binney J., Tremaine S., 1987, Galactic dynamics
Bishop C. M., 2006, Pattern recognition and machine learning. springer
Bonnell I., Bate M., 2006, MNRAS, 370, 488
Bonnell I. A., Bate M. R., Vine S. G., 2003, MNRAS, 343, 413
Bressert E., et al., 2010, MNRAS, 409, L54
Cartwright A., Whitworth A. P., 2004, MNRAS, 348, 589
Cunningham A. J., Krumholz M. R., McKee C. F., Klein R. I., 2018, MNRAS, 476, 771
Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium
Ester M., Kriegel H.-P., Sander J., Xu X., et al., 1996, in Kdd. pp 226–231
Fall S. M., Krumholz M. R., Matzner C. D., 2010a, ApJ, 710, L142
Figure 13. Evolution of the dominant cluster in a set of runs with different levels of initial turbulence (M2e4, a1, M2e4 and M2e4, a4, see Table 1), similar to Figure 11. The top row shows the total stellar mass in the cluster (left), number of cluster members (middle) and the cluster size scales (right, Eq. 1) respectively. The middle row shows the cluster velocity dispersion (left, Eq. 2), virialization state (middle, Eq. 4) and the gas mass fraction within the cluster radius (like Figure 5 panel e). The bottom row shows the $\bar{v}_{\text{rad}}$ mass-weighted mean radial velocity for the cluster (left, Eq. 8), as well as the mass segregation ratio (right, Eq. 9) and mass segregation offset (right, Eq. 10). For an analysis of the main trends see §4.2 in the main text.

Fall S. M., Krumholz M. R., Matzner C. D., 2010b, ApJ, 710, L142
Farias J. P., Fellhauer M., Smith R., Domínguez R., Dabringhausen J., 2018, MNRAS, 476, 5341
Federrath C., Klessen R. S., 2012, ApJ, 761, 156
Federrath C., Roman-Duval J., Klessen R. S., Schmidt W., Mac Low M.-M., 2010, A&A, 512, A81
Federrath C., Schön M., Banerjee R., Klessen R. S., 2014a, ApJ, 790, 128
Federrath C., Schober J., Bovino S., Schleicher D. R. G., 2014b, ApJ, 797, L19
Ferland G. J., et al., 2013, Rev. Mex. Astron. Astrofis., 49, 137
Foster J. B., et al., 2015, ApJ, 799, 136
Geen S., Watson S. K., Rosdahl J., Bieri R., Klessen R. S., Hennebelle P., 2018, MNRAS, 481, 2548
Gieles M., Sana H., Portegies Zwart S. F., 2010, MNRAS, 402, 1750
Gouliermis D. A., 2018, PASP, 130, 072001
Gouliermis D., Keller S. C., Kontizas M., Kontizas E., Bellas-Velidis I., 2004, A&A, 416, 137
Gouliermis D. A., de Grijs R., Xin Y., 2009, ApJ, 692, 1678
Grudić M. Y., Guszejnov D., Hopkins P. F., Lamberts A., Boylan-Kolchin M., Murray N., Schwiz D., 2018, 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., Kruijssen J. M. D., Faucher-Giguère C.-A., Hopkins P. F., Ma X., Quataert E., Boylan-Kolchin M., 2021b, MNRAS, 506, 3239
Grudić M. Y., Guszejnov D., Offner S. S. R., Rosen A. L., Raju A. N., Faucher-Giguère C.-A., Hopkins P. F., 2022, arXiv e-prints, p.
\[ \Sigma = 10 \times \Sigma_{MW} = 630 \text{M}_\odot/\text{pc}^2 \] 

Figure 14. Same as Figure 8 but for a run with 10 times higher initial surface density (M2e4_R3).

This paper has been typeset from a \LaTeX file prepared by the author.
Figure 15. Evolution of the most dominant cluster in a set of runs with different initial surface densities (M2e4 and M2e4_R3, see Table 1), similar to Figure 11. To make the runs easier to compare, we normalized the time evolutions with the freefall times of the respective initial clouds (3.7 and 0.6 Myr for M2e4 and M2e4_R3 respectively). The top row shows the total stellar mass in the cluster (left), number of cluster members (middle) and the cluster size scales (top right, Eq. 1) respectively. The middle row shows the cluster velocity dispersion (left, Eq. 2), virialization state (middle, Eq. 4) and the gas mass fraction within the cluster radius (like Figure 5 panel e). The bottom row shows the $\bar{v}_{\text{rad}}$ mass-weighted mean radial velocity for the cluster (left, Eq. 8), as well as the mass segregation ratio (right, Eq. 9) and mass segregation offset (right, Eq. 10). For an analysis of the main trends see §4.3 in the main text.