



Star cluster formation from turbulent clumps – V. Stellar clustering around massive stars

Downloaded from: <https://research.chalmers.se>, 2026-10-16 18:14 UTC

Citation for the original published paper (version of record):

Gautam, A., Farias, J., Tan, J. (2026). Star cluster formation from turbulent clumps – V. Stellar clustering around massive stars. *Monthly Notices of the Royal Astronomical Society*, 552(1).
<http://dx.doi.org/10.1093/mnras/stag1721>

N.B. When citing this work, cite the original published paper.

Star cluster formation from turbulent clumps – V. Stellar clustering around massive stars

Aayush Gautam^{1,2★}, Juan P. Farias^{3,4} and Jonathan C. Tan^{1,2,5}

¹Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA

²Virginia Institute for Theoretical Astronomy, University of Virginia, Charlottesville, VA 22904, USA

³Department of Astronomy, University of Texas at Austin, Austin, TX 78712, USA

⁴Department of Physics and Astronomy, McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4M1, Canada

⁵Department of Physics & Astronomy, Chalmers University of Technology, Gothenburg SE-41293, Sweden

Accepted 2026 September 7. Received 2026 August 15; in original form 2026 April 29

ABSTRACT

Massive stars ($\geq 8 M_{\odot}$) are known to have high degrees of multiplicity, e.g. with about 60 per cent in triples or higher-order multiples. Such high levels of multiplicity may arise during formation (primary multiplicity) or through dynamical processing of already formed stars in dense clusters (secondary multiplicity). The level of primary multiplicity is an important metric to help distinguish between different formation scenarios, such as core accretion and competitive accretion. The level of secondary multiplicity is expected to evolve with time and be sensitive to local cluster environment. Here we analyse a suite of N -body simulations to study bound multiplicity and local projected stellar density, $N_{\star}$, around massive stars within gradually forming star clusters with 50 per cent primordial binaries in the Turbulent Clump Core Accretion (TCCA) paradigm. We find that massive stars rapidly gather triple or higher-order bound companions and enhancements in local $N_{\star}$ via dynamical processes. We study these metrics as a function of environment in a given cluster, quantifying the increasing multiplicity that arises towards cluster centres. We find that secondary multiplicity tends to decrease in more massive clusters due to their higher velocity dispersions, but rises as the mean density of the bound cluster increases. We find our $N_{\star}$ radial profiles are shallower compared to those in the STARFORGE simulations, which form massive stars via competitive accretion. A comparison to the AFGL 5180 system suggests it is better described by TCCA models. However, a larger number of observed systems is needed to better discriminate between these formation models.

Key words: stars: formation – stars: kinematics and dynamics – methods: numerical.

1 INTRODUCTION

Massive stars, i.e. those with masses $\geq 8 M_{\odot}$, are both products and regulators of the process of star formation across the Universe. While short-lived and less numerous than their low-mass counterparts (P. Kroupa 2001; G. Chabrier 2003), massive stars strongly influence their star-forming environment, through protostellar outflows (K. E. I. Tanaka, J. C. Tan & Y. Zhang 2017; A. Kölligan & R. Kuiper 2018; J. E. Staff, K. E. I. Tanaka & J. C. Tan 2019; J. E. Staff et al. 2023; S. M. Appel et al. 2026), ionizing and dissociating radiation, and stellar winds and supernovae (e.g. C. D. Matzner 2002; J. E. Dale, B. Ercolano & I. A. Bonnell 2012; H. Rogers & J. M. Pittard 2013; J. E. Dale et al. 2014; M. J. Butler et al. 2017; M. C. Smith, D. Sijacki & S. Shen 2018; S. C. Lewis et al. 2023). Supernova explosions lead to the chemical enrichment of the interstellar medium through the dispersal of heavy elements synthesized inside massive stars (K. Nomoto, C. Kobayashi & N. Tominaga 2013), shaping the next generation of stars that will

form in both isolated and clustered environments. Other feedback mechanisms affect star formation by restructuring the local gas distribution, possibly dispersing the parent cloud (S. C. Lewis et al. 2023; S. M. Appel et al. 2026), however they could also trigger the formation of new stars during the process (e.g. R. I. Klein, M. T. Sandford & R. W. Whitaker 1980; B. G. Elmegreen 2011; S. W. Kinoshita, F. Nakamura & B. Wu 2021). Considering their importance, the question of how massive stars form has been extensively studied both theoretically and observationally (see e.g. J. C. Tan et al. 2014; H. Beuther, R. Kuiper & M. Tafalla 2025, for reviews).

The multiplicity of stars across the mass spectrum is an important constraint on formation theories. Observationally, the number of bound companions is known to rise with stellar mass (D. Raghavan et al. 2010; H. Sana et al. 2012; S. S. R. Offner et al. 2023). In particular, massive stars have been found to have a high degree of multiplicity across observational surveys: 90 per cent of stars above $8 M_{\odot}$ have at least one companion, and 55 per cent of them are part of triples and/or higher-order multiple systems (M. Moe & R. Di Stefano 2017; S. S. R. Offner et al. 2023). These higher-order multiple systems could have different configura-

★ E-mail: gravitas908@gmail.com

tions for their members and be hierarchical in nature: for example, a triple system could have an inner binary system with a wide tertiary (A. Tanikawa et al. 2026; A. Tokovinin 2026); and a quadruple system could have two inner binaries bound to each other (M. Zummer et al. 2025). These multiple systems could be formed through multiple pathways: filament fragmentation (S.-i. Inutsuka & S. M. Miyama 1997; J. E. Pineda et al. 2015), core fragmentation (D. Guszejnov, P. F. Hopkins & M. R. Krumholz 2017; H. Kirk et al. 2017), and disc fragmentation (C. F. Gammie 2001; N. K. Reynolds et al. 2021). Remnant star-forming gas in the core or disc can lead to gas-driven migration of primordial companions during their formation (M. R. Bate & I. A. Bonnell 1997; A. Tokovinin & M. Moe 2020). In addition, this primordial multiplicity is subject to further dynamical processing, with companions gathered or lost during dynamical interactions with other stars in the forming star cluster (D. C. Heggie 1975; J. G. Hills 1975; J. E. Wall et al. 2019).

The relative contribution of these mechanisms to the observed multiplicity remains an area of active research (S. S. R. Offner et al. 2023). These mechanisms should leave imprints on the separation and mass-ratio distribution of companions, however the completeness and bias of these observations should be carefully considered. Moreover, these primordial signatures could be masked by subsequent dynamical processing. So, it is important to study the multiplicity of massive stars in young star clusters before significant dynamical interactions have occurred (S. S. R. Offner et al. 2023). To complement this, one could use N -body simulations to quantify the level of dynamical processing that occurs on a population of primordial companions around massive stars.

Different theories of massive star formation have different predictions about the primordial multiplicity of massive stars. For example, Turbulent Core Accretion (TCA; C. F. McKee & J. C. Tan 2003), especially in magnetically supported cores, can produce initially relatively isolated massive cores and thus stars. A certain level of clustering of companions around these massive stars could then develop subsequently via dynamical evolution in the environment of a forming star cluster. On the other hand, competitive accretion (I. A. Bonnell et al. 2001; P. Wang et al. 2010; M. Y. Grudić et al. 2022) predicts the ubiquitous presence of many low-mass companions around a massive star, even from its earliest stages of formation. This presents a way to distinguish between these two massive star formation models by studying the clustering of lower-mass companions around massive stars in young star-forming regions.

We expect that this may be accomplished by the consideration of surface density profiles of young stellar objects (YSOs) around massive protostars (i.e. while still accreting) and recently formed massive stars. Such surface number density profiles around massive stars have been studied using both observations (A. R. Costa Silva et al. 2022; S. Crowe et al. 2024), and simulations of star cluster formation (M. Y. Grudić et al. 2022). Such comparisons with observed systems help us to understand whether massive stars can form in relative isolation or need to have companions present, shedding light on their dominant formation mechanism.

An important complement to studies of primordial clustering around massive stars would be to investigate the effect of the star cluster environment on the ability of massive stars to capture companions through dynamical interactions. Such interactions are expected to be highly dependent on the local environment, particularly the local stellar velocity dispersion and stellar density. Both of these variables evolve with time and vary throughout

the cluster, regulating the degree of gravitational focusing around massive stars, which would increase stellar captures. Furthermore, massive stars naturally segregate towards the centres of star clusters as they lose energy through interactions with lower-mass stars (S. Chandrasekhar 1943; R. J. Allison et al. 2009; R. Domínguez et al. 2017). Consequently, even within the same cluster, massive stars are expected to have greater chances to capture companions as they could move into the cluster centre, where the local stellar density is much higher, and the dynamical environment is significantly different from the sparse outskirts. This effect may be even more pronounced if mass segregation is primordial (I. A. Bonnell et al. 2001) rather than dynamical.

To capture all these processes, N -body simulations of gradually forming star clusters of different formation time-scales are necessary. These simulations should also include accurate treatment of close encounters and regularization of close binary systems. In this paper, we use such existing N -body simulations of star cluster formation developed by J. P. Farias, J. C. Tan & S. Chatterjee (2019). In these simulations, star clusters form gradually within turbulent clumps in giant molecular clouds (GMCs) with a semi-analytic potential representing the gas that is slowly dissipated away. Massive stars are formed randomly in the star cluster, i.e. approximating the process of TCA in the limit of no primordial mass segregation of massive star formation. We perform an analysis of clustering of companions around these massive stars. We also compare with observations and massive stars formed by competitive accretion in STARFORGE simulations. In addition, the simulations involve 50 per cent of systems being primordial binaries, but with no preference for massive stars to be favoured as being in multiples. There are no primordial triples or higher-order multiples, so such systems must form dynamically during the evolution of the cluster. We track the formation of such bound systems of companions around massive stars to study the effect of star cluster environment on the dynamical capture of companions. This exploration yields benchmark statistics that can be used to compare the results of more sophisticated MHD + N -body simulations to be developed in future studies.

2 METHODS

In Section 2.1, we describe the N -body models of gradual star cluster formation used for this paper. We have post-processed these simulations to study the evolution of stellar clustering around massive stars. In Section 2.2, we describe the process of identifying bound systems around massive stars, which are later compared to study how the environment and formation time-scale affect these bound systems. In Section 2.3, we describe how we construct projected stellar density profiles around massive stars to compare different models of massive star formation with each other and with observations.

2.1 Star cluster formation simulations

In this work, we use the Turbulent Clump Core Accretion (TCCA) star cluster formation models developed by J. P. Farias, J. C. Tan & S. Chatterjee (2017, 2019) and J. P. Farias & J. C. Tan (2023), where star clusters are formed inside magnetized, turbulent, gravitationally bound, and initially starless clumps within GMCs. These consist of a suite of unique N -body simulations focused on accurately following stellar dynamics in the early stages of star cluster formation, i.e. where there is still ongoing

formation of stars. These models assume individual stars and primordial binary stars form via stochastic core accretion with initial clump properties based on the TCA model (C. F. McKee & J. C. Tan 2003). J. P. Farias et al. (2017, hereafter Paper I) presented the extreme case of instantaneous star cluster formation, where all stars are formed together at once and remaining gas is assumed to leave the system instantaneously according to a given global star formation efficiency (ϵ). J. P. Farias et al. (2019) and J. P. Farias & J. C. Tan (2023, hereafter Papers II and III, respectively) presented the case of gradual star cluster formation. In these simulations, we assume that a fraction (ϵ) of the gas is gradually converted into stars while the remaining gas is also steadily removed from the system due to stellar feedback. This is modelled as a linear decrease of the gas reservoir (equation 9, Paper III) until it reaches zero at the end of star formation, t_* . The time-scale over which this process occurs is set by the star formation efficiency per free-fall time, ϵ_{ff} , such that higher values of ϵ_{ff} result in shorter t_* (see Paper III for details). The case of $\epsilon_{\text{ff}} = \infty$ corresponds to the instantaneous case where all stars are formed at once, whereas the slowest case considered assumes $\epsilon_{\text{ff}} = 0.01$. For this paper, we post-processed clusters formed from N -body models outlined in Paper III.

In these pure N -body models, we emulate the parent cloud gas using an analytic potential that follows the structure from the TCCA model, i.e. a turbulent gas clump of mass M_{cl} embedded in the surrounding molecular cloud. As mentioned above, the mass of this gas clump is linearly reduced as the star cluster forms, and the gas mass reaches zero at t_* . The parent cloud has a mass surface density, Σ_{cloud} , which sets the bounding pressure of the clump: a clump of a given mass in a higher Σ_{cloud} environment will be subject to higher pressure and thus be smaller and more dense. The characteristic radius of the clump (R_{cl}) in a Σ_{cloud} environment is

$$R_{\text{cl}} = 0.365 \left(\frac{M_{\text{cl}}}{3000 M_{\odot}} \right)^{1/2} \left(\frac{\Sigma_{\text{cloud}}}{1 \text{ g cm}^{-2}} \right)^{-1/2} \text{ pc}. \quad (1)$$

The clump is modelled as a polytropic sphere of gas with a three-dimensional density profile of the form:

$$\rho_{\text{cl}}(r) = \rho_{\text{s,cl}} \left(\frac{r}{R_{\text{cl}}} \right)^{-k_{\rho}}. \quad (2)$$

The density thus varies radially within the clump and is defined within the clump radius R_{cl} . Here, $\rho_{\text{s,cl}}$ is the density of clump at its edge ($r = R_{\text{cl}}$). It is given by

$$\rho_{\text{s,cl}} = \frac{(3 - k_{\rho}) M_{\text{cl}}}{4\pi R_{\text{cl}}^3}. \quad (3)$$

Similarly, the velocity dispersion profile of the clump is given by

$$\sigma_{\text{cl}}(r) = \sigma_{\text{s}} \left(\frac{r}{R_{\text{cl}}} \right)^{(2-k_{\rho})/2}, \quad (4)$$

where σ_{s} is the velocity dispersion at the surface of the clump, given by

$$\sigma_{\text{s}} = 3.04 \left(\frac{M_{\text{cl}}}{3000 M_{\odot}} \right)^{1/4} \left(\frac{\Sigma_{\text{cloud}}}{1 \text{ g cm}^{-2}} \right)^{1/4} \text{ km s}^{-1}. \quad (5)$$

We set the exponent $k_{\rho} = 1.5$ following Paper II, which are motivated by the observational results of M. J. Butler & J. C. Tan (2012). The stars are formed following the above gas density (equation 2) and velocity dispersion (equation 4) profiles of the parent clump. It is evident that clusters formed in higher M_{cl}

and Σ_{cloud} models have larger stellar velocity dispersion σ_{s} . These simulations have no primordial mass segregation, however some degree of mass segregation is developed through dynamical processes, especially in the lower mass clusters (refer to section 4.4 of Paper III).

The time taken for the star cluster to form inside the clump is given by

$$t_* = \frac{\epsilon}{\epsilon_{\text{ff}}} t_{\text{ff},0}, \quad (6)$$

where $t_{\text{ff},0}$ is the initial global free-fall time-scale of the clump,

$$t_{\text{ff},0} = 0.069 \left(\frac{M_{\text{cl}}}{3000 M_{\odot}} \right)^{1/4} \left(\frac{\Sigma_{\text{cloud}}}{1 \text{ g cm}^{-2}} \right)^{-3/4} \text{ Myr}, \quad (7)$$

in the TCCA model (equation 9, Paper II). We find that for the same value of ϵ_{ff} , the time-scale t_* is shorter for higher Σ_{cloud} and/or smaller M_{cl} , because the free-fall time-scale $t_{\text{ff},0}$ is shorter in higher Σ_{cloud} and/or smaller M_{cl} environments (see equation 7).

For our simulations, we fix ϵ to 0.5. For instance, for $M_{\text{cl}} = 3000 M_{\odot}$ and $\epsilon = 0.5$, we form $1500 M_{\odot}$ in stellar mass. We choose three different values of M_{cl} : 300, 3000, and 30 000 $M_{\odot}$. For each M_{cl} , we consider star clusters formed from clumps in two different cloud environments: set ‘L’ formed from clumps in lower mass surface density clouds ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$), and set ‘H’ formed from clumps in higher mass surface density clouds ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$). Next, for each Σ_{cloud} , we consider five values for ϵ_{ff} : 0.01, 0.03, 0.10, 0.30, and 1.00. Table 1 shows the full set of cluster parameters (M_{cl} , Σ_{cloud} and ϵ_{ff}) modelled in this work. For each row of Table 1 we have run multiple realizations (N_{sims}) shown in Column 2. The number of stars formed in a realization varies because of the random sampling of the IMF, so the average number of stars (and thus the average stellar mass) also differs slightly across models. However, on average, our models form around $\langle N_{\text{stars}} \rangle \approx 4 \times 10^2$, 4×10^3 and 4×10^4 stars in a realization of M300, M3000, and M30000 models, respectively. The resulting average stellar mass in our models is $0.375 M_{\odot}$.

Each realization has an IMF sampling (P. Kroupa 2001) and 50 percent primordial binary population as described in J. P. Farias & J. C. Tan (2023). However, we do not have primordial triples, quadruples or higher-order multiples. Any systems with three or more bound stars must have formed via dynamical processing. We track the formation of such systems around massive stars studied for this work. To quantify the formation of such bound systems, we calculate the fraction of massive stars that have triple or higher-order companions around them. This fraction called the Triple/Higher-order fraction (THF) is given by

$$\text{THF} = \frac{T + Q + P + \dots}{S + B + T + Q + P + \dots}, \quad (8)$$

where S, B, T, Q, P are the number of single stars, binaries, triples, quadruples, pentuples, and so on (S. S. R. Offner et al. 2023).

2.2 Clustering around massive stars

We examine the stellar clustering around massive stars ($\geq 8 M_{\odot}$) formed in our simulations. To study the effect of cluster environment on such clustering, we divide these massive stars into three different subsets according to their location and boundedness. First, we determine stars that are bound to the star cluster using a snowballing algorithm developed in J. P. Farias & J. C.

Table 1. Overview of star cluster formation simulations. Column 1 is the name of the set. Column 2 is the number of simulations (N_{sims}) run in each set. Column 3 is the star formation efficiency per free-fall time (ϵ_{ff}) for the set. Column 4 is the mass surface density of the parent cloud (Σ_{cloud}). Column 5 is the mass of the gas clump (M_{cl}). Column 6 is the number of stars ($\langle N_{\text{stars}} \rangle$) formed on average in the model. Columns 7 and 8 are the star formation time and initial free-fall time of the clump, respectively. Column 9 is the radius of the gas clump (R_{cl}). Column 10 is the velocity dispersion at the surface of the clump.

Set name	N_{sims}	ϵ_{ff}	Σ_{cloud} (g cm^{-2})	M_{cl} ($M_{\odot}$)	$\langle N_{\text{stars}} \rangle$	$t_{\star}$ (Myr)	$t_{\text{ff},0}$ (Myr)	R_{cl} (pc)	σ_s (km s^{-1})
M300L	200	0.01	0.1	300	4×10^2	10.91	0.22	0.36	0.96
	200	0.03	0.1	300	4×10^2	3.64	0.22	0.36	0.96
	200	0.1	0.1	300	4×10^2	1.09	0.22	0.36	0.96
	200	0.3	0.1	300	4×10^2	0.36	0.22	0.36	0.96
	200	1.0	0.1	300	4×10^2	0.11	0.22	0.36	0.96
M300H	200	0.01	1.0	300	4×10^2	1.94	0.039	0.115	1.71
	200	0.03	1.0	300	4×10^2	0.65	0.039	0.115	1.71
	200	0.1	1.0	300	4×10^2	0.19	0.039	0.115	1.71
	200	0.3	1.0	300	4×10^2	0.06	0.039	0.115	1.71
	200	1.0	1.0	300	4×10^2	0.02	0.039	0.115	1.71
M3000L	20	0.01	0.1	3000	4×10^3	19.40	0.39	1.15	1.71
	20	0.03	0.1	3000	4×10^3	6.47	0.39	1.15	1.71
	20	0.1	0.1	3000	4×10^3	1.94	0.39	1.15	1.71
	20	0.3	0.1	3000	4×10^3	0.65	0.39	1.15	1.71
	20	1.0	0.1	3000	4×10^3	0.19	0.39	1.15	1.71
M3000H	20	0.01	1.0	3000	4×10^3	3.45	0.069	0.365	3.04
	20	0.03	1.0	3000	4×10^3	1.15	0.069	0.365	3.04
	20	0.1	1.0	3000	4×10^3	0.35	0.069	0.365	3.04
	20	0.3	1.0	3000	4×10^3	0.12	0.069	0.365	3.04
	20	1.0	1.0	3000	4×10^3	0.03	0.069	0.365	3.04
M30000L	2	0.01	0.1	30 000	4×10^4	34.50	0.69	3.65	3.04
	2	0.03	0.1	30 000	4×10^4	11.50	0.69	3.65	3.04
	2	0.1	0.1	30 000	4×10^4	3.45	0.69	3.65	3.04
	2	0.3	0.1	30 000	4×10^4	1.15	0.69	3.65	3.04
	2	1.0	0.1	30 000	4×10^4	0.34	0.69	3.65	3.04
M30000H	2	0.01	1.0	30 000	4×10^4	6.14	0.123	1.154	5.41
	2	0.03	1.0	30 000	4×10^4	2.05	0.123	1.154	5.41
	2	0.1	1.0	30 000	4×10^4	0.61	0.123	1.154	5.41
	2	0.3	1.0	30 000	4×10^4	0.20	0.123	1.154	5.41
	2	1.0	1.0	30 000	4×10^4	0.06	0.123	1.154	5.41

Tan (2023). Then, we refer to bound massive stars in the central regions within 20 per cent of the half-mass radius of their star cluster as *central stars*. Similarly, massive stars beyond 20 per cent of the half-mass radius which are in the outer regions of the star cluster are referred to as *offset stars*. Finally, massive stars that are unbound from the star cluster are referred to as *ejected stars*.

After dividing massive stars into subsets, we study the properties of bound and unbound companions around them. For this purpose, we have developed routines to identify bound companions around any given massive star. The routine works as follows: First, we identify the closest bound companion (a binary partner) to the massive star. Then, we replace the two stars with a centre-of-mass particle. Next, we search for the next closest star that is bounded to this centre-of-mass particle, hence forming a triple. If present, we replace the triple with its corresponding centre of mass particle. Then, we search for the next closest bound companion. We perform this process iteratively to identify all the stars that are bound to the massive star. We start this search around the most massive star, identify its bound companions and tag them. Then, we move to the next most massive untagged star and search for its bound companions, ensuring there is no double-counting of stars during this iterative process. This allows us to identify all the bound stars that are centred on independent primary massive stars within our star cluster.

At each step, the bound systems are identified based on the total kinetic and potential energies of the components, combined with a tidal force criterion (equation 9) to remove unstable companions. If m_S and m_* are the masses of the two particles under consideration separated by distance d_1 , and m_p is the mass of a perturber star at a distance d_2 from m_* , we are checking if m_* is bound to m_S under perturbation from a perturber star m_p :

$$\left| \frac{m_S m_*}{d_1^2} \right| > \left| K \frac{m_p m_*}{d_2^2} \right|. \quad (9)$$

Here m_S could be a single star or a centre-of-mass particle in the case of an inner system of bound stars. We set $K = 10$. To calculate the internal and external forces, we iterate and sum up contributions from all stars internal and external to the bound system, respectively.

We denote N_{bound} as the total number of bound companions for a given massive star. We study any correlations between N_{bound} and the mass of the massive star M_* , and its radial location in the star cluster. We define the radial location of the star in its star cluster using a radial coordinate $r[r_h] = r_*/r_h$, where we normalize the radial distance of the star (r_*) from the centre by the half-mass radius (r_h) of the bound cluster. The centre of the bound cluster is computed using the snowballing algorithm of J. P. Farias & J. C. Tan (2023) mentioned above.

2.3 Projected stellar density profiles

After the identification of bound and unbound stars around every massive star, we construct bound and total (bound + unbound) projected stellar density profiles for the massive star. For this, we create annular bins centred on the massive star. For the linear case, we create annular bins of width 0.05 pc centred on the massive star – the innermost bin is a circle of radius 0.05 pc. For logarithmic case, we choose N_{bins} 30 bins between 10^{-5} pc and 1 pc to probe the small-scale distribution near the massive star. The distances between the stars are 2D separations measured in projection along the z -axis in the simulation. Then, we count the number of stars $n_*(r)$ in each radial bin of area $A(r)$ located at radius r from the massive star. We convert this into a projected 2D stellar density N_* :

$$N_*(r) = \frac{n_*(r)}{A(r)}. \quad (10)$$

Here, $n_* = n_{\text{bound}} + n_{\text{unbound}}$, where n_{bound} is the number of bound companions and n_{unbound} is the number of unbound companions in a given radial bin. Adding up bound companions in all bins, we recover N_{bound} . Using equation (10), we construct the bound and total projected stellar density profiles around every massive star. We compare these profiles across our full range of models and different cluster ages. We also compare them with observational data to explore our questions about theoretical models of massive star formation.

3 RESULTS

In Section 3.1, we present our results on the variation of the number of bound companions around a massive star with its location and mass. In Section 3.2, we compare massive stars across star clusters forming at different time-scales in our $(M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}})$ parameter space. In Section 3.3, we present the evolution of projected stellar density profiles around these massive stars. In Section 3.4, we compare these profiles and the properties of bound companions with observations.

Before we begin, we note that typical, well-studied young star-forming regions (ages $\lesssim 10$ Myr) have a wide range of ages: 1–2 Myr (e.g. Ophiuchus; B. A. Wilking, M. Gagné & L. E. Allen 2008), 2–5 Myr (e.g. Taurus, Lupus, Chamaeleon I; S. J. Kenyon & L. Hartmann 1995; K. L. Luhman 2004; J. M. Alcalá et al. 2014), and 5–10 Myr (e.g. Upper Scorpius; M. J. Pecaut, E. E. Mamajek & E. J. Bubar 2012). One of the most notable is the Orion Nebula Cluster with stellar ages between ~ 1 and ~ 5 Myr (e.g. N. Da Rio et al. 2010; M. Reggiani et al. 2011; M. Fajrin et al. 2025). These age estimates have their own systematic uncertainties (e.g. T. Preibisch 2012), which are beyond the scope of our work. So, we have picked 3 Myr as a fiducial age for a typical young star-forming region. This adopted reference time can influence the trends we observe during our model comparisons at a given cluster age. Our different star clusters will have reached different stages at 3 Myr in terms of how advanced the star formation process is. However, this is also true for the observed systems; and unlike the simulations, we do not know which exact stage of the process a given star-forming region is currently in. So, in order to make a fair comparison between theoretical models in this work, we need to pick a fixed physical time (3 Myr). When we compare models with observations as in Section 3.4, we generate the metrics at different ages to account for the uncertainty in our assumption for the typical age of the star-forming region.

3.1 Effect of location and stellar mass

We classified massive stars into three subsets: *central stars*, *offset stars* and *ejected stars* to study the effect of cluster environment on their bound companions. Table 2 presents statistics of primary massive stars across these subsets in our simulations across all $(M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}})$ models at $t = 3$ Myr. Figures in this work show the results for our fiducial $\epsilon_{\text{ff}} = 0.03$ model, unless otherwise specified. These primary massive stars have their own independent systems of bound companions (if any). We find that *ejected* subset comprised a range of 1–30 per cent of the primary massive star population at $t = 3.0$ Myr, depending on the model. The dynamical state of the forming star cluster determined the relative fraction of *central* and *offset* stars in the remaining stars that are still bound to the cluster.

In Fig. 1, we show the mass M_* of the massive stars versus the number of bound companions (N_{comp}) around them. We find that more massive stars tend to have more companions than lower-mass stars, as expected. However, N_{bound} depends on the environment of the massive star as well. For similar M_* , two massive stars can have different N_{bound} in the same star cluster depending on their location. We find that a *central star* tends to have more companions than an *ejected star* of the same mass. For instance, we find that *ejected stars* tend to be mostly singles ($N_{\text{bound}} = 0$) or binaries ($N_{\text{bound}} = 1$) across all our models. Only a few of them have more than one companion, with those companions having small mass ratios compared to the massive star.

To further illustrate these effects, in Fig. 2, we show N_{comp} versus the normalized radial location ($r[r_h]$) of the massive stars. First, we can infer the distribution of the three subsets of massive stars inside their star clusters. For instance, the *ejected stars* are located mostly beyond 0.5 times the half-mass radius of their cluster. Similarly, we observe that many massive stars ($\geq 20M_{\odot}$) have segregated to the centre of their star cluster ($r[r_h] \leq 0.2$). Next, we find that there is some correlation between N_{bound} and $r[r_h]$. *Central stars* have more companions in general than *offset stars*. Also, some *central stars* deeper in the star cluster have more companions than *central stars* of similar mass. In addition to the variation of N_{bound} between massive stars in the same star cluster, there are clear variations in N_{bound} across star clusters in different models (compare different panels of Figs 1 and 2) of our parameter space $(M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}})$. We discuss these differences in Section 3.2.

3.2 Comparison across models

In order to understand the effect of the environment of a forming star cluster on stellar clustering around massive stars, we compared the statistics of bound companions around these massive stars across our full range of models $(M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}})$ outlined in Table 1. For this, we averaged the number of bound companions N_{bound} across all massive stars in each subset to find the average number of bound companions ($\langle N_{\text{bound}} \rangle$). Then, we studied the variation of $\langle N_{\text{bound}} \rangle$ with our model parameters (M_{cl} , Σ_{cloud} and ϵ_{ff}).

Fig. 3 shows $\langle N_{\text{bound}} \rangle$ around massive stars in different M_{cl} models at a cluster age of $t = 3.0$ Myr. In addition to *central stars* having more companions than *ejected stars* for all M_{cl} models, we find that massive stars in the smaller M_{cl} clusters have larger $\langle N_{\text{bound}} \rangle$ compared to those in higher-mass clusters. This is attributed to the lower velocity dispersion (σ_s) for smaller M_{cl} clusters in our TCM models (refer to Column 10 of Table 1).

Table 2. Statistics of primary massive stars ($M_{\text{prim}} \geq 8 M_{\odot}$) categorized into *central*, *offset* and *ejected* subsets in our star cluster formation simulations. Primary massive stars are massive stars that act as primaries for their own independent bound stellar systems, so there is no double counting of stars. The statistics presented in this table were calculated at a cluster age $t = 3.0$ Myr. Since these are models of gradual star cluster formation (varying ϵ_{ff} , Column 2), some models are still forming stars (including massive stars) while others have fully formed their star clusters by this age, leading to a variation in N_{prim} at $t = 3.0$ Myr (Column 3). Columns 4, 5, and 6 show the relative number of *central*, *offset*, and *ejected* stars. Columns 7–9 show the number of *central* stars that are singles (S: $N_{\text{bound}} = 0$), in binaries (B: $N_{\text{bound}} = 1$), or part of triple/higher-order bound stellar systems (TH: $N_{\text{bound}} \geq 2$). Columns 10–12 show the same for *offset* stars, and Columns 13–15 for *ejected* stars, respectively.

Set name	ϵ_{ff}	N_{prim} ($M_{\text{prim}} \geq 8 M_{\odot}$)	N_{prim} (central)	N_{prim} (offset)	N_{prim} (ejected)	$N_{\text{prim,central}}$			$N_{\text{prim,offset}}$			$N_{\text{prim,ejected}}$		
						S	B	TH	S	B	TH	S	B	TH
M300L	0.01	204	69	134	1	19	28	22	58	58	18	1	0	0
	0.03	214	128	73	13	39	33	56	27	34	12	7	6	0
	0.10	255	161	76	18	31	44	86	29	29	18	13	3	2
	0.30	248	138	75	35	28	38	72	24	27	24	17	16	2
	1.00	268	134	88	46	33	37	64	22	27	39	21	19	6
M300H	0.01	276	166	75	35	23	43	100	33	19	23	22	11	2
	0.03	190	122	42	26	17	29	76	12	13	17	16	7	3
	0.10	243	146	56	41	20	36	90	12	16	28	24	14	3
	0.30	256	157	52	47	26	41	90	6	12	34	17	20	10
	1.00	270	132	67	71	15	34	83	10	15	42	30	23	18
M3000L	0.01	57	5	47	5	3	2	0	34	12	1	4	1	0
	0.03	140	11	122	7	4	3	4	69	50	3	4	3	0
	0.10	251	39	170	42	11	14	14	95	69	6	21	20	1
	0.30	264	64	133	67	31	25	8	77	44	12	40	25	2
	1.00	259	61	121	77	23	27	11	68	40	13	44	30	3
M3000H	0.01	197	48	136	13	14	25	9	81	50	5	8	5	0
	0.03	278	63	139	76	31	23	9	98	36	5	53	23	0
	0.10	258	60	118	80	26	24	10	73	35	10	50	25	5
	0.30	254	71	106	77	32	24	15	51	40	15	48	23	6
	1.00	261	63	110	88	26	20	17	61	30	19	52	31	5
M30000L	0.01	20	3	17	0	2	1	0	11	6	0	0	0	0
	0.03	68	9	57	2	5	4	0	35	22	0	2	0	0
	0.10	225	19	185	21	12	7	0	114	70	1	13	8	0
	0.30	259	37	167	55	22	11	4	101	62	4	37	18	0
	1.00	255	34	160	61	20	13	1	95	62	3	32	28	1
M30000H	0.01	142	10	130	2	7	1	2	89	40	1	2	0	0
	0.03	268	43	181	44	25	17	1	109	71	1	31	13	0
	0.10	260	38	174	48	24	13	1	112	58	4	27	21	0
	0.30	262	55	142	65	34	19	2	90	51	1	39	26	0
	1.00	273	38	169	66	28	8	2	116	49	4	49	17	0

Finally, the *central* massive stars in higher Σ_{cloud} models have slightly more companions than lower Σ_{cloud} models. This is due to the smaller half-mass radius of these higher Σ_{cloud} models (refer to Column 9 of Table 1), which leads to higher stellar densities ($n_{\text{s,b}}$) in the bound cluster. We investigate the variation of $\langle N_{\text{bound}} \rangle$ with σ_{S} and $n_{\text{s,b}}$ later in this section.

We studied the potential effect of observational limits on the statistics of $\langle N_{\text{bound}} \rangle$. These limits mean only companions above a certain mass would be observable. For this, we considered three cases for companion masses: $M_{\text{comp}} \geq 0.01 M_{\odot}$ (fiducial case), $M_{\text{comp}} \geq 0.1 M_{\odot}$ and $M_{\text{comp}} \geq 1.0 M_{\odot}$. Here, $0.01 M_{\odot}$ is the lowest possible mass of objects in our star cluster simulations and $0.1 M_{\odot}$ is the observational limit on companion mass from the AFGL 5180 observations, which we later compare with our simulated companions. We chose $1.0 M_{\odot}$ to follow a uniform logarithmic order in companion mass. Fig. 3 shows the variation of $\langle N_{\text{bound}} \rangle$ with these choices for the low and high Σ_{cloud} models in top and bottom rows, respectively. We find that if we consider only higher-mass companions by increasing the companion mass threshold (M_{comp}), the metric $\langle N_{\text{bound}} \rangle$ goes down as expected. The decline is more prominent for the lower M_{cl} models, which indicates that massive stars in these models have many low-mass companions below $0.1 M_{\odot}$. In addition, the $\langle N_{\text{bound}} \rangle$ around *ejected* stars is

not as affected as much as *central* or *offset* stars. This indicates that these *ejected* stars usually have higher-mass companions ($> 1.0 M_{\odot}$), if any.

Fig. 4 shows $\langle N_{\text{bound}} \rangle$ around massive stars at cluster age of $t = 3.0$ Myr compared across ϵ_{ff} models. We find that *central* stars have larger N_{bound} than *offset* and *ejected* stars across almost all models. We also notice there is a gradual rise in N_{bound} with ϵ_{ff} in the M300L model. However, the variation of N_{bound} with ϵ_{ff} is not monotonous in other models. Nonetheless, this hints that the number of bound companions around a massive star can be affected by the formation time-scales of its star cluster.

This is primarily due to the time evolution of stellar velocity dispersion (σ_{S}) and stellar number density ($n_{\text{s,b}}$) around the massive star during the formation of the star cluster. Fig. 5 shows the variation of $\langle N_{\text{bound}} \rangle$ with these environmental variables for both *central* stars (left column) and *offset* stars (right column). For *central* stars, we calculate σ_{S} and $n_{\text{s,b}}$ in the central core, within 20 per cent of the half-mass radius, of the bound cluster. For *offset* stars, we calculate σ_{S} and $n_{\text{s,b}}$ in a spherical shell extending from 0.2 to 3.0 times the half-mass radius. We observe that the stellar densities ($n_{\text{s,b}}$) are smaller around *offset* stars than *central* stars, as expected by their location.

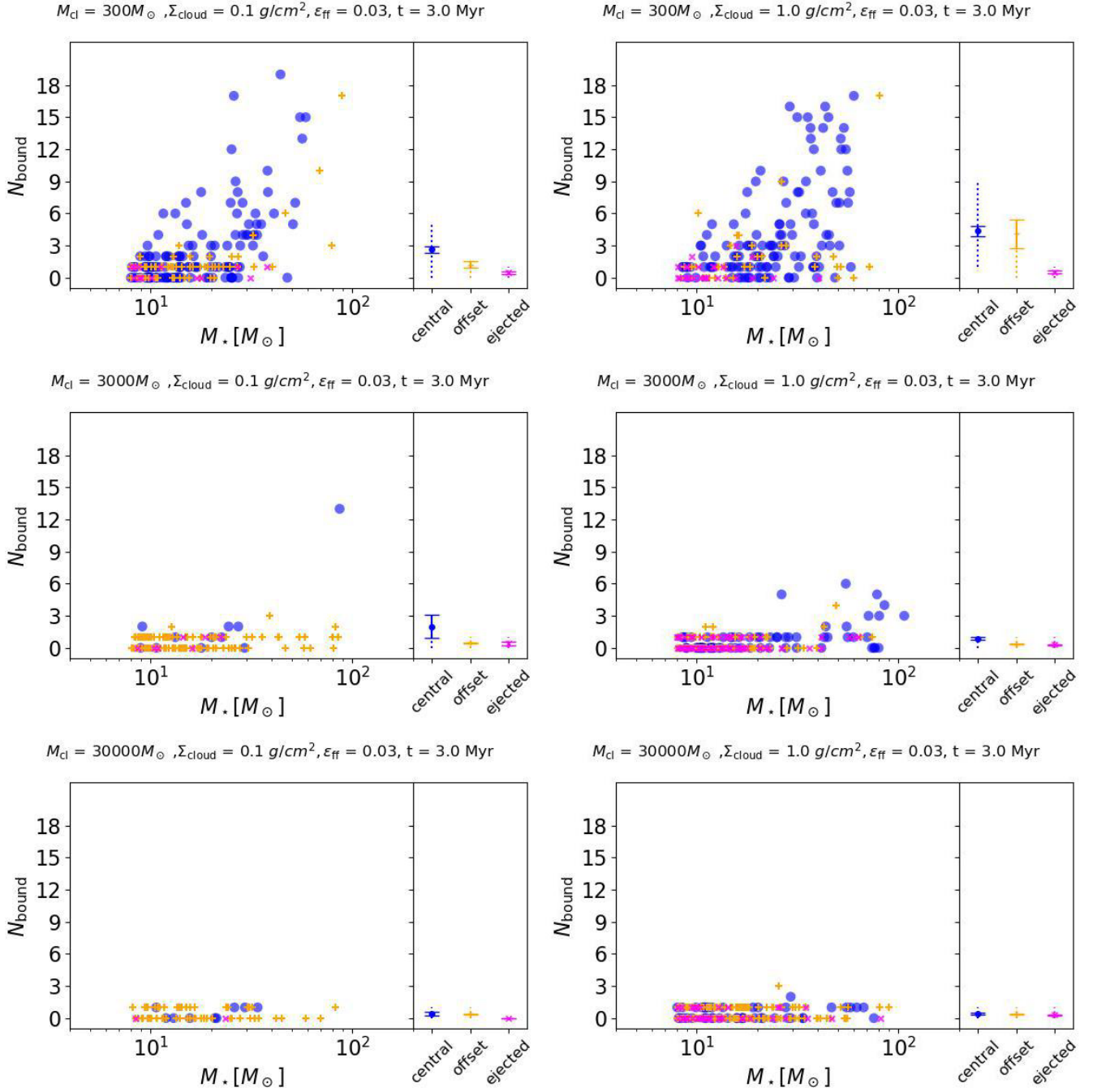


Figure 1. Masses of massive stars (M_*) versus the number of bound companions (N_{bound}) around them, in our fiducial $\epsilon_{\text{ff}} = 0.03$ models at cluster age $t = 3.0$ Myr. From top to bottom, the rows show M300, M3000, and M30000 models, respectively. The left and right columns show low ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) and high ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) cloud mass surface density cases, respectively. In each panel, the coloured symbols (blue circles, orange pluses, magenta crosses) show the three subsets of massive stars: *central* stars, *offset* stars, and *ejected* stars, respectively. In each side panel, we show the mean number of companions, $\langle N_{\text{bound}} \rangle$, averaged around massive stars in each subset. The dotted errorbar indicates the spread of N_{bound} from the 16th percentile to the 84th percentile, while the solid errorbar shows the standard error in $\langle N_{\text{bound}} \rangle$. We find that more massive stars have a larger number of bound companions, however this also depends on their location within the star cluster.

For both *central* and *offset* stars, we find that $\langle N_{\text{bound}} \rangle$ is greater for smaller values of stellar velocity dispersion (σ_s , see first row). The dependence of $\langle N_{\text{bound}} \rangle$ on stellar density ($n_{s,b}$) is more complex (see second row). If we check the third row, we find that $\langle N_{\text{bound}} \rangle$ is maximum for models with lower velocity dispersion and higher stellar density (tending towards the upper left cor-

ner). These happen to be the M300H models with the highest $\langle N_{\text{bound}} \rangle$, as also seen in Fig. 3. This is true for both *central* and *offset* stars. This shows that the prevailing stellar velocity dispersion and number density around a massive star in a forming star cluster can determine the number of bound companions it gathers.

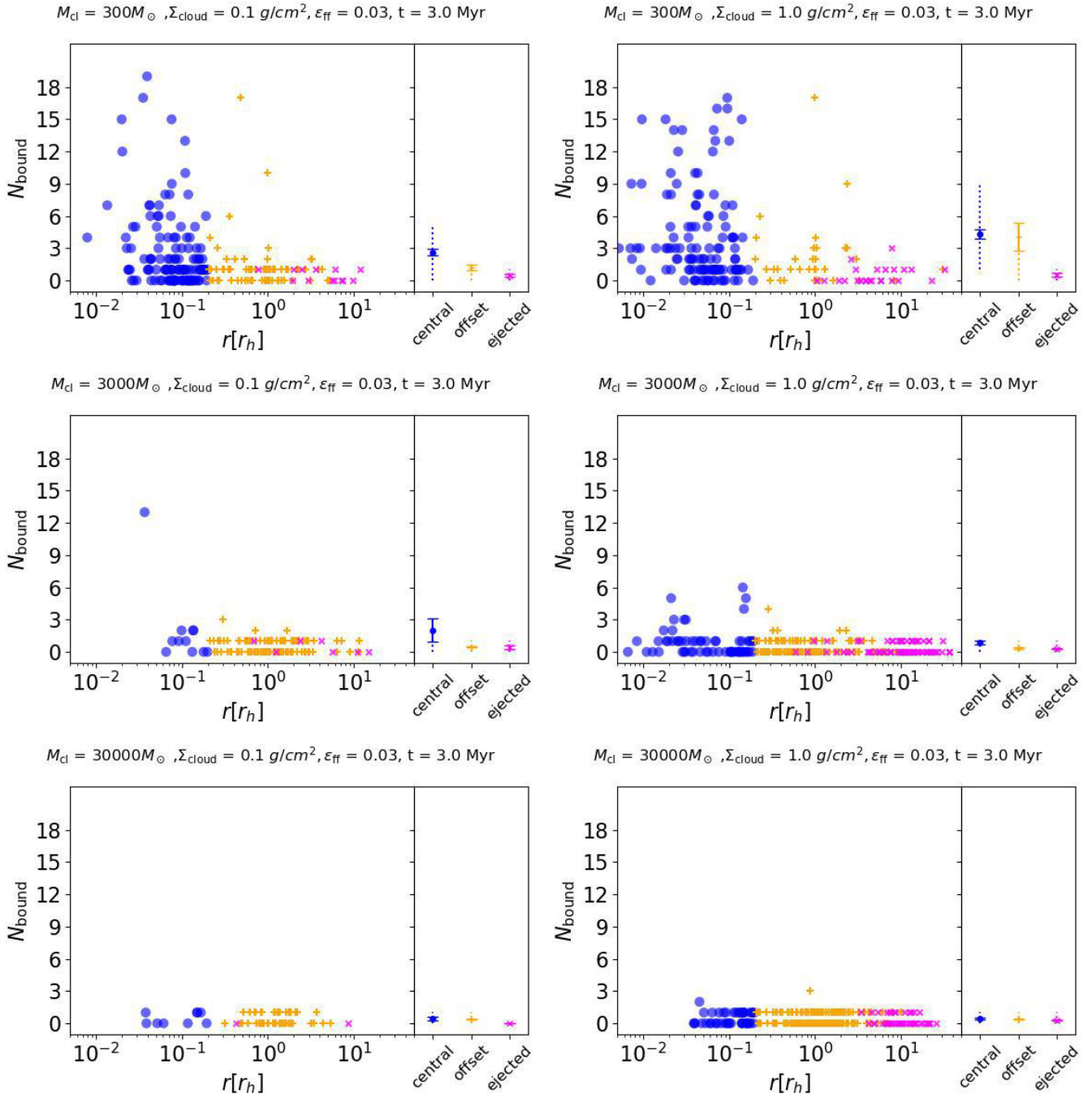


Figure 2. Normalized radial locations of massive stars (r_h) versus number of bound companions (N_{bound}) around them, in our fiducial $\epsilon_{\text{ff}} = 0.03$ models at cluster age $t = 3.0$ Myr. From top to bottom, the rows show M300, M3000, and M30000 models, respectively. The left and right columns show low ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) and high ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) cloud mass surface density cases, respectively. In each panel, the coloured symbols (blue circles, orange pluses, magenta crosses) show three subsets of massive stars: *central* stars, *offset* stars, and *ejected* stars, respectively. In each side panel, we show the mean number of companions, $\langle N_{\text{bound}} \rangle$, averaged around massive stars in each subset. The dotted errorbar indicates the spread of N_{bound} from the 16th percentile to the 84th percentile, while the solid errorbar shows the standard error in $\langle N_{\text{bound}} \rangle$. We find that massive stars which are closer to the centre of the star cluster generally have more bound companions.

The reasoning is as follows: bound systems of multiple stars form and disrupt through stellar encounters. The size of these bound systems will be determined by the balance of captures and disruptive encounters. The number of possible interactions will depend on the number density ($n_{\text{s,b}}$) around the massive star, because it determines the availability of stars which could

undergo such encounters (both captures and disruptions). However, the velocity dispersion (σ_s) of the environment determines whether captures or disruptions win. This is because σ_s controls the strength of the gravitational focusing (v_{max}/σ_s) experienced by companions during capture. A smaller σ_s environment increases the gravitational focusing by increasing the impact pa-

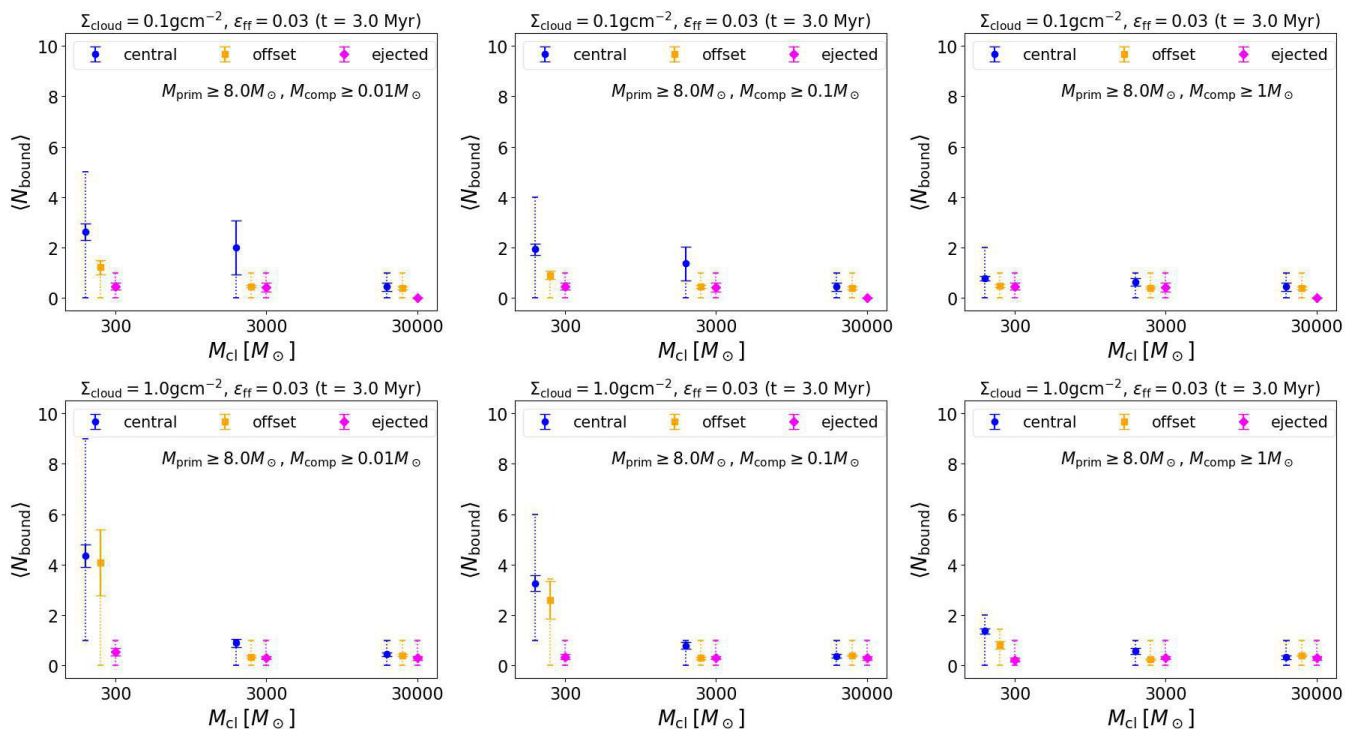


Figure 3. Each panel in the grid shows the comparison between different M_{cl} models showing the average number of bound companions ($\langle N_{\text{bound}} \rangle$) around massive stars in that model. Panels in the first and second rows show the low ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) and high ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) cloud mass surface density cases, respectively, for the fiducial ($\epsilon_{\text{ff}} = 0.03$) models at cluster age of $t = 3.0 \text{ Myr}$. In the columns, we vary the threshold for companion mass (M_{comp}) between 0.01 (fiducial), 0.1, and $1 M_{\odot}$. Each panel, on the y-axis, shows $\langle N_{\text{bound}} \rangle$ around massive stars in three subsets: *central* stars (blue), *offset* stars (orange) and *ejected* stars (magenta). (...) indicates averaging over massive stars in all realizations of each model. The points for different subsets are offset by [0.1, 0.0, -0.1] along x-axis to allow for visual comparison. For each point, the dotted errorbar indicates the spread of N_{bound} from the 16th percentile to the 84th percentile, while the solid errorbar shows the standard error in $\langle N_{\text{bound}} \rangle$. We find that massive stars in smaller M_{cl} star clusters with lower velocity dispersion have more bound companions.

parameter. In denser environments with larger velocity dispersions, captures become more difficult, as encounters are on average more energetic. Among our models, the low M_{cl} clusters provide environments that favour captures over disruptions because of their lower velocity dispersions, and thus M300 models have higher $\langle N_{\text{bound}} \rangle$.

We note that dynamical mass segregation does contribute to the increase in $\langle N_{\text{bound}} \rangle$, as more massive stars move to the more crowded centre (higher $n_{\text{S,b}}$) with more potential companions that could be captured. While we have found mass segregation to be more efficient in our low M_{cl} clusters (see Paper II, section 4.3.1), $\langle N_{\text{bound}} \rangle$ will depend more on the σ_{S} of the environment rather than mass segregation alone. This is because velocity dispersion ultimately determines the balance of captures that form the multiple systems versus the encounters that disrupt them.

Variations in these environmental variables (σ_{S} and $n_{\text{S,b}}$) in the same cluster and between star clusters forming at different time-scales lead to the variation in $\langle N_{\text{bound}} \rangle$ between massive stars at a given cluster age. This also means that as the cluster forms and evolves, $\langle N_{\text{bound}} \rangle$ will evolve with time due to changes in σ_{S} and $n_{\text{S,b}}$. In Fig. 6, we show the evolution of $\langle N_{\text{bound}} \rangle$ around massive stars with cluster age t in Myr across our full set of models ($M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}}$). Each model has its own formation time-scale ($t_{\star}$) indicated by the dashed vertical line in each panel. The star cluster is still embedded in background gas before $t_{\star}$. We find that across all models, there is a gradual rise in N_{bound} before a

decline and/or flattening as the cluster ages. This indicates that the formation time-scale and stellar environment (density and velocity dispersion) have a significant effect on $\langle N_{\text{bound}} \rangle$.

Besides $\langle N_{\text{bound}} \rangle$, the build-up of bound stellar systems around massive stars is also reflected in the triple/higher-order fraction of massive stars (THF; equation 8). Our simulations have only 50 per cent primordial binaries, and triple/higher-order companions are gathered dynamically over time. These companions can be lost or replenished over time in the star cluster environment, so THF evolves over time. Fig. 7 shows the time evolution of THF for the full set of our models ($M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}}$). We find that for *central* and *offset* stars, THF usually rises before flattening and/or declining as the star cluster ages. Depending on the time evolution of σ_{S} and $n_{\text{S,b}}$, and thus dynamical captures of the model, this rise in THF occurs quickly in the first Myr, whereas for others, it takes more time. There is also variation across models in the flattening or fall of THF once gas expulsion is complete. M300 models have higher THF, which reflects their larger $\langle N_{\text{bound}} \rangle$. In these smaller M_{cl} clusters, THF rises quickly and flattens. In contrast, THF in M3000 clusters rises and then falls gradually over time. In the M30000 clusters, triple and higher-order bound systems are rarely formed, so THF remains at a small level and does not change much over time ($\text{THF} < 0.2$). Finally, we note that the evolution of THF around *offset* stars is more gradual with the THF levels mostly flattening over time, whereas for *central* stars, there can be flattening or decline. THF remains low around *ejected* stars across all models.

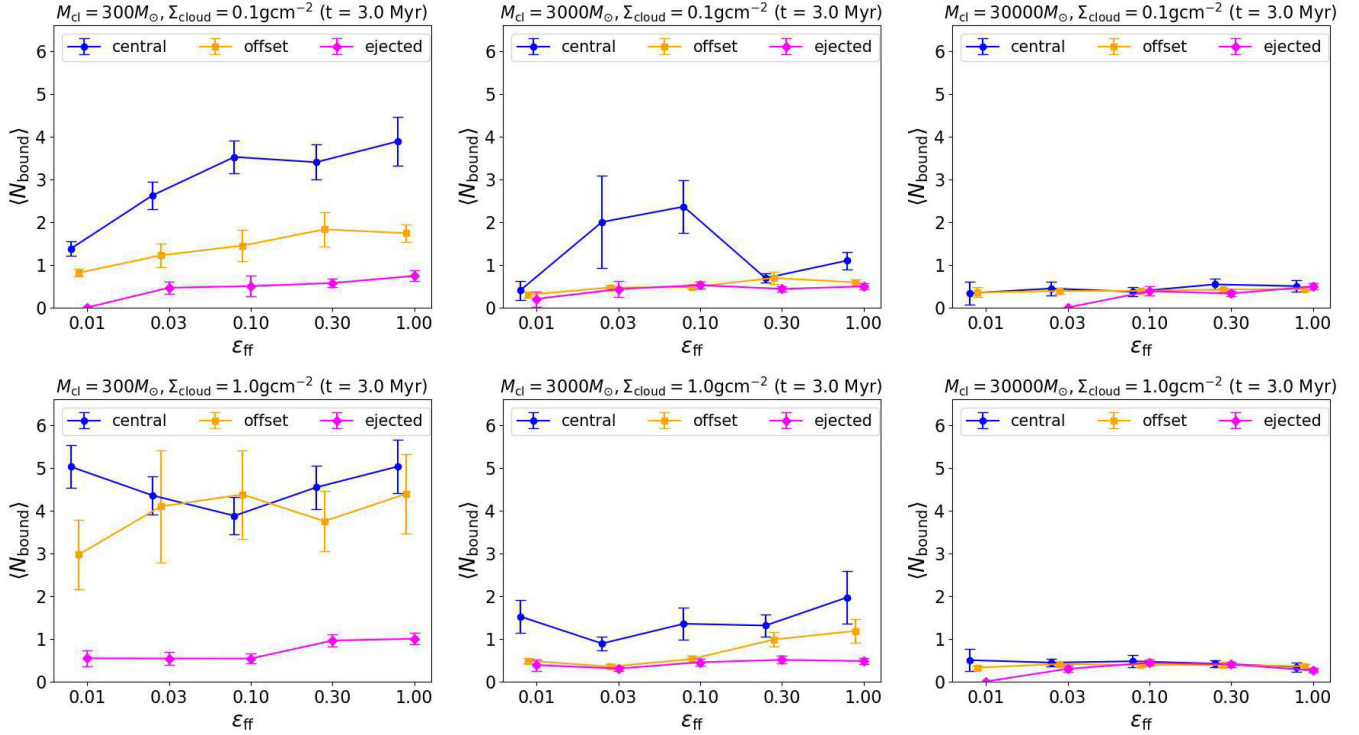


Figure 4. Each panel shows comparison between different ϵ_{ff} models showing $\langle N_{\text{bound}} \rangle$ around massive stars at cluster age of $t = 3.0$ Myr. Panels in the first and second rows show low ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) and high ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) cloud mass surface density cases, respectively. From left to right, the columns show M300, M3000, and M30000 models, respectively. Each panel shows $\langle N_{\text{bound}} \rangle$ around massive stars in three subsets: *central* stars (blue), *offset* stars (orange) and *ejected* stars (magenta). $\langle \dots \rangle$ indicates averaging over massive stars in all realizations of each model. The points for different subsets are offset by [0.1, 0, -0.1] along x-axis to allow for visual comparison. We find that massive stars in different ϵ_{ff} clusters have variations in the number of bound companions at the same cluster age, possibly due to variations in the star cluster's dynamical history.

3.3 Evolution of projected stellar density profiles

In Fig. 8, we show the averaged projected stellar density profiles around *central*, *offset* and *ejected* massive stars. We first constructed projected stellar density profiles ($N_*(r)$) around every primary massive star as described in Section 2.3. Then, we averaged the individual projected stellar density profiles over all massive stars in each subset to generate an averaged projected stellar density profile shown in the figure. We find that in all fiducial models, there is a background density of stars, the blue profile, which includes all stars surrounding the massive star. The blue profile remains mostly flat at large r however, near the massive star, there is a gradual rise in N_* in the innermost bins ($r < 0.01 \text{ pc}$). This is due to the presence of bound companions near the massive star, indicated by the brown profile. However, we find that the brown profile quickly drops off beyond 0.1 pc, contributing only a very small fraction of the total N_* . We note that the bound profile stretches out to further distances around massive stars in M300 models. This is consistent with the fact that these models have small velocity dispersion and thus massive stars can gather more companions, possibly at larger separations, leading to more extended bound stellar systems. Finally, we also note that the background projected stellar density (blue profile) around *ejected* stars is lower in comparison to the *central* and *offset* stars, as expected from their location in the cluster. The bound companions (if any) around these *ejected* massive stars show up as the rise in N_* in the innermost bin.

To compare our projected stellar density profiles with observations in Section 3.4.1, we considered *central* massive stars in the mass range from 10 to 20 $M_\odot$. Then, we constructed projected stellar density profiles around these stars at different cluster ages to study their time evolution as the cluster is formed. Figs 9 and 10 show time evolution of $\langle N_* \rangle$ profiles for different M_{cl} models (rows) in low and high Σ_{cloud} cases, respectively. These profiles include all bound and unbound companions around the massive stars. To reproduce the effect of observational detection limits, we vary the companion mass threshold between $M_{\text{comp}} \geq 0.01 M_\odot$, $M_{\text{comp}} \geq 0.1 M_\odot$, and $M_{\text{comp}} \geq 1.0 M_\odot$ across the columns. In the rows, we compare different M_{cl} models for the fiducial $\epsilon_{\text{ff}} = 0.03$ case.

Let us consider the first column in each figure. We find that across each model, the $\langle N_* \rangle$ profile around $t = 0.5$ Myr starts at a baseline level determined by the mass (hence, number of stars formed) and size of the star cluster, which are set by the initial conditions ($M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}}$). For instance, the M30000 models are more compact and have a higher initial $\langle N_* \rangle$ than M300 or M3000 models. In general, the profile rises with time as more stars are formed and the massive stars, which have already formed, gather companions around them. This continues as long as the cluster is embedded in its gas clump, which is gradually dissipated away. Once the gas is depleted at $t = t_*$, the $\langle N_* \rangle$ profile starts to decline. This is because of the loss of binding potential provided by the gas clump, which causes the star cluster to expand and the projected

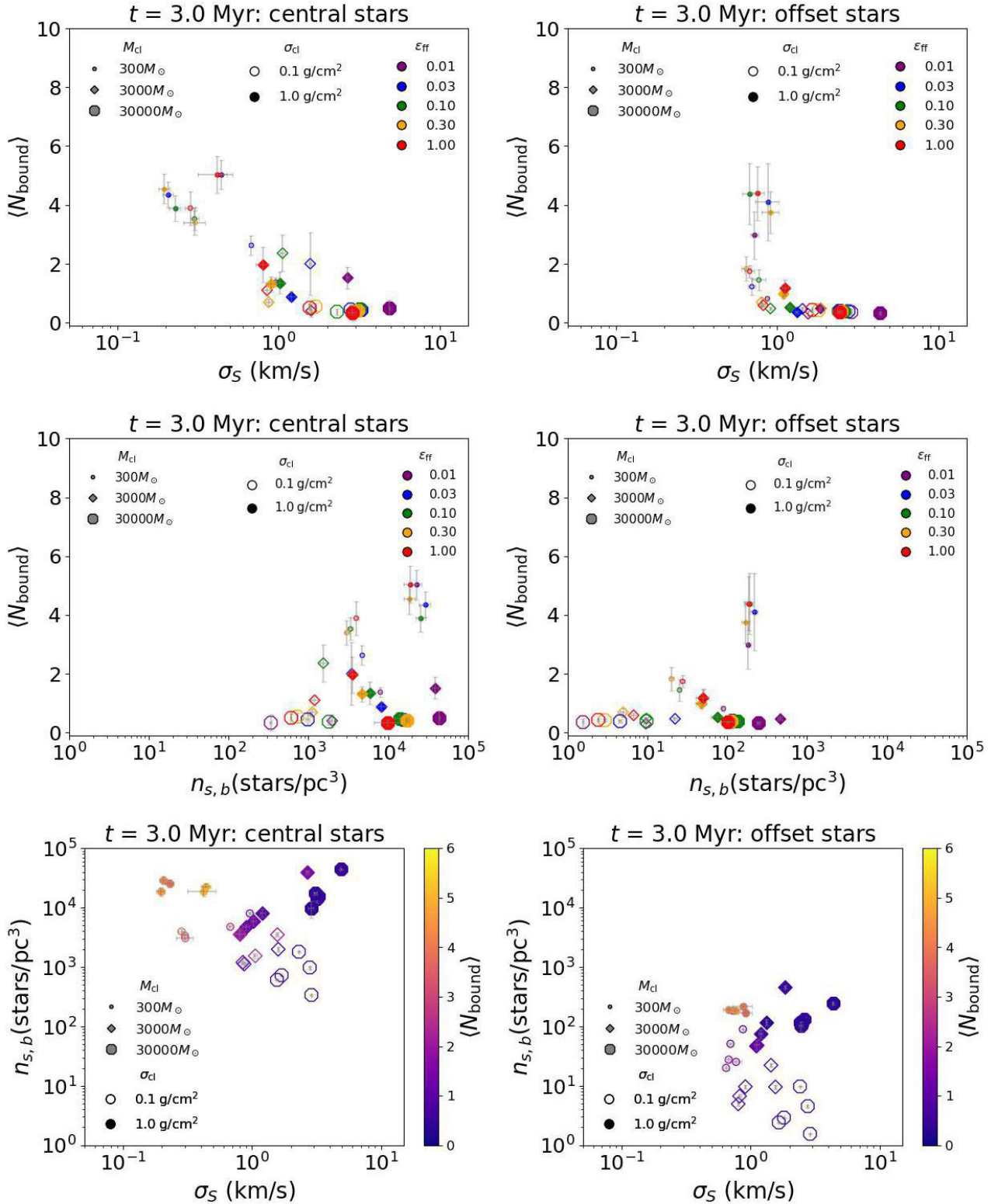


Figure 5. Comparison between different $(M_{\text{cl}}, \Sigma_{\text{cloud}}, \epsilon_{\text{ff}})$ models showing how $\langle N_{\text{bound}} \rangle$ around *central* stars (left column) and *offset* stars (right column) is affected by the stellar velocity dispersion (σ_s) and stellar number density ($n_{s,b}$) within their respective radial regions in the bound star cluster. The first and second rows show $\langle N_{\text{bound}} \rangle$ for each model versus the corresponding σ_s and $n_{s,b}$, respectively, at a cluster age of $t = 3.0$ Myr. The grey lines show the standard errors on the means calculated from the multiple realizations (N_{sims}) of each model. In the third row, a scatter plot of σ_s and $n_{s,b}$ for each $(M_{\text{cl}}, \Sigma_{\text{cloud}})$ model is colour-coded by the value of $\langle N_{\text{bound}} \rangle$ for that model. The colourbar shows the range of $\langle N_{\text{bound}} \rangle$. We find that in star clusters with lower velocity dispersion and higher stellar densities, massive stars will gather more bound companions.

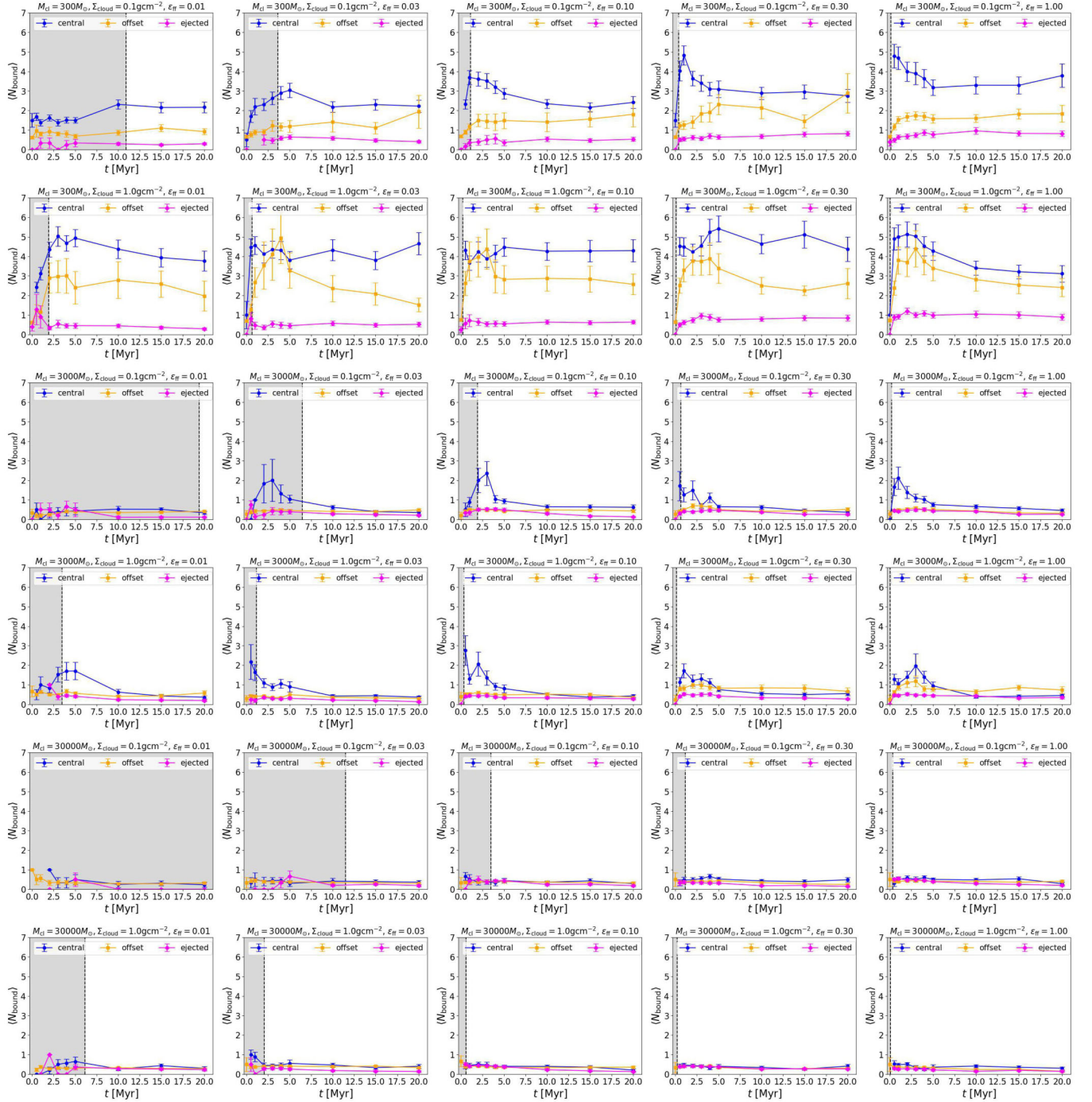


Figure 6. Comparison between different models ($M_{cl}, \Sigma_{cloud}, \epsilon_H$) across our full set showing the evolution of average number of bound companions ($\langle N_{bound} \rangle$) around massive stars versus cluster age t in Myr. The columns show the five different values of ϵ_H . The rows (from top to bottom) show M300L, M300H, M3000L, M3000H, M30000L, and M30000H sets, respectively. Each panel shows $\langle N_{bound} \rangle$ around massive stars in three subsets: *central* stars (blue), *offset* stars (orange) and *ejected* stars (magenta). (...) indicates averaging over massive stars in all realizations of each model. The dashed black vertical line represents the formation time-scale (t_*) of the model. The star cluster is still embedded in background gas before t_* , represented by the shaded region. We find that the number of bound companions around massive stars will be affected by the stellar environment (σ_S and $n_{S,b}$) of their star cluster during its formation, gas expulsion and subsequent expansion.

stellar density around each massive star to fall. The time evolution of the $\langle N_* \rangle$ profile is thus affected by the formation time-scale of the star cluster.

Now, if we consider the same M_{cl} model (row) and compare the columns, we find that on increasing the minimum companion

mass, the $\langle N_* \rangle$ declines because we miss the lower-mass companions around the massive star. The fall in $\langle N_* \rangle$ is around a factor of 10 between $M_{comp} \geq 0.1M_\odot$ and $M_{comp} \geq 1.0M_\odot$, and a factor of 2 between $M_{comp} \geq 0.01M_\odot$ and $M_{comp} \geq 0.1M_\odot$. This indicates that observational studies of stellar clustering around

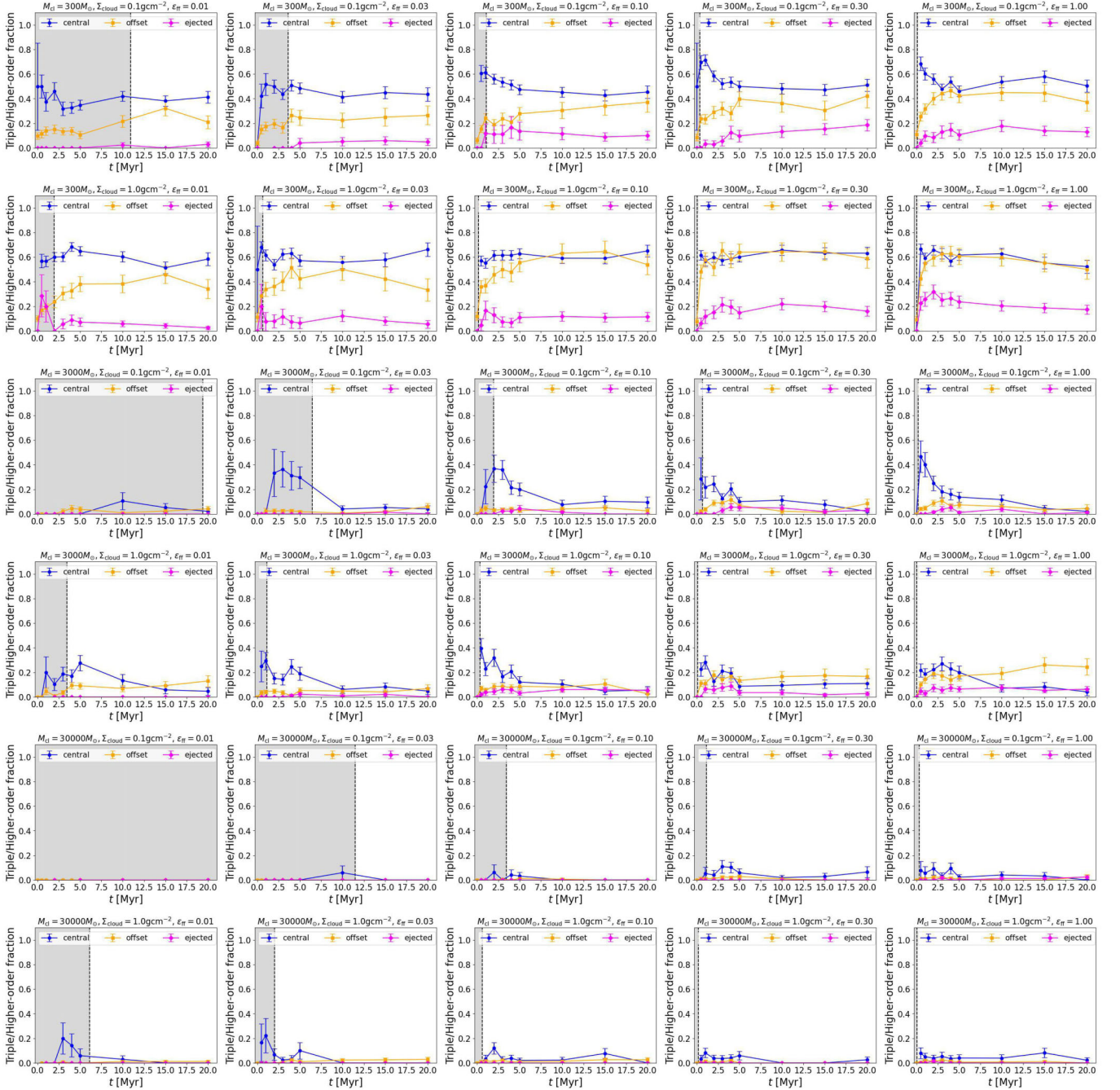


Figure 7. Comparison between different models ($M_{cl}, \Sigma_{cloud}, \epsilon_{ff}$) across our full set showing the evolution of Triple/Higher-order fraction (THF) of massive stars versus cluster age t in Myr. The columns show the five different values of ϵ_{ff} . The rows (from top to bottom) show M300L, M300H, M3000L, M3000H, M30000L, and M30000H sets, respectively. Each panel shows THF of massive stars in three subsets: *central* stars (blue), *offset* stars (orange) and *ejected* stars (magenta). THF was calculated by aggregating massive stars across all realizations of each model and by counting the fraction of massive stars with triple or higher-order bound companions. The dashed black vertical line represents the formation time-scale (t_*) of the model. The star cluster is still embedded in background gas before t_* , represented by the shaded region. We find that bound systems with three or more stars can form dynamically in a forming star cluster. The rise and decline of the fraction of these systems depends on the evolving stellar environment (σ_S and $n_{S,b}$) of the star cluster.

massive stars should be able to detect low-mass companions as small as $0.01 M_\odot$ to properly test models of high-mass star formation. Conversely, any comparison of simulated models with observations should consider both the cluster age and detection limit, considering that $\langle N_* \rangle$ is affected by time evolution and the detected companion masses, respectively.

3.4 Comparison with observations

3.4.1 Comparison of high-mass star formation models

We compare the projected stellar density profiles ($\langle N_* \rangle$) in our simulations with observations in this section. We consider $N_*(r)$ observations around high-mass stars in two star-forming regions:

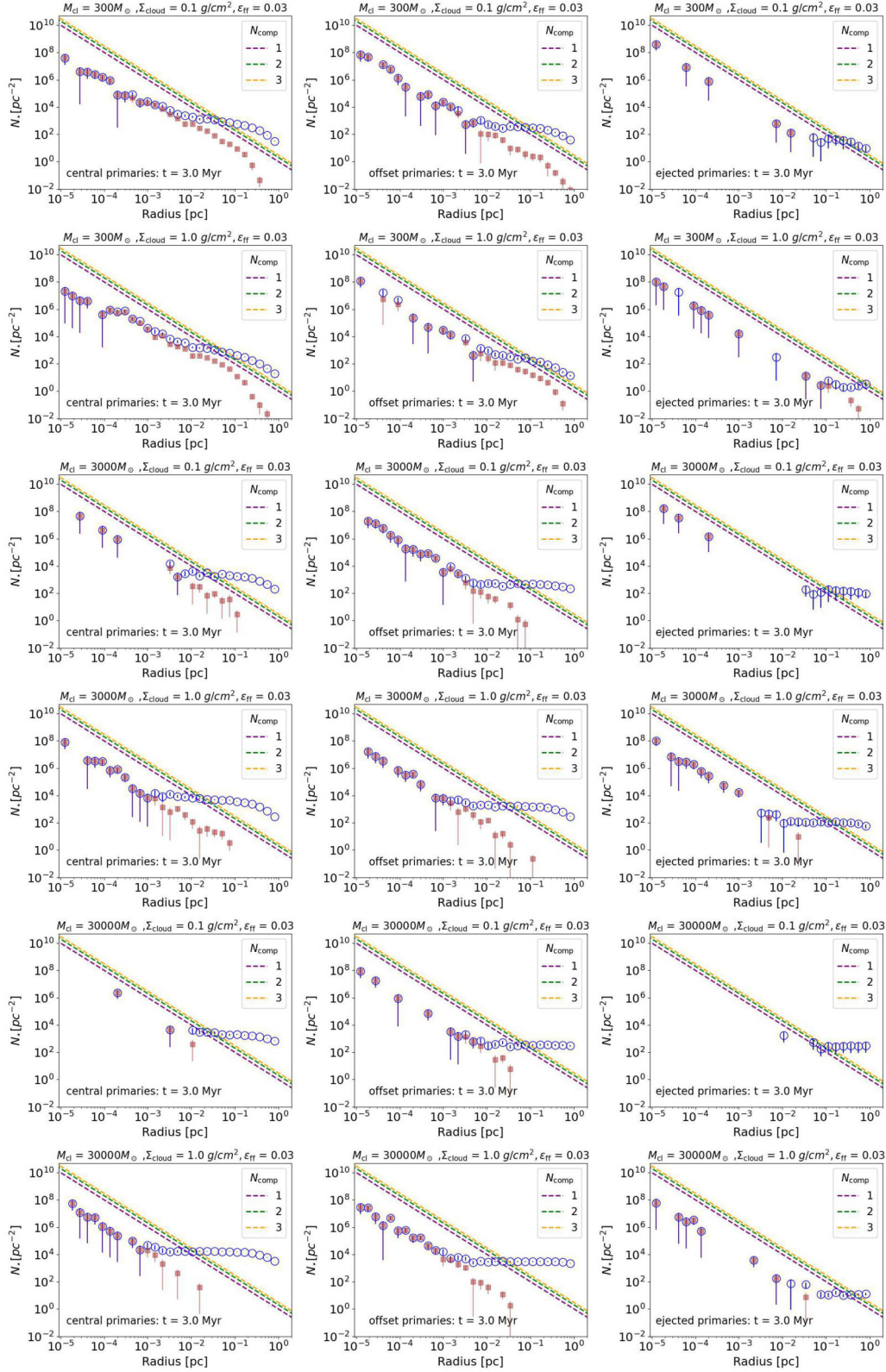


Figure 8. Each panel shows averaged projected stellar density profiles ($\langle N_*(r) \rangle$) around massive stars in the fiducial ($\epsilon_{\text{ff}} = 0.03$) model at cluster age of $t = 3.0$ Myr. The rows show different sets in the following order from top to bottom: M300L, M300H, M3000L, M3000H, M30000L, and M30000H, respectively. The columns (from left to right) show the profiles around *central* stars, *offset* stars, and *ejected* stars, respectively. In each panel, the brown points include only the bound companions, whereas the blue points include all nearby companions. The errorbars are the standard errors on the $\langle N_*(r) \rangle$ in the bin at each radius r . The dashed lines show the expected projected stellar densities if the bin had one, two, or three stars, respectively. The stars are counted in logarithmically spaced annular bins ($N_{\text{bins}} = 30$). We find that the projected stellar density profile of all companions is relatively flat, while the bound companions are found close ($r < 0.1$ pc) to the massive star.

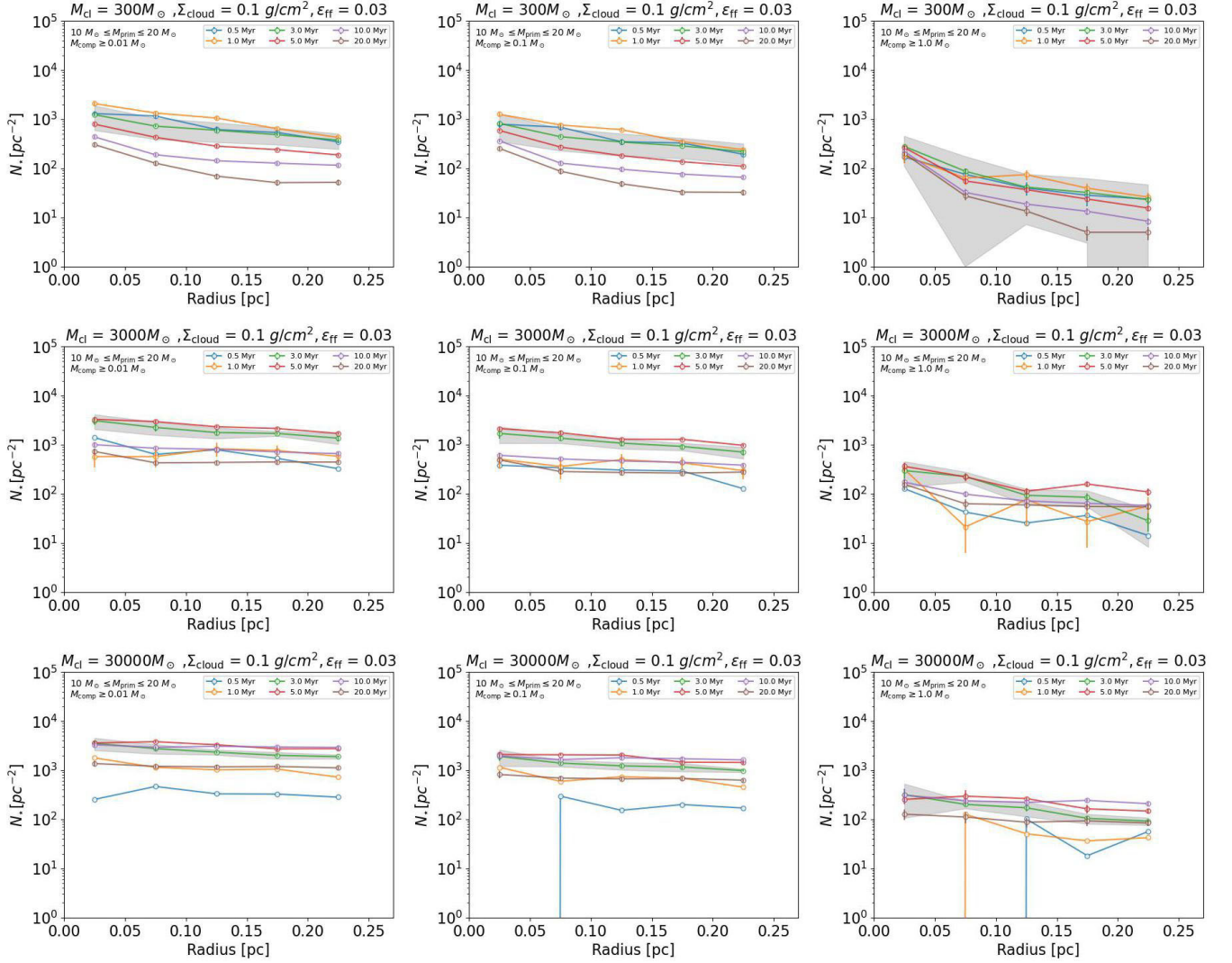


Figure 9. Each panel shows the time evolution of projected stellar density profiles ($\langle N_*(r) \rangle$) around massive stars in simulated clusters for the fiducial ($\epsilon_{\text{ff}} = 0.03$) case in low cloud mass surface density ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) sets. The rows show M300L, M3000L, and M30000L models, respectively. The columns show the results considering companions with masses above 0.01, 0.1, and $1.0 M_{\odot}$, respectively. In every panel, each line represents averaged values of projected stellar density around *central* stars in the $10\text{--}20 M_{\odot}$ mass range at the given cluster age, across multiple realizations for that model. The errorbars represent the standard errors. The shaded region represents the standard deviation for the $t = 3.0$ Myr profile. We find that the projected stellar density profile around a massive star rises during ongoing star formation, reaches a maximum before gas expulsion, and declines as the star cluster expands following gas removal.

a massive protostar in AFGL 5180, studied by S. Crowe et al. (2024), and a high-mass star-forming complex G19.88-0.53, studied by A. R. Costa Silva et al. (2022). For these previous observations, the surface number density was derived by counting the number of YSOs within a given annulus, each annulus of width 0.05 pc. For a fair comparison of simulations with observations, we have counted companions around massive stars in the same linearly spaced bins. Since we are studying the projected number density profiles of stars, we choose the final stellar mass (M_{stars}) as the common variable when comparing the variation of $\langle N_* \rangle$ between simulations and observations during our analysis.

The total gas mass of the AFGL 5180 star-forming clump is estimated from spectral energy distribution (SED) fitting to be around $300\text{--}500 M_{\odot}$ in the literature (A. K. Maity et al. 2023; Z. Telkamp et al. 2025). For our analysis, we assume the gas mass of AFGL 5180 clump to be $M_{\text{cl}} = 400 M_{\odot}$. Since we do not know

the actual final star-formation efficiency (ϵ) of the AFGL 5180 region, we assume a range of $\epsilon = [0.1, 0.3, 0.5, 0.8]$ to estimate the final stellar mass ($M_{\text{stars}} = \epsilon M_{\text{cl}}$) expected to form in the region. The rough lower limit of $M_{\text{stars}} = 40 M_{\odot}$ ($\epsilon = 0.1$) is set because Z. Telkamp et al. (2025) identified four protostars in AFGL 5180 by SED fitting: three in the range $12\text{--}16 M_{\odot}$ and one with $5 M_{\odot}$. The rough upper limit of $M_{\text{stars}} = 320 M_{\odot}$ ($\epsilon = 0.8$) assumes that 20 per cent of gas mass is lost due to effects of stellar feedback. So, the order-of-magnitude estimate of M_{stars} for AFGL 5180 is around $100 M_{\odot}$. Among the four protostars, the best-fit mass of central massive protostar (designated S4) in AFGL 5180 is inferred to be $11.3^{+6.1}_{-4.0} M_{\odot}$ by SED fitting (S. Crowe et al. 2024). The observed $\langle N_* \rangle$ profile was calculated by those authors around this protostar. So, we constructed averaged projected stellar density profiles around massive stars in the $10\text{--}20 M_{\odot}$ mass range across multiple realizations. These observational studies had a detection

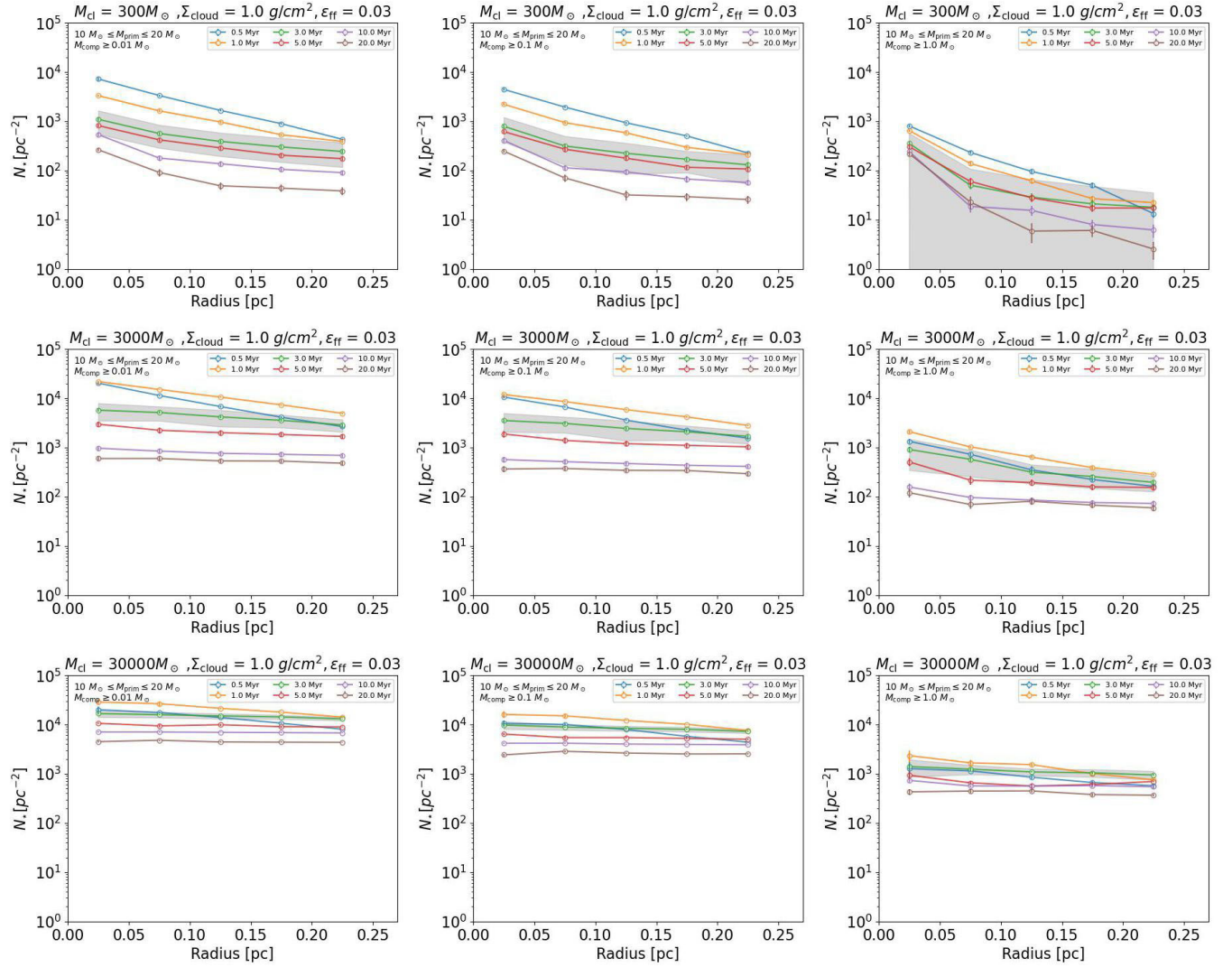


Figure 10. Each panel shows the time evolution of projected stellar density profiles ($\langle N_*(r) \rangle$) around massive stars in simulated clusters for the fiducial ($\epsilon_{\text{ff}} = 0.03$) case in high cloud mass surface density ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) sets. The rows show M300H, M3000H, and M30000H models, respectively. The columns show the results considering companions with masses above 0.01, 0.1, and $1.0 M_{\odot}$, respectively. In every panel, each line represents averaged values of projected stellar density around *central* stars in the $10\text{--}20 M_{\odot}$ mass range at the given cluster age, across multiple realizations for that model. The errorbars represent the standard errors. The shaded region represents the standard deviation for the $t = 3.0$ Myr profile. In addition to the time evolution, observational limits also affect the shape of the projected stellar density profiles.

limit of $0.1 M_{\odot}$. So, we consider profiles generated for the case with $M_{\text{comp}} \geq 0.1 M_{\odot}$. However, for completeness, we also include the cases of $M_{\text{comp}} \geq 0.01 M_{\odot}$ and $M_{\text{comp}} \geq 1.0 M_{\odot}$ in Table 3.

On the other hand, the G19.88-0.53 region is likely a protostellar cluster where multiple massive stars are forming and driving the numerous bipolar outflows observed in the radio and near-infrared (N. Issac et al. 2020; A. R. Costa Silva et al. 2022). Nonetheless, by assuming a single protostellar source dominates the luminosity of the region, the authors performed SED fitting to derive best-fit masses ranging from 4 to $8 M_{\odot}$ for a proto-massive star forming in an initial gas core of mass ranging from 100 to $480 M_{\odot}$, respectively. They also derived a YSO number density of $170 \pm 46.5 \text{ pc}^{-2}$ in the region, which we include in our comparison. For completeness, we note that the above SED fitting is performed with radiative transfer models from Y. Zhang & J. C. Tan (2018), which are based on the TCA model for massive star formation (C. F. McKee & J. C. Tan 2003). The formation of

high-mass stars in our TCCA simulations can be considered to be occurring stochastically via TCA within our clumps.

In the top row of Fig. 11, we present projected stellar density profiles in our M300L and M300H models, which we found to be the best-fit models to the observations, across the parameter space of our models (M_{cl} , Σ_{cloud} , ϵ_{ff}) and cluster age (t). Our TCCA simulations form a star cluster with $M_{\text{stars}} = 0.5 M_{\text{cl}}$, so M300 models have $M_{\text{stars}} = 150 M_{\odot}$. Since we are comparing with young star-forming regions, we examined snapshots at $t = 1.0, 2.0, 3.0, 4.0$, and 5.0 Myr in our simulations. As a result of our multiple realizations for each model, we are able to obtain a reasonably large set of massive stars in the $10\text{--}20 M_{\odot}$ mass range so that we can create averaged profiles at each cluster age.

For comparison to other types of models, we include projected stellar density profiles constructed around massive stars in the STARFORGE simulations from M. Y. Grudić et al. (2022). Their simulations model the collapse of a giant molecular cloud with

Table 3. Calculated values of constants A and α for power-law fits of the form $\langle N_\star \rangle(r) = Ar^\alpha$ to projected stellar density profiles from our simulations (TCCA), STARFORGE simulations (SF) and AFGL 5180 observations.

Set	$\langle N_\star \rangle$ profile	$M_{\text{comp}} \geq 0.01M_\odot$		$M_{\text{comp}} \geq 0.1M_\odot$		$M_{\text{comp}} \geq 1.0M_\odot$	
		A	α	A	α	A	α
AFGL 5180		–	–	147.84	–0.690	–	–
STARFORGE (SF)	10 $M_\odot$	27.71	–1.166	19.50	–1.228	1.59	–1.531
	15 $M_\odot$	25.40	–1.214	18.30	–1.269	1.24	–1.680
	20 $M_\odot$	22.20	–1.256	15.74	–1.321	0.84	–1.815
TCCA M300L	1.0 Myr	170.75	–0.727	98.79	–0.729	8.59	–0.828
	2.0 Myr	240.92	–0.617	134.09	–0.638	10.24	–0.907
	3.0 Myr	182.02	–0.531	98.13	–0.581	4.18	–1.141
	4.0 Myr	93.08	–0.655	45.05	–0.771	2.75	–1.241
	5.0 Myr	74.23	–0.651	36.23	–0.763	2.44	–1.262
TCCA M300H	1.0 Myr	102.01	–0.971	49.14	–1.063	2.16	–1.552
	2.0 Myr	111.47	–0.738	56.69	–0.814	2.83	–1.341
	3.0 Myr	92.40	–0.679	40.70	–0.804	1.57	–1.448
	4.0 Myr	63.19	–0.763	27.21	–0.906	1.36	–1.480
	5.0 Myr	61.91	–0.710	30.30	–0.822	1.70	–1.399
TCCA M3000L	1.0 Myr	686.54	0.035	274.40	–0.171	0.12	–2.133
	2.0 Myr	no fit	no fit	no fit	no fit	no fit	no fit
	3.0 Myr	901.88	–0.342	446.85	–0.398	12.74	–0.952
	4.0 Myr	996.12	–0.335	658.79	–0.248	74.39	–0.286
	5.0 Myr	1221.89	–0.283	652.82	–0.331	47.76	–0.530
TCCA M3000H	1.0 Myr	2058.91	–0.699	1084.18	–0.720	76.96	–0.943
	2.0 Myr	1876.13	–0.470	1015.06	–0.501	82.34	–0.755
	3.0 Myr	1990.77	–0.313	1107.73	–0.339	70.46	–0.721
	4.0 Myr	1608.57	–0.288	903.71	–0.298	77.88	–0.549
	5.0 Myr	1166.46	–0.255	683.29	–0.275	65.30	–0.528
TCCA M30000L	1.0 Myr	no fit	no fit	no fit	no fit	no fit	no fit
	2.0 Myr	694.05	–0.337	443.44	–0.300	67.01	–0.159
	3.0 Myr	1209.10	–0.304	622.00	–0.315	34.48	–0.668
	4.0 Myr	2016.21	–0.120	1024.41	–0.175	108.40	–0.225
	5.0 Myr	2303.86	–0.147	1131.54	–0.208	117.05	–0.293
TCCA M30000H	1.0 Myr	8769.36	–0.401	3833.51	–0.542	334.76	–0.621
	2.0 Myr	13799.42	–0.186	7398.29	–0.209	847.74	–0.241
	3.0 Myr	11282.22	–0.119	6071.45	–0.141	736.41	–0.190
	4.0 Myr	10258.03	–0.124	5618.23	–0.137	636.06	–0.199
	5.0 Myr	8009.56	–0.076	4327.21	–0.104	702.76	0.021

an initial mass of 20 000 $M_\odot$ and size of 10 pc into a young stellar cluster with the formation of protostars as sink particles and incorporating their feedback when they become stars. At $\epsilon = 0.02$, this STARFORGE simulation forms a cluster with $M_{\text{stars}} = 400 M_\odot$ (M. Y. Grudić et al. 2022; S. Crowe et al. 2024), which has the closest resemblance to both AFGL 5180 ($\sim 100 M_\odot$) and M300 models of TCCA (150 $M_\odot$) among the publicly available STARFORGE simulation runs. As mentioned above, we assume M_{stars} as the common variable when comparing models and observations. The formation of massive stars in STARFORGE simulations resembles the competitive accretion model of high-mass star formation. In the case of STARFORGE, massive stars build up their mass over time by gas accretion while competing against neighbouring low-mass companions. So, STARFORGE simulations had massive stars in the required mass range of 10–20 $M_\odot$ only from $t = 3.0$ Myr onwards. At early cluster ages ($3.0 \text{ Myr} \leq t \leq 6.0 \text{ Myr}$), only one or two protostars reach this mass threshold. If we generated profiles at these ages, they would be highly subject to random scatter because we do not have a significant number of massive stars in the required mass range. This is further complicated by the fact that the STARFORGE model has only one realization, so we can only analyse the set of 28 protostars that reach this mass range at some point during this simulation. Therefore, we chose to analyse projected number density profiles when these massive

stars reached a certain mass instead. Then, we average the profiles around the full set of 28 massive stars that reach the same 10–20 $M_\odot$ mass range during their formation. In Fig. 11, we present the averaged projected stellar density profiles expected when a massive star reaches masses of 10, 15, and 20 $M_\odot$ in STARFORGE simulations.

To quantify the shape of the $\langle N_\star \rangle$ profiles, we fit power-laws of the form: $\langle N_\star \rangle(r) = Ar^\alpha$. The calculated values of A and α for each profile are presented in Table 3. The columns for the $M_{\text{comp}} \geq 0.1M_\odot$ case represent the profiles presented in the middle row of Fig. 11. However, we also explore the effect of observational limits on M_{comp} on the shape of $\langle N_\star \rangle$ profiles for the other two cases ($M_{\text{comp}} \geq 0.01M_\odot$ and $M_{\text{comp}} \geq 1.0M_\odot$), and present their fits in Table 3. While the first two rows of Fig. 11 focus on M300 models, we also present fits for M3000 and M30000 models in Table 3.

From Fig. 11, we find that there is a wide scatter in both our and STARFORGE projected stellar density profiles (indicated by the shaded regions). However, profiles from STARFORGE models are generally steeper than our profiles, and fall quickly beyond $r > 0.1$ pc. In contrast, our projected stellar density profiles remain relatively flat out to 0.25 pc, which is also seen in Figs 9 and 10. The STARFORGE profiles for the 10, 15, and 20 $M_\odot$ cases are quite similar even though they get slightly steeper as the massive star accretes mass to rise from 10 $M_\odot$ to 20 $M_\odot$ (see power-law

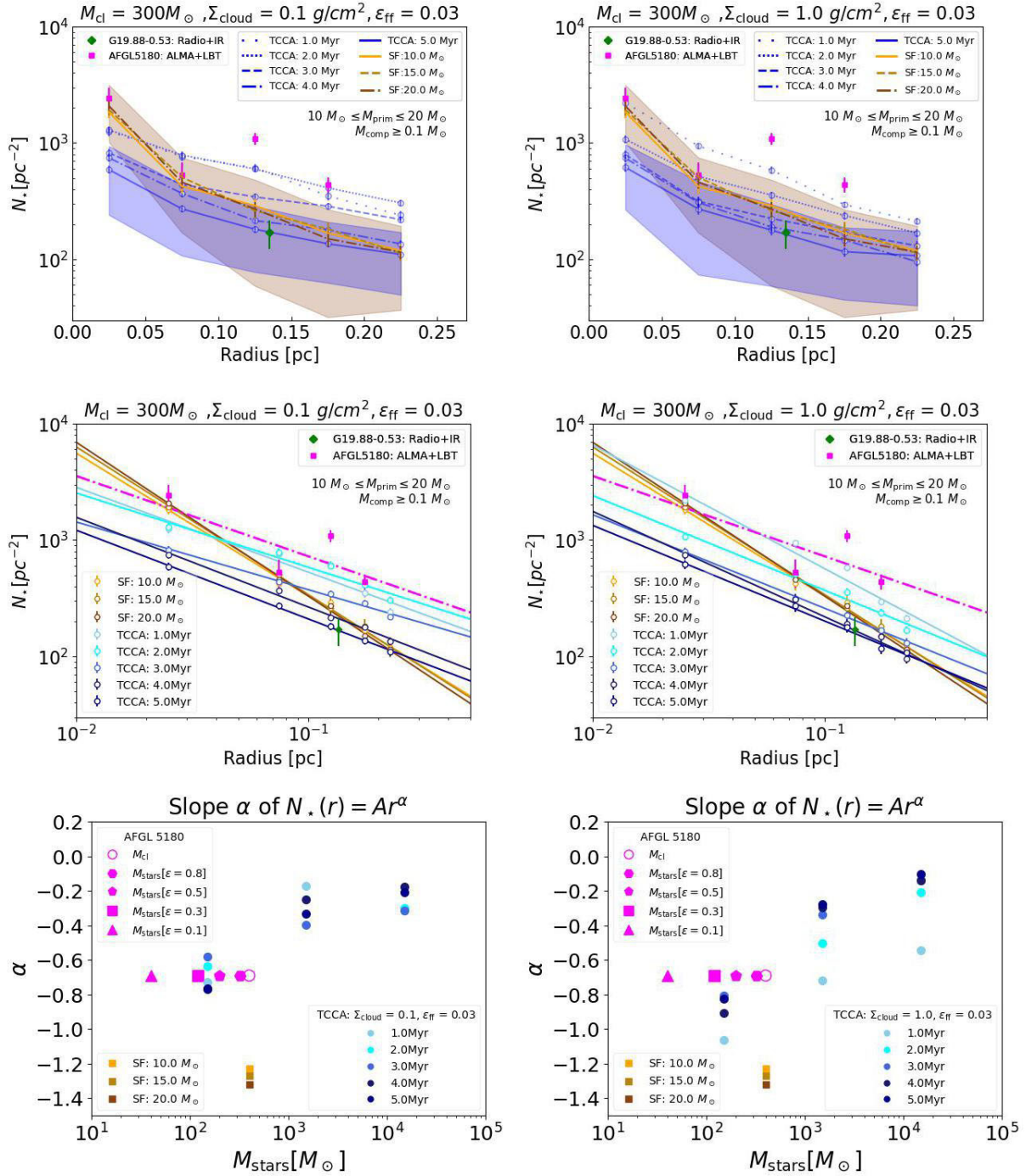


Figure 11. *Top row:* Projected stellar density profiles for our fiducial ($\epsilon_{\text{ff}} = 0.03$) M300L and M300H models (TCCA models, shown in blue), which were found to be the best match to the observations from AFGL 5180 (magenta squares) and G19.88-0.53 (green diamond). Blue lines with different line styles represent TCCA profiles at different cluster ages. The errorbars represent the standard errors. The blue shaded region represents the standard deviation for the $t = 5.0$ Myr profile. These profiles were created around *central* massive stars which fall in the mass range $10\text{--}20 M_{\odot}$ and their companions have masses $M_{\text{comp}} \geq 0.1 M_{\odot}$, matching the observational limits. We also compare with simulated profiles averaged around massive stars from STARFORGE simulations in Grudic et al. 2022. Orange lines of different shades represent the profiles when the massive star reaches different masses. The orange shaded region represents the standard deviation for the case of $20 M_{\odot}$ profile. *Middle row:* Power-law fits of the form $\langle N_{\star} \rangle(r) = Ar^{\alpha}$ to the projected stellar density profiles presented in the top row. Blue lines of different shades represent fits for our TCCA models at different cluster ages. Orange lines of different shades represent fits for STARFORGE models when the massive star reaches different masses. The magenta line represents the fit to AFGL 5180 observations. The constants A and α for these power-law fits are provided in Table 3. *Bottom row:* Index (α) of the power-law fits presented in the middle row versus the final stellar mass (M_{stars}) formed in each model. TCCA models form $0.5 M_{\text{cl}}$ in stellar mass and this particular STARFORGE model also forms $400 M_{\odot}$ of stars. The AFGL 5180 clump has a mass of $M_{\text{cl}} = 400 M_{\odot}$: we estimate M_{stars} for AFGL 5180 assuming different ϵ , where $M_{\text{stars}} = \epsilon M_{\text{cl}}$. We find that the observed AFGL 5180 projected stellar density profile has a flat shape consistent with our TCCA models.

indices (α) in Columns 5 and 6 of Table 3). We also note the time evolution of our models in Fig. 11. The profile rises in the first Myr as new stars form in the vicinity of the massive star. The star formation is complete by $t_* = 3.64$ Myr and 0.65 Myr in the M300L and M300H models, respectively (see Table 1). Then, the star cluster expands following gas removal. Thus, $\langle N_* \rangle$ profile decreases with time. If we consider all low-mass companions ($M_{\text{comp}} > 0.01 M_\odot$), we find that our profiles at $t = 5.0$ Myr are in general flatter than corresponding early profiles at $t = 1.0$ Myr. We also note the effect of observational limits (M_{comp}) on the shape of profiles. We find that on increasing M_{comp} from $0.01 M_\odot$ to $1.0 M_\odot$, the profile steepens for STARFORGE profiles. This is generally true for our models as well [refer to power-law indices (α) in Table 3].

Now, we turn to the comparison of simulated profiles with observations. In the two innermost bins ($r < 0.1$ pc), we find that STARFORGE profile is closer to the AFGL 5180 data points. While our profiles are generally lower than AFGL 5180 in the innermost bin, our M300H profile reaches similar values around an early cluster age of $t = 1.0$ Myr. In addition, the second AFGL 5180 data point is reached by our M300L and M300H models between 1.0 and 3.0 Myr. Moreover, we find that the AFGL 5180 profile remains relatively flat, with our models being a better match in the power-law index (α ; see Table 3). The STARFORGE profiles are considerably steeper so the outer AFGL 5180 data points ($r > 0.1$ pc) are much higher than the corresponding STARFORGE point. The closest overall match of our M300 models with the AFGL 5180 observations occurs at a cluster age between 1.0 and 2.0 Myr. This is in good agreement with the reported age between 1.0 and 3.0 Myr for the AFGL 5180 region (K. E. Devine et al. 2008; T. Vasyunina 2010; A. K. Maity et al. 2023). Even though we cannot constrain the shape of the profile for G19.88-0.53 with only one point, we note that their measurement aligns well with both our and STARFORGE models within the range of one standard deviation.

Finally, the third row of Fig. 11 helps us quantify how the power-law index (α) of the $\langle N_* \rangle$ profile changes with the final stellar mass (M_{stars}) of the model. We find that the index becomes less negative with rising M_{stars} for our TCCA models, in other words, the profiles are flatter for M3000 and M30000 models than M300 models. The AFGL 5180 profile has a flatter power-law index ($\alpha = -0.690$), which is closer to TCCA models ($\alpha \sim -0.7$ to -0.8), as observed in the first two rows of Fig. 11 and Table 3. The STARFORGE profiles with indices $\alpha \sim -1.3$ are much steeper than both AFGL 5180 and TCCA models.

3.4.2 Comparison of companion properties

Fig. 12 shows the separation and mass-ratio of bound companions around all massive stars with masses $\geq 16.0 M_\odot$ in our simulations. We compare them with observed companions around O-stars from the Southern Massive Stars at High Angular Resolution survey (SMASH+; F. Tramper et al. 2025). This high angular resolution interferometric survey targeted nearby O-stars in the Southern hemisphere ($\delta < 0^\circ$) and identified companions around them. In addition to the comparison of full set of companions, we compare the different orders of companions separately, focusing on the binary and tertiary companions. We found M300L and M300H models to be the most comparable to their observations, because massive stars in higher M_{cl} models have fewer companions (mostly singles or binaries, hence the smaller $\langle N_{\text{bound}} \rangle$).

The survey had reported a relative scarcity of massive companions at large separations beyond 100 au, with few companions above $q > 0.8$ at such distances, which were mostly upper limits (F. Tramper et al. 2025). We observe a similar trend in our simulated companions in the M300 models, with large-distance companions having small mass ratios ($q < 0.4$) beyond 1000 au. On the other hand, the SMASH+ survey found close binary and tertiary companions within 10 au of some of these O-stars. In our simulations, we reproduce close binaries however, we lack close triple companions. Our closest triple companions are at separations of 100s of au. We discuss the implications of this lack of close triple companions in Section 4.3.

For M300L and M300H models, we ran a two-sample Kolmogorov–Smirnov (K-S) test to check the similarity of the simulated and observed companion distributions for both separations and mass ratios. The results of the K-S test are presented in Table 4. We found that both M300L and M300H models were similar to the observed triple separations while only M300L was similar to the observed binary separations. For mass ratios, however, we did not find similarity between simulated and observed companions in these models.

4 DISCUSSION

In Section 4.1, we discuss the effect of the star cluster environment on the bound systems around massive stars. In Section 4.2, we discuss how projected stellar density profiles reflect the clustering of companions predicted by different massive star formation models. In Section 4.3, we discuss the comparison of our bound companions with observed companions around O-stars. In Section 4.4, we mention the caveats of our modelling and the scope for future improvements.

4.1 Massive stars in the environment of a star cluster

We found that stars in the central regions of the star cluster have more bound companions than stars of similar mass in the periphery or those that had been ejected from the star cluster. This indicates that, in addition to its mass, the location of a massive star also determines the extent of bound systems that can form around it. This is related to the prevailing number density of stars that could be captured by the massive star. The crowded environment in the centre of the star cluster enables massive stars to gather more companions during their formation and/or migration during dynamical mass-segregation. However, this is also true for encounters that could disrupt the bound system. We found that this process also depends on the velocity dispersion of the bound cluster. In particular, we found that massive stars in central regions of star clusters with lower values of stellar velocity dispersion gathered more companions than those in clusters with a higher stellar velocity dispersion. This is because the lower velocity dispersion enhances the gravitational focusing experienced by a potential companion star during the encounter with the massive star and its bound multiple system. This means that captures are favoured over disruptions in lower velocity dispersion environments. In higher velocity dispersion environments, the captures become harder because encounters are more energetic on average. Therefore, the extent of bound systems is determined by the stellar velocity dispersion and number density of the star cluster environment.

The time evolution of stellar velocity dispersion and number density of systems in the bound cluster means that the bound

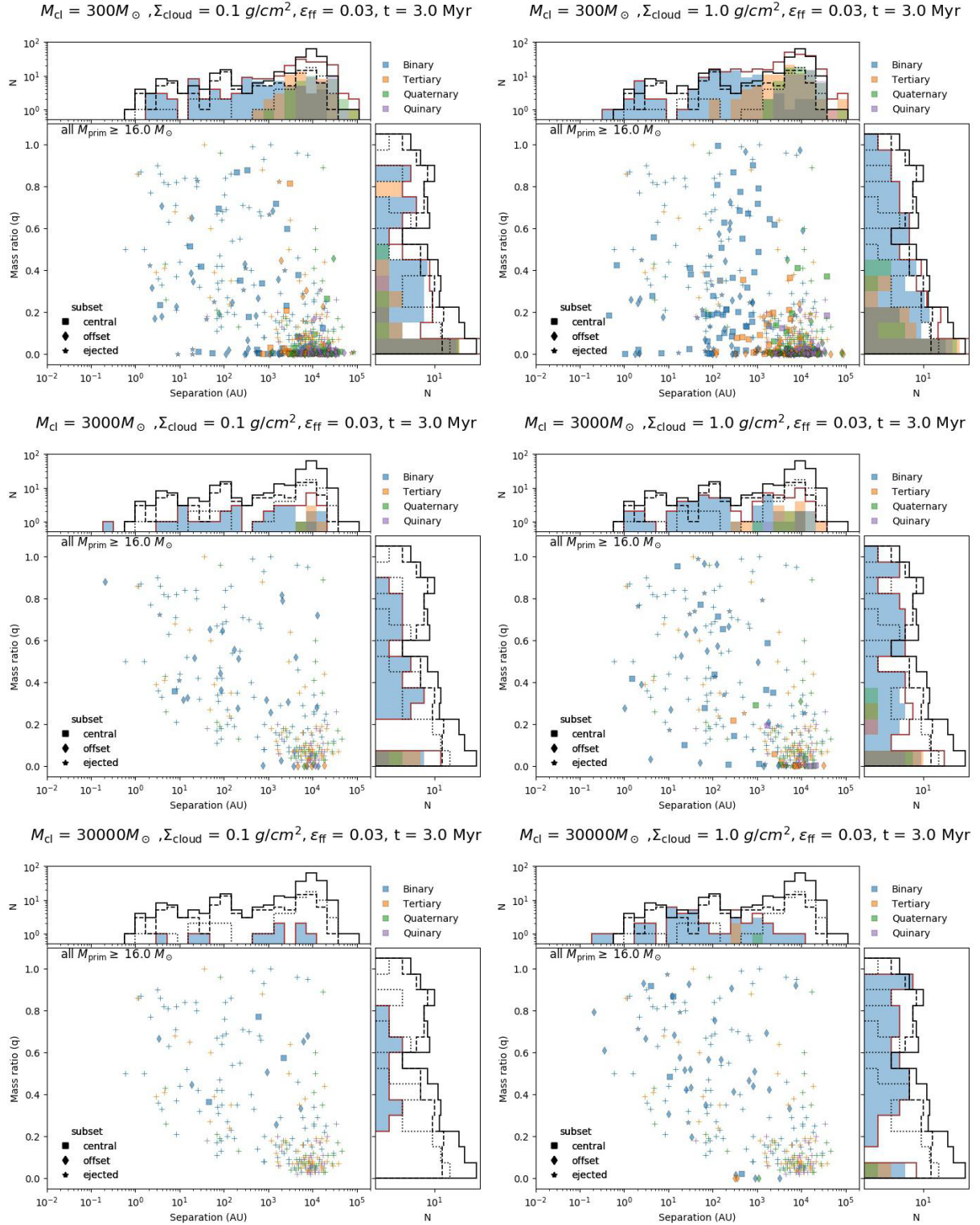


Figure 12. Each panel shows the distribution of the separation and mass ratios (with respect to primary mass) of bound companions around all massive stars with masses $\geq 16.0 M_{\odot}$ in our fiducial ($\epsilon_{\text{ff}} = 0.03$) simulations. From top to bottom, the rows show M300, M3000, and M30000 models, respectively. The left and right columns show low ($\Sigma_{\text{cloud}} = 0.1 \text{ g cm}^{-2}$) and high ($\Sigma_{\text{cloud}} = 1.0 \text{ g cm}^{-2}$) cloud mass surface density cases, respectively. In the scatter plot in each panel, the filled symbols represent our simulated companions, whereas the plus symbols denote observed companions around O-stars from the SMASH+ survey. The square, diamond, and star markers represent companions around central, offset, and ejected primaries, respectively, in our simulations. The companions are colour-coded based on their multiplicity order (binary, triple, quaternary and quinary companions). In the histograms for separation (top panel) and mass-ratio (right panel), the black histogram represents the observed companions, while the red histogram represents the simulated companions. We further plot separate colour-coded filled histograms for different orders of simulated companions. For comparison, the binary and tertiary subsets of observed companions are denoted by the dashed and dotted black histograms, respectively. We find that some observed massive stars have close triple companions, which might need to be formed primordially.

Table 4. Results from the two-sample Kolmogorov–Smirnov (K-S) test between the separation and mass-ratio distributions for observed and simulated companions. The null hypothesis (H_0) states that both distributions are drawn from the same population. D is the K-S distance between the two distributions being tested for similarity. Each row shows the simulated distribution, compared with the corresponding distribution from SMASH+ observations.

Set	Property	D	p -value	H_0
M300L	Binary separations	0.183	0.097	True
	Triple separations	0.203	0.190	True
	Binary mass ratios	0.426	<0.001	False
	Triple mass ratios	0.756	<0.001	False
M300H	Binary separations	0.325	<0.001	False
	Triple separations	0.191	0.133	True
	Binary mass ratios	0.267	<0.001	False
	Triple mass ratios	0.738	<0.001	False

systems around massive stars evolve with time. In early stages of formation, when new stars are being formed and number densities are high, the bound systems build up. However, once the star cluster expands after gas removal and the number densities fall, bound systems lose members due to dynamical encounters and/or stripping of weakly bound members by the rest of the star cluster. These stripped companions would become part of the field star population that remains bound to the star cluster.

During our analysis, we identified ejected stars that were unbound from their star cluster. They comprised around 1–30 per cent of the primary massive star population at $t = 3.0$ Myr, depending on the model. These stars could have become unbound during the phase of cluster expansion after loss of gas potential. In addition, some stars might have been ejected through three-body encounters in the crowded region of the cluster or during the chaotic breakup of non-hierarchical multiple systems (A. Poveda, J. Ruiz & C. Allen 1967; J. P. Anosova 1986; B. Reipurth et al. 2010, 2026), in the dynamical ejection scenario (DES). Some stars might have received supernova kicks at the end of stellar evolution, either as single stars or part of a binary system, in the binary supernova (BSS) scenario (A. Blaauw 1961; B. Dinçel et al. 2025). M. Carretero-Castrillo et al. (2025) conducted a study of galactic runaway O-stars studying the signatures of the DES and BSS scenarios on the runaway speed and rotation of such runaway massive stars.

4.2 Projected stellar density profiles and theories of massive star formation

Our TCCA models can be considered to represent a limiting case of the TCA (C. F. McKee & J. C. Tan 2003) model of massive star formation because massive stars are born randomly across the star cluster. They are not necessarily formed in regions of existing high stellar density, in which case they would form with many low-mass companions also forming around them and competing for gas accretion, as in models of competitive accretion (I. A. Bonnell et al. 2001; M. Y. Grudić et al. 2022). We found that the TCCA model, which develops enhanced stellar multiplicity and clustering around massive stars dynamically after individual stars have formed, and competitive accretion models that have these features primordially as a crucial part of the massive star formation process, predict different shapes for projected stellar density profiles around massive stars. The competitive accretion models have more substructure due to the close low-mass com-

panions. Hence, their projected stellar density profiles are steeper in comparison with the TCCA models, where star formation is more homogeneous.

In addition, we noted the time evolution of projected stellar density profiles around massive stars in young forming star clusters. In TCCA, the projected stellar density profile builds up quickly over time as more stars are formed around the massive star. However, once the star cluster is formed and gas is dissipated away, the projected stellar density falls as the cluster expands. In star clusters that form quicker (for instance, higher ϵ_{ff} models), the projected stellar density profiles can build up and then fall to a baseline level quickly within the first few Myr. This is true for competitive accretion as well, however the massive star takes time to grow in mass and gather enough companions, so the projected stellar density profile is still rising in the first few Myr in those models.

We also compared power-law fits between simulated and observed profiles from AFGL 5180 and G19.88-0.53 star-forming regions. We found that G19.88-0.53 was in agreement with both TCCA and competitive accretion models. For AFGL 5180, we found that the competitive accretion models fitted the observed profile in the two innermost bins ($r < 0.1$ pc), however they were much steeper and fell lower at larger distances. In contrast, the TCCA models were relatively flat with the power-law indices being closer to those of the AFGL 5180 profile. The closest fit of the TCCA models (M300L and M300H) to AFGL 5180 observations occurred at $t \lesssim 3.0$ Myr.

However, we have only compared with two high-mass star-forming regions here; observations of more such regions will be necessary to generalize our findings. In addition, the current observations do not measure the profile at smaller scales ($r < 0.01$ pc) due to resolution limits. Our work aimed to test whether, given an observed companion distribution, we could distinguish profiles resulting from different high-mass star formation scenarios. Higher-resolution observations, which can probe the small-scale distribution of companions, will be necessary to confirm the dominant mechanism. It is possible that different scenarios could contribute to the high-mass stellar population in different parts of the profile. If we find that both mechanisms operate together in the same star-forming region, this may indicate a scale-dependent behaviour: small-scale clustering around massive stars could be shaped by localized fragmentation and accretion processes, while the larger-scale radial structure reflects the overall mass distribution. Alternatively, the flatter observed slopes may result from partial dynamical mixing that smoothens any primordial substructure. To explore this, simulations of TCCA and competitive accretion tailored for those specific star-forming regions would help to fairly ascertain the relative contributions of these mechanisms to the formation of the overall massive star population. Ideally, these simulations will have enough independent realizations to provide a statistically large sample of massive stars at cluster ages comparable to young star-forming regions.

4.3 Companion properties

We have 50 per cent primordial binaries, but no primordial triples and/or multiple systems. However, we found that massive stars gathered many lower-mass companions via dynamical interactions after they were formed in the simulation. We compared the projected separations and mass ratios of companions around massive stars above $16 M_{\odot}$ in our simulations with observed

companions around O-stars in the SMASH+ survey. We found a similar drop-off in equal-mass companions beyond a certain distance. They noted a decline after 100 au, whereas we found a decline after 1000 au. The fact that we find massive companions further out is probably because our massive stars are part of a forming star cluster with higher stellar densities around them, whereas SMASH+ survey includes field stars. So, some dynamical processing may have stripped the companions beyond 100 au around these stars after star cluster formation and expansion was complete.

In addition, we find close binary companions to massive stars within 10 au of our massive stars, some of which have formed dynamically. S. Chon & A. Vigna-Gómez (2026) used their star cluster simulations to note that such close binary systems can form through the hardening of the binary due to circumbinary disc-driven migration. On the other hand, we lack close triple companions in our simulations. We infer that these close triple companions around observed O-stars must have been formed primordially. These primordial companions could be formed via fragmentation in a core or disc, with the remnant gas possibly driving migration to tighten the orbits of the bound triple system.

4.4 Model caveats and future work

We have studied the clustering of companions around massive stars in star clusters by post-processing simulations of gradual star cluster formation. These are pure N -body simulations with a semi-analytical gas potential. This was necessary because our focus was on studying N -body dynamics of stars using multiple realizations for star clusters that were forming at different time-scales. Pure hydrodynamical simulations would be too computationally expensive for this purpose and miss the detailed N -body dynamics. Recent works using the TORCH framework have made progress in this regard (J. E. Wall et al. 2019; C. Cournoyer-Cloutier et al. 2024), even though these coupled N -body and hydro simulations are still computationally expensive in general. Nonetheless, capturing the dynamical formation of multiple systems in an evolving gas potential of a forming star cluster will need self-consistent modelling of stellar dynamics, gas hydrodynamics, and stellar feedback, which we aim to include in future work.

Such self-consistent simulations are needed to study the survival of young star clusters as they emerge from their parent gas clouds. During this period, these young clusters can dissolve rapidly due to the change in gravitational potential caused by gas expulsion through effects of stellar feedback (e.g. C. J. Lada & E. A. Lada 2003). However, they can also be dissolved by the disruptive effects of continuous tidal shocks generated during interactions of the young star cluster with the surrounding dense gaseous environment (‘cruel cradle effect’; B. G. Elmegreen & D. A. Hunter 2010; J. M. D. Kruijssen et al. 2011, 2012). In addition, star cluster formation could occur through the merging of fractal substructures formed during the turbulent fragmentation of a single parent molecular cloud (I. A. Bonnell, M. R. Bate & S. G. Vine 2003; J. Karam & A. Sills 2022), or those formed during cloud-cloud collisions (B. Wu et al. 2017). These simulations will allow us to study how multiple systems inside such rapidly evolving star clusters react to the changing stellar number density and velocity dispersion of those environments.

Although our simulations did not include primordial mass segregation, some level of mass segregation is developed dynamically during the formation phase, especially in lower-mass clus-

ters (see J. P. Farias & J. C. Tan 2023). We found that this migration of massive stars to regions of higher number densities in the centre contributed partially to the formation of multiple systems. Therefore, if massive stars do form preferentially in the crowded regions (e.g. I. A. Bonnell et al. 2001), it is possible that the number of multiples forming in such environments is higher than what we report here under the same global conditions. Finally, we did not have any primordial triples and/or multiple systems in our simulations. However, observations indicate a population of close triple companions besides close binaries. We aim to include these primordial triples in our future simulations. This will allow us to explore the evolution of young star clusters with simultaneous presence of mass segregation and primordial higher-order multiplicity, similar to recent work by V. Amiri et al. (2026).

5 CONCLUSIONS

In this paper, we studied the clustering of bound and unbound companions around massive stars inside gradually forming star clusters via TCCA by post-processing existing N -body simulations. In these simulations, star clusters are formed in different formation time-scales (t_*) parametrized by star formation efficiency per free-fall time (ϵ_{ff}), mass of star-forming clump (M_{cl}) and mass surface densities (Σ_{cloud}) of the parent cloud.

We found that massive stars in the centre of bound clusters gather more companions than massive stars that have been ejected from the cluster. We found that ejected stars are mostly single or in binaries. The number of companions is greater for massive stars in star clusters with lower stellar velocity dispersion and higher stellar densities. When we compared the clustering of companions around our massive stars with projected stellar density profiles, we found a good match to YSO profiles in the AFGL 5180 and G19.88-0.53 star-forming regions. Finally, we noted a lack of close triple companions around our simulated O-stars when compared to companions around O-stars in the SMASH+ survey. Since these companions were not captured dynamically, they are likely formed via primordial routes like core or disc fragmentation.

In the future, we aim to extend our gradual star cluster formation models with hydrodynamical coupling and an accurate treatment of primordial triples. We are also currently working on extending the analysis of dynamically gathered companions around primary stars across the full stellar mass spectrum.

ACKNOWLEDGEMENTS

We thank the anonymous reviewer for their comments and constructive feedback that greatly improved the manuscript. We thank Mike Grudić and Hugues Sana for helpful discussions. AG acknowledges invaluable support from the Chalmers Astrophysics and Space Sciences Summer (CASSUM) research program. JCT acknowledges support from the European Research Council (ERC) Advanced Grant 788829: Massive Star Formation through the Universe (MSTAR), National Science Foundation (NSF) grant AST-2009674, and funding from the Virginia Institute for Theoretical Astrophysics (VITA), supported by the College and Graduate School of Arts and Sciences at the University of Virginia. JPF acknowledges support from the National Aeronautics and Space Administration (NASA) grant 80NSSC20K0507 and NSF Career award 1748571. The simulations presented in this work were performed on the Origins computing cluster at Chalmers. The authors acknowl-

edge Research Computing at the University of Virginia for providing computational resources and technical support that have contributed to the results reported within this publication (<https://rc.virginia.edu>). This research made use of NASA's Astrophysics Data System.

DATA AVAILABILITY

The data underlying this article will be shared upon reasonable request to the corresponding author.

REFERENCES

- Alcalá J. M. et al., 2014, *A&A*, 561, A2
- Allison R. J., Goodwin S. P., Parker R. J., de Grijs R., Portegies Zwart S. F., Kouwenhoven M. B. N., 2009, *ApJ*, 700, L99
- Amiri V., Flammini Dotti F., Pang X., Kamlah A. W. H., Berczik P., Shukir-galiyev B., Spurzem R., 2026, preprint ([arXiv:2606.04509](https://arxiv.org/abs/2606.04509))
- Anosova J. P., 1986, *Ap&SS*, 124, 217
- Appel S. M. et al., 2026, *AJ*, 172, 12
- Bate M. R., Bonnell I. A., 1997, *MNRAS*, 285, 33
- Beuther H., Kuiper R., Tafalla M., 2025, *ARA&A*, 63, 1
- Blaauw A., 1961, *Bull. Astron. Inst. Netherlands*, 15, 265
- Bonnell I. A., Bate M. R., Clarke C. J., Pringle J. E., 2001, *MNRAS*, 323, 785
- Bonnell I. A., Bate M. R., Vine S. G., 2003, *MNRAS*, 343, 413
- Butler M. J., Tan J. C., 2012, *ApJ*, 754, 5
- Butler M. J., Tan J. C., Teyssier R., Rosdahl J., Van Loo S., Nickerson S., 2017, *ApJ*, 841, 82
- Carretero-Castrillo M., Ribó M., Paredes J. M., Holgado G., Martínez-Sebastián C., Simón-Díaz S., 2025, preprint ([arXiv:2510.21577](https://arxiv.org/abs/2510.21577))
- Chabrier G., 2003, *PASP*, 115, 763
- Chandrasekhar S., 1943, *ApJ*, 97, 255
- Chon S., Vigna-Gómez A., 2026, preprint ([arXiv:2601.06251](https://arxiv.org/abs/2601.06251))
- Costa Silva A. R. et al., 2022, *A&A*, 659, A23
- Cournoyer-Cloutier C. et al., 2024, *ApJ*, 977, 203
- Crowe S. et al., 2024, *A&A*, 682, A2
- Da Rio N., Robberto M., Soderblom D. R., Panagia N., Hillenbrand L. A., Palla F., Stassun K. G., 2010, *ApJ*, 722, 1092
- Dale J. E., Ercolano B., Bonnell I. A., 2012, *MNRAS*, 424, 377
- Dale J. E., Ngoumou J., Ercolano B., Bonnell I. A., 2014, *MNRAS*, 442, 694
- Devine K. E., Churchwell E. B., Indebetouw R., Watson C., Crawford S. M., 2008, *AJ*, 135, 2095
- Dinçel B. et al., 2025, preprint ([arXiv:2511.14686](https://arxiv.org/abs/2511.14686))
- Domínguez R., Fellhauer M., Blafla M., Farias J. P., Dabringhausen J., 2017, *MNRAS*, 472, 465
- Elmegreen B. G., 2011, in Charbonnel C., Montmerle T., eds, *EAS Publications Series Vol. 51, Triggered Star Formation*. EDP, p. 45
- Elmegreen B. G., Hunter D. A., 2010, *ApJ*, 712, 604
- Fajrin M., Armstrong J. J., Tan J. C., Farias J. P., Eyer L., 2025, *MNRAS*, 537, 1320
- Farias J. P., Tan J. C., 2023, *MNRAS*, 523, 2083 (Paper III)
- Farias J. P., Tan J. C., Chatterjee S., 2017, *ApJ*, 838, 116 (Paper I)
- Farias J. P., Tan J. C., Chatterjee S., 2019, *MNRAS*, 483, 4999 (Paper II)
- Gammie C. F., 2001, *ApJ*, 553, 174
- Grudić M. Y., Guszejnov D., Offner S. S. R., Rosen A. L., Raju A. N., Faucher-Giguère C.-A., Hopkins P. F., 2022, *MNRAS*, 512, 216
- Guszejnov D., Hopkins P. F., Krumholz M. R., 2017, *MNRAS*, 468, 4093
- Heggie D. C., 1975, *MNRAS*, 173, 729
- Hills J. G., 1975, *AJ*, 80, 1075
- Inutsuka S.-i., Miyama S. M., 1997, *ApJ*, 480, 681
- Issac N., Tej A., Liu T., Varricatt W., Vig S., Ishwara Chandra C. H., Schultheis M., Nandakumar G., 2020, *MNRAS*, 497, 5454
- Karam J., Sills A., 2022, *MNRAS*, 513, 6095
- Kenyon S. J., Hartmann L., 1995, *ApJS*, 101, 117
- Kinoshita S. W., Nakamura F., Wu B., 2021, preprint ([arXiv:2108.05554](https://arxiv.org/abs/2108.05554))
- Kirk H. et al., 2017, *ApJ*, 838, 114
- Klein R. I., Sandford M. T. II, Whitaker R. W., 1980, *Space Sci. Rev.*, 27, 275
- Kölligan A., Kuiper R., 2018, *A&A*, 620, A182
- Kroupa P., 2001, *MNRAS*, 322, 231
- 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., Maschberger T., Moeckel N., Clarke C. J., Bastian N., Bonnell I. A., 2012, *MNRAS*, 419, 841
- Lada C. J., Lada E. A., 2003, *ARA&A*, 41, 57
- Lewis S. C. et al., 2023, *ApJ*, 944, 211
- Luhman K. L., 2004, *ApJ*, 602, 816
- Maity A. K., Dewangan L. K., Bhadari N. K., Ojha D. K., Chen Z., Pandey R., 2023, *MNRAS*, 523, 5388
- Matzner C. D., 2002, *ApJ*, 566, 302
- McKee C. F., Tan J. C., 2003, *ApJ*, 585, 850
- Moe M., Di Stefano R., 2017, *ApJS*, 230, 15
- Nomoto K., Kobayashi C., Tominaga N., 2013, *ARA&A*, 51, 457
- Offner S. S. R., Moe M., Kratter K. M., Sadavoy S. I., Jensen E. L. N., Tobin J. J., 2023, in Inutsuka S., Aikawa Y., Muto T., Tomida K., Tamura M., eds, *ASP Conf. Ser. Vol. 534, Protostars and Planets VII*. Astron. Soc. Pac., San Francisco, p. 275
- Pecaut M. J., Mamajek E. E., Bubar E. J., 2012, *ApJ*, 746, 154
- Pineda J. E. et al., 2015, *Nature*, 518, 213
- Poveda A., Ruiz J., Allen C., 1967, *Bol. Obs. Tonantzintla Tacubaya*, 4, 86
- Preibisch T., 2012, *Res. Astron. Astrophys.*, 12, 1
- Raghavan D. et al., 2010, *ApJS*, 190, 1
- Reggiani M., Robberto M., Da Rio N., Meyer M. R., Soderblom D. R., Ricci L., 2011, *A&A*, 534, A83
- Reipurth B., Mikkola S., Connelley M., Valtonen M., 2010, *ApJ*, 725, L56
- Reipurth B. et al., 2026, preprint ([arXiv:2602.06465](https://arxiv.org/abs/2602.06465))
- Reynolds N. K. et al., 2021, *ApJ*, 907, L10
- Rogers H., Pittard J. M., 2013, *MNRAS*, 431, 1337
- Sana H. et al., 2012, *Science*, 337, 444
- Smith M. C., Sijacki D., Shen S., 2018, *MNRAS*, 478, 302
- Staff J. E., Tanaka K. E. I., Tan J. C., 2019, *ApJ*, 882, 123
- Staff J. E., Tanaka K. E. I., Ramsey J. P., Zhang Y., Tan J. C., 2023, *ApJ*, 947, 40
- Tan J. C., Beltrán M. T., Caselli P., Fontani F., Fuente A., Krumholz M. R., McKee C. F., Stolte A., 2014, in Beuther H., Klessen R. S., Dullemond C. P., Henning T., eds, *Protostars and Planets VI*. Univ. Arizona Press, Tucson, p. 149
- Tanaka K. E. I., Tan J. C., Zhang Y., 2017, *ApJ*, 835, 32
- Tanikawa A., Tajitsu A., Honda S., Maehara H., Sato B., Masuda K., Omiya M., Izumiura H., 2026, preprint ([arXiv:2601.21125](https://arxiv.org/abs/2601.21125))
- Telkamp Z. et al., 2025, *ApJ*, 986, 15
- Tokovinin A., 2026, preprint ([arXiv:2601.05006](https://arxiv.org/abs/2601.05006))
- Tokovinin A., Moe M., 2020, *MNRAS*, 491, 5158
- Tramper F., Sana H., de Koter A., Pauwels T., 2025, preprint ([arXiv:2509.18431](https://arxiv.org/abs/2509.18431))
- Vasyunina T., 2010, PhD thesis, Ruprecht-Karls University of Heidelberg, Germany
- Wall J. E., McMillan S. L. W., Mac Low M.-M., Klessen R. S., Portegies Zwart S., 2019, *ApJ*, 887, 62
- Wang P., Li Z.-Y., Abel T., Nakamura F., 2010, *ApJ*, 709, 27
- Wilking B. A., Gagné M., Allen L. E., 2008, in Reipurth B., ed., *Handbook of Star Forming Regions, Volume II*, ASP Monograph Publications, Vol. 5. Astron. Soc. Pac., San Francisco, p. 351
- Wu B., Tan J. C., Christie D., Nakamura F., Van Loo S., Collins D., 2017, *ApJ*, 841, 88
- Zhang Y., Tan J. C., 2018, *ApJ*, 853, 18
- Zummer M., Harmanec P., Barlow B., Blackford M., Švrčková J., 2025, preprint ([arXiv:2511.15544](https://arxiv.org/abs/2511.15544))

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.