Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **522,** 5340–5357 (2023) https://doi.org/10.1093/mnras/stad1287
Advance Access publication 2023 May 04

The mass-loss rates of star clusters with stellar-mass black holes: implications for the globular cluster mass function

Mark Gieles [1,2]★ and Oleg Y. Gnedin [3]

[1] *ICREA, Pg. Lluís Companys 23, E-08010 Barcelona, Spain*
[2] *Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona (IEEC-UB), Martí i Franquès 1, E-08028 Barcelona, Spain*
[3] *Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*

Accepted 2023 April 25. Received 2023 April 25; in original form 2023 March 7

ABSTRACT

Stellar-mass black holes (BHs) can be retained in globular clusters (GCs) until the present. Simulations of GC evolution find that the relaxation driven mass-loss rate is elevated if BHs are present, especially near dissolution. We capture this behaviour in a parametrized mass-loss rate, bench marked by results from *N*-body simulations, and use it to evolve an initial GC mass function (GCMF), similar to that of young massive clusters in the Local Universe, to an age of 12 Gyr. Low-metallicity GCs ([Fe/H] $\lesssim -1.5$) have the highest mass-loss rates, because of their relatively high BH masses, which combined with their more radial orbits and stronger tidal field in the past explains the high turnover mass of the GCMF ($\sim 10^5\,\mathrm{M}_\odot$) at large Galactic radii ($\gtrsim 10$ kpc). The turnover mass at smaller Galactic radii is similar because of the upper mass truncation of the initial GCMF and the lower mass-loss rate due to the higher metallicities. The density profile in the Galaxy of mass lost from massive GCs ($\gtrsim 10^5\,\mathrm{M}_\odot$) resembles that of nitrogen-rich stars in the halo, confirming that these stars originated from GCs. We conclude that two-body relaxation is the dominant effect in shaping the GCMF from a universal initial GCMF, because including the effect of BHs reduces the need for additional disruption mechanisms.

Key words: stars: black holes – globular clusters: general – galaxies: star clusters: general.

1 INTRODUCTION

Globular cluster (GC) systems in the Milky Way and external galaxies have peaked logarithmic mass and luminosity functions, with a typical luminosity $M_V \simeq -7.5$ (for example, Harris 2001; Jordán et al. 2007), corresponding to a peak mass $\sim 2\times10^5\,\mathrm{M}_\odot$ and a dispersion $\sigma_{\log_{10}M} \simeq 0.5$. This is markedly different from young star clusters, which form with a power-law mass function with a slope of about -2 (see Portegies Zwart, McMillan & Gieles 2010; Krumholz, McKee & Bland-Hawthorn 2019, for reviews). Old GCs may have formed with a similar mass function, because low-mass clusters had time to dissolve as the result of various disruptive effects, such as two-body relaxation, tidal shocks, and interaction with dense molecular gas clouds. This disruption could turn over a power-law initial GC mass function (GCMF) and impose a typical mass scale of $\sim 10^5\,\mathrm{M}_\odot$ in the surviving GCs (for example, Okazaki & Tosa 1995; Fall & Zhang 2001; Prieto & Gnedin 2008; Elmegreen 2010; Kruijssen 2015). Adopting the hypothesis that the physics of cluster formation in giant molecular clouds is similar at all cosmic times (Harris & Pudritz 1994; Elmegreen & Efremov 1997), several recent studies have confirmed that massive star clusters formed in high-redshift galaxies would evolve into old clusters matching the age–metallicity distribution and the spatial and kinematic distributions of observed GC systems (Choksi, Gnedin & Li 2018; Pfeffer et al. 2018; Kruijssen et al. 2019; Rodriguez et al. 2023). However, the resulting GCMF in these models tends to be skewed towards lower masses compared to the observed GCMF. Reproducing the shape of the GCMF is, therefore, one of the last remaining hurdles to confirm that star cluster formation is a universal mechanism in all epochs and environments.

Some studies have suggested that GCs had a typical mass scale imprinted at formation (for example, Peebles & Dicke 1968; Fall & Rees 1985; Bromm & Clarke 2002; Kimm et al. 2016), and indeed an initially peaked mass distribution would preserve its shape as clusters lose mass (Vesperini 2000; Fall & Zhang 2001). However, such scenarios rely on physical conditions in the galactic interstellar medium that are unlikely to produce giant molecular clouds massive and dense enough to host proto-GCs (e.g. Forbes et al. 2018). We therefore prefer the hypothesis that the initial GCMF is universal across cosmic time and that disruption is responsible for the current shape of the GCMF.

The relative contribution of various disruption mechanisms is still debated. Because of the high rate of close stellar encounters in GCs, it is natural first to explore the effect of two-body relaxation in the large-scale galactic tidal field (hereafter, 'evaporation'). The mass-loss rate due to evaporation depends on the strength of the tidal field (for example, Lee & Ostriker 1987; Chernoff & Weinberg 1990; Baumgardt & Makino 2003), and therefore it predicts an anticorrelation between the turnover mass ($M_{\rm TO}$) and galactocentric radius (R) for clusters in a static galactic potential and with constant velocity anisotropy. In contrast, the observed $M_{\rm TO}$ varies only mildly with R in the Milky Way (see, for example, fig. 8 in McLaughlin & Fall 2008) and is also remarkably constant across galactic environments (Jordán

★ E-mail: mgieles@icc.ub.edu

et al. 2007). We refer to this tension between model predictions and the observed near-universality of the GCMF as 'the GCMF problem'.

Various studies have attempted to resolve the GCMF problem by either assuming a strong radially biased velocity anisotropy at large Galactic radii (for example, Fall & Zhang 2001); invoking additional universal disruptive effects, such as stellar evolution (Vesperini & Zepf 2003) and gas expulsion (Baumgardt, Kroupa & Parmentier 2008); or assuming that the mass-loss rate depends mainly on present-day density (McLaughlin & Fall 2008). However, these assumptions are all in tension with results from observations (Vesperini et al. 2003; Vasiliev 2019), theory (Hénon 1961; Gieles, Heggie & Zhao 2011), and numerical simulations of cluster evolution in tidal fields (Lee & Ostriker 1987; Baumgardt & Makino 2003; Gieles & Baumgardt 2008).

In recent years, much attention has been given to the disruption by tidal interactions with giant molecular clouds in the first ~ 1 Gyr (Elmegreen 2010; Kruijssen 2015; Pfeffer et al. 2018). Tidal shocks preferentially destroy low-density clusters (Spitzer 1958; Ostriker, Spitzer & Chevalier 1972) and therefore not necessarily low-mass clusters, but relaxation leads to an expansion of low-mass clusters, reducing their densities, such that the combined effect of relaxation and tidal perturbations leads to a mass-dependence of the disruption time-scale that is similar to that of evaporation (Gieles & Renaud 2016).

Most cluster population studies mentioned above rely on prescriptions for evaporation based on theory of equal-mass clusters by Hénon (1961), or results of numerical N-body simulations of clusters with a stellar mass function and stellar evolution (for example, Baumgardt & Makino 2003), but without stellar-mass black holes (BHs). However, BH candidates have been reported in several Milky Way GCs (Strader et al. 2012; Chomiuk et al. 2013; Miller-Jones et al. 2015; Giesers et al. 2018; Kamann et al. 2020) and in extragalactic clusters (Maccarone et al. 2007, 2011; Barnard, Garcia & Murray 2012; Saracino et al. 2022).

These discoveries of BHs in clusters led to various modelling efforts of GCs with BHs, that showed that clusters in a tidal field dissolve faster if they retain a significant fraction of their BHs after natal kicks (Chatterjee, Rodriguez & Rasio 2017; Peuten et al. 2017; Giersz et al. 2019; Kremer et al. 2020; Wang 2020; Gieles et al. 2021). Apart from shortening the total lifetime, the mass evolution over time is also different, in the sense that the (absolute) mass-loss rate increases towards dissolution. This is because tidally limited clusters with a BH mass fraction at a critical value of a few per cent will lose BH mass at the same rate as stellar mass and therefore maintain that constant BH mass fraction (Breen & Heggie 2013). If the mass fraction in BHs is higher (lower), the BH fraction continues to increase (decrease) (Banerjee & Kroupa 2011; Gieles et al. 2021). An increasing BH fraction with time leads to an increasing (absolute) mass-loss rate and such an abrupt dissolution leads to a concave shape of the mass evolution with time $M(t)$ (Giersz et al. 2019). In contrast, for clusters without BHs the shape of $M(t)$ is convex (Baumgardt 2001; Gieles & Baumgardt 2008). We will refer to the concave and convex shapes of $M(t)$ as 'jumping' and 'skiing', respectively, following terminology from Contenta, Varri & Heggie (2015). In this work we propose an analytical prescription for the mass-loss rate that allows for different shapes of $M(t)$, with a flexible dependence of the total lifetime on the initial mass, informed by results of a grid of direct N-body simulations. We then use it to model the evolution of the GCMF in a Milky Way-like galaxy.

This paper is organized as follows. In Section 2, we analyse the mass-loss rate in N-body simulations of star clusters with BHs. In Section 3, we parametrize the mass-loss rate guided by the N-body simulations. We present a model for the evolution of the GCMF in Section 4, present the results in Section 5, and then discuss broader implications of our results in Section 6. Our conclusions are summarized in Section 7.

2 INSIGHT FROM N-BODY MODELS

2.1 Description of the models

To quantify the effect of BHs on the mass-loss rate ($\dot{M}$) of star clusters, we use the N-body models presented in Gieles et al. (2021) that were performed with NBODY6++GPU (Aarseth 2003; Wang et al. 2015). The grid of models in that work was intended to find the initial conditions of the Milky Way GC Palomar 5 (Pal 5, hereafter), hence all clusters are on the same orbit in a three-component Milky Way potential, with an apocentre distance ≈ 15.5 kpc and a pericentre distance ≈ 6.5 kpc (implying an orbital eccentricity $\epsilon \simeq 0.41$). We use the first 11 models from their table 1, which is a grid of models with different initial density within the half-mass radius ($r_{\rm h,0}$) of $\rho_{\rm h,0} \equiv 3M_0/(8\pi r_{\rm h,0}^3) = \{30, 100, 300, 1000\}\,{\rm M}_\odot\,{\rm pc}^{-3}$ and number of stars $N = \{0.5, 1, 2\} \times 10^5$, which for the adopted Kroupa (2001) stellar initial mass function (IMF) in the range $0.1 - 100\,{\rm M}_\odot$ corresponds to initial cluster masses $M_0 = \{0.32, 0.64, 1.28\} \times 10^5\,{\rm M}_\odot$. The models adopt the rapid supernova mechanism (Fryer et al. 2012) with the natal kicks lowered by the amount of fallback such that momentum is conserved. As a result, 63 per cent (73 per cent) of the number (mass) of BHs do not receive a natal kick for the adopted IMF and the metallicity of $Z = 0.0006$ ([Fe/H] $\simeq -1.4$, using $Z_\odot = 0.014$ for the solar metallicity; Asplund, Amarsi & Grevesse 2021). The model with $\rho_{\rm h,0} = 10^3\,{\rm M}_\odot\,{\rm pc}^{-3}$ and $N = 2 \times 10^5$ was not run for that study, and we run it here with the same settings as the other models. We also run here two additional models with higher metallicity: $Z = 0.006$ ([Fe/H] $\simeq -0.4$) and $Z = 0.017$ ([Fe/H] $\simeq 0$), both with $N = 10^5$ and $\rho_{\rm h,0} = 300\,{\rm M}_\odot\,{\rm pc}^{-3}$.

We compare the mass-loss rates of these models to the frequently-cited N-body models of Baumgardt & Makino (2003). These models consider a galactic tidal field due to a singular isothermal sphere (SIS) with circular velocity $V_{\rm c} = 220\,{\rm km\,s}^{-1}$, the effects of stellar evolution, and a stellar IMF truncated at $15\,{\rm M}_\odot$ such that no BHs form. Their mass-loss rates can be approximated as (Lamers, Baumgardt & Gieles 2013; Choksi et al. 2018)

$$\dot{M} \simeq -30\,{\rm M}_\odot\,{\rm Myr}^{-1} \left(\frac{M}{2\times10^5\,{\rm M}_\odot}\right)^{1/3} \frac{\Omega_{\rm tid}}{0.32\,{\rm Myr}^{-1}}. \quad (1)$$

The influence of the tidal field is captured by $\Omega_{\rm tid}$, which depends on the tidal and centrifugal forces. For circular orbits, it can be expressed through the first and third eigenvalues of the tidal tensor as $\Omega_{\rm tid} = \sqrt{\lambda_1 - \lambda_3}$ (Renaud, Gieles & Boily 2011; Chen & Gnedin 2023). For the SIS $\lambda_1 = -\lambda_3 = V_{\rm c}^2/R^2$ such that $\Omega_{\rm tid} = \sqrt{2}V_{\rm c}/R$. For eccentric orbits we use $\Omega_{\rm tid} = \sqrt{2}V_{\rm c}/R_{\rm eff}$, where $R_{\rm eff}$ is the radius of the circular orbit with the same lifetime. Baumgardt & Makino (2003) show that for relaxation driven mass-loss of clusters in an SIS this effective radius is $R_{\rm eff} \equiv R_{\rm p}(1+\epsilon) = R_{\rm a}(1-\epsilon)$, where $R_{\rm p}$ and $R_{\rm a}$ are the pericentre and apocentre distance of the orbit, respectively (see also Cai et al. 2016). We normalize $\Omega_{\rm tid}$ to a value corresponding to $V_{\rm c} = 220\,{\rm km\,s}^{-1}$ and $R_{\rm eff} = 1$ kpc.

For clarity, in this work we also adopt the SIS to approximate the Galaxy potential. The Pal 5 models did not evolve in an SIS, but in a more realistic three-component Milky Way, so we can use those results as an anchor point from which we extrapolate to larger and smaller $R_{\rm eff}$ by using the simple analytic properties of the SIS. For the Pal 5 models $R_{\rm eff} \simeq 9.15$ kpc and therefore $\Omega_{\rm tid} \simeq 0.035\,{\rm Myr}^{-1}$.

Table 1. Parameters of the *N*-body models shown in Fig. 1. For each initial density, models with $N = \{0.5, 1, 2\} \times 10^5$ stars were run, corresponding to initial masses of $M_0 \simeq \{0.32, 0.64, 1.28\} \times 10^5\,\mathrm{M}_\odot$. All models adopt a metallicity of $Z = 0.0006$ ([Fe/H] $\simeq -1.4$) and two additional models with higher metallicity were run for $\mathcal{R}_0 = 29$. The parameters $\dot{M}_{\rm ref}$ and y in the last two columns are used to approximate $\dot{M}(M, M_0)$ (equation 4) in Fig. 2 (dashed line, top row).

$\rho_{\rm h,\,0}$ ($\mathrm{M}_\odot\,\mathrm{pc}^{-3}$)	$\mathcal{R}_0 = \frac{\rho_{\rm h,0}}{\rho_{\rm h,f}}$	$\frac{r_{\rm h,0}}{r_{\rm J,eff}}$	$\dot{M}_{\rm ref}$ ($\mathrm{M}_\odot\,\mathrm{Myr}^{-1}$)	y
30	2.9	0.10	−95	2
100	9.5	0.068	−60	1.75
300	29	0.047	−45	1.33
1000	95	0.032	−30	0.67

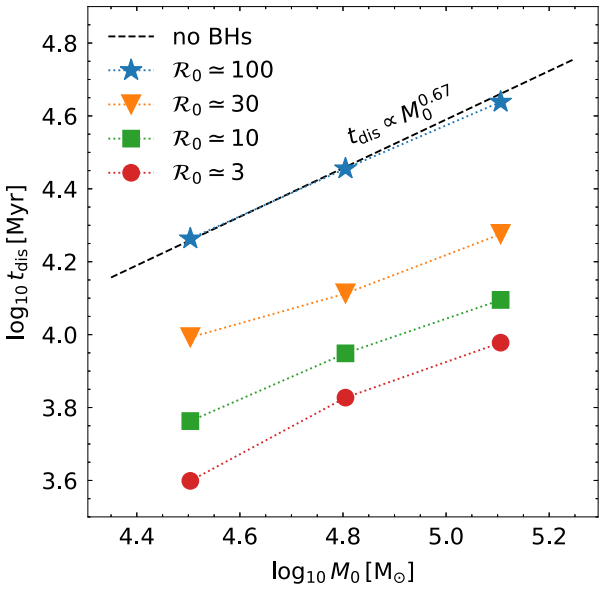

Figure 1. Disruption times ($t_{\rm dis}$) for different M_0 and $\mathcal{R}_0$ (see Table 1) of all 12 *N*-body models on the orbit of Pal 5 from Gieles et al. (2021). The dashed line shows the expected $t_{\rm dis}$ for models without BHs, as derived from equation (1), see the text for detail. The densest *N*-body models ($\mathcal{R}_0 = 100$) have similar $t_{\rm dis}$ because they eject all their BHs dynamically (Section 2.2). Lower density clusters dynamically retain a BH population until dissolution and disrupt up to 4.5 times faster.

The density within the effective Jacobi radius ($r_{\rm J,\,eff}$) for an SIS is given by $\rho_{\rm J,eff} = 3/(2\pi G)\Omega_{\rm tid}^2$, where G is the gravitational constant. For the models on the orbit of Pal 5 it is $\rho_{\rm J,eff} \simeq 0.064\,\mathrm{M}_\odot\,\mathrm{pc}^{-3}$.

To describe the dependence of our results on the initial cluster density, we introduce a dimension-less parameter $\mathcal{R} \equiv \rho_{\rm h}/\rho_{\rm h,f}$, where $\rho_{\rm h,\,f}$ is the half-mass density of a Roche-filling cluster. Roche filling in the context of clusters is ill-defined, but here we take it as a cluster that has $r_{\rm h}/r_{\rm J} = 0.145$, which is the filling factor in the tidally limited cluster on a circular orbit of Hénon (1961). The Pal 5 models have $\rho_{\rm h,f} = 0.5 \times 0.145^{-3}\,\rho_{\rm J,eff} \simeq 10.5\,\mathrm{M}_\odot\mathrm{pc}^{-3}$. The relation between density and filling factor in terms of radius is $r_{\rm h}/r_{\rm J,eff} = 0.145\mathcal{R}^{-1/3}$. Table 1 relates $\rho_{\rm h,\,0}$ of the *N*-body models to these more physically relevant quantities.

In Fig. 1 we show the disruption times ($t_{\rm dis}$) of all 12 *N*-body models with low metallicity. We define $t_{\rm dis}$ as the time needed for the mass to reach zero, although in the *N*-body models we only determine the time when the cluster reaches a low enough mass of 300 $\mathrm{M}_\odot$. The dashed line shows $t_{\rm dis}$ following from $\dot{M}$ for models without BHs (equation 1), which for this mass dependence of $\dot{M}$ is $t_{\rm dis}(M_0) = 1.5 M_{\rm i}/|\dot{M}(M_{\rm i})|$ (Lamers et al. 2005), where

$$M_{\rm i} \equiv \mu_{\rm sev} M_0 \tag{2}$$

is the initial mass after (most) stellar evolution related mass-loss has occurred. From hereon we make the simplifying assumption that stellar evolution happens independently from evaporation, which is justified by the different time-scales on which they operate (several 10 Myr for most stellar evolution mass-loss to occur versus several Gyr for evaporation). Here $\mu_{\rm sev} \simeq 0.55$ is the remaining mass fraction after stellar evolution for a metallicity of $Z = 0.0006$ ([Fe/H] $\simeq -1.4$). Expressing $t_{\rm dis}$ in terms of the initial mass after stellar evolution is needed because equation (1) only describes the mass-loss due to evaporation. This prediction for $t_{\rm dis}$ agrees well with the results from the densest clusters, while the clusters with $\mathcal{R}_0 \simeq \{30, 10, 3\}$ dissolve approximately a factor of $\{2, 3, 4.5\}$ faster. In the next section, we discuss in more detail the role of BHs in this trend of $t_{\rm dis}(\mathcal{R}_0)$.

2.2 Mass-loss rates

In this section, we describe how $\dot{M}$ depends on the initial conditions of the clusters. Because all clusters lose about 45 per cent of their initial mass by stellar evolution, mostly in the first Gyr, and we are here interested in evaporation, we determine $\dot{M}$ in the range $500\,\mathrm{M}_\odot \leq M < M_{\rm i}$ in mass bins with widths of $3 \times 10^3\,\mathrm{M}_\odot$. In Fig. 2 we show $\dot{M}$ from the *N*-body models in the top row, with different M_0 (different colours and symbols) and different initial densities (different columns). The bottom row shows the remaining mass in BHs ($M_{\rm BH}$). The clusters with relatively low densities (left two columns) keep a significant fraction of their BHs and the mass-loss rate of these models increases towards dissolution.

There is a clear trend for higher density clusters to lose more of their BHs, which is the result of their shorter relaxation time (Breen & Heggie 2013). The densest clusters ($\mathcal{R}_0 \simeq 100$, right-hand column) eject almost all BHs early and evolve along similar tracks as models without BHs (equation 1), shown as black-dashed lines in the top row. This is why their $t_{\rm dis}$ is similar to those of clusters without BHs (Fig. 1). Breen & Heggie (2013) explain that for tidally limited clusters there exists a critical $f_{\rm BH} \simeq 0.1$ at which the fraction of the total mass that is lost is in the form of BHs equals 0.1, such that $f_{\rm BH}$ remains constant. If $f_{\rm BH} \lesssim 0.1$ then all BHs are ejected, while if $f_{\rm BH} \gtrsim 0.1$ the cluster evolves towards a 100 per cent BH cluster (see also Banerjee & Kroupa 2011). This was derived for idealized two-component models. In our models we find that this critical fraction is lower: $f_{\rm BH} \simeq 0.025$. This has consequences for clusters with higher metallicity, because they form with a lower $f_{\rm BH}$ than metal-poor clusters and therefore drop more easily below the critical $f_{\rm BH} \simeq 0.025$ (Section 2.3).

The fact that the different coloured $\dot{M}$ points in each panel do not overlap shows that at the same remaining mass M, models with different M_0 have different $\dot{M}$, because their $f_{\rm BH}$ are different. This behaviour is reproduced for most parts of the evolution by the dotted lines, which are a simple parametrization of $\dot{M}$, described in more detail in Section 3 (equation 4). They are power-law relations for $\dot{M}(M, M_0)$ of the form $\dot{M} \propto (M/M_{\rm i})^a M_{\rm i}^{1/3}$, where the value of a required to described (most of) the data ranges from $a = -1$ to $a = 1/3$. This relation results in a dependence $t_{\rm dis} \propto M_0^{2/3}$ (see Fig. 1) independent of the value of a, that is, the same M_0 dependence as was found for models without BHs (equation 1).

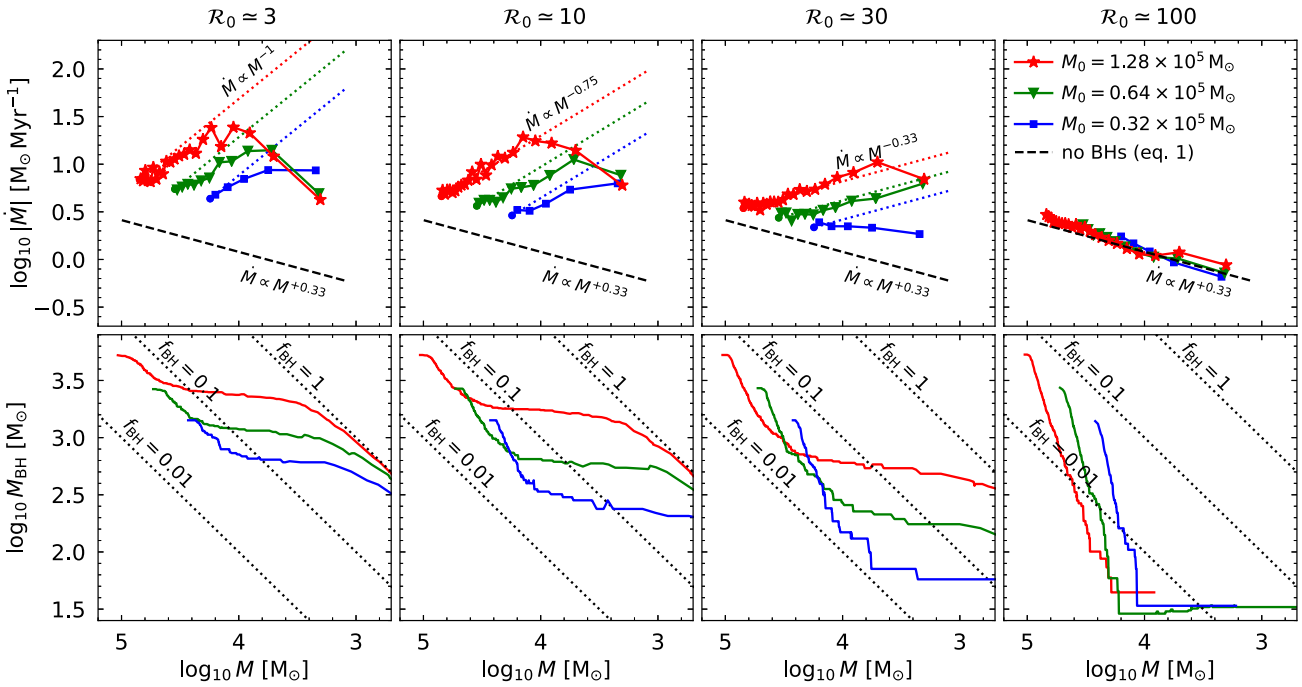

Figure 2. Top: Mass-loss rates of N-body models on the orbit of Pal 5 from Gieles et al. (2021). Current cluster mass $M(t)$ acts as a proxy for time, flowing from left to right. The dotted lines show $\dot{M}$ curves from equation (4) that match the N-body points. The dashed line is the same in every top panel and shows $\dot{M}$ without the effect of BHs. Bottom: Mass in BHs in the same N-body models. The diagonal dotted lines mark constant BH fractions of 1 per cent, 10 per cent, and 100 per cent.

For $f_{\rm BH} \gtrsim 0.3$, $|\dot{M}|$ decreases again, because then the stars become less important and the cluster evolves from a two-component model (stars and BHs) to a single-component model (only BHs). This leads to a narrower mass spectrum and a slower evolution, but still $\sim$10 times faster than for a cluster with only stars. In fact, $\dot{M}(M)$ is then evolving parallel to the dashed line shown in the top row of Fig. 2, but above it because of the higher mean mass, which reduces $\tau_{\rm rh}$.

Fig. 3 shows $\dot{M}$ as a function of the remaining BH fraction $f_{\rm BH}$ for all low-metallicity models. The increase of $|\dot{M}|$ with $f_{\rm BH}$ can be approximated by a linear relation $\dot{M} \simeq -12.5\,{\rm M}_\odot\,{\rm Myr}^{-1}\,(f_{\rm BH}/0.1)$. As shown earlier, clusters evolve at roughly constant $M_{\rm BH}$ at late stages, such that a relation $\dot{M} \propto f_{\rm BH}$ implies $\dot{M} \propto M^{-1}$. This leads to a strongly 'jumping' $M(t)$, very different from the result of models without BHs ($\dot{M} \propto M^{1/3}$).

Clusters lose most of their BHs in the early expansion phase. Once the cluster density has become comparable to (some multiple of) the tidal density, the cluster evolves at approximately constant $M_{\rm BH}$ (see bottom row of Fig. 2). We can estimate how the BH loss in the expansion phase depends on the initial conditions. Breen & Heggie (2013) showed that the mass-loss rate of the BH population depends on the cluster properties as $\dot{M}_{\rm BH} \propto M/\tau_{\rm rh}$, where $\tau_{\rm rh}$ is the half-mass relaxation time-scale. We approximate the total BH mass lost as $\Delta M_{\rm BH} = \dot{M}_{\rm BH}\Delta t$, where Δt is the time the cluster needs to fill the Roche volume. In the expansion phase the density reduces in time as $\rho(t) \simeq \rho_{\rm h,0}(t/\tau_{\rm rh,0})^{-2}$, where $\tau_{\rm rh,0}$ is the initial $\tau_{\rm rh}$ (Hénon 1965; Gieles et al. 2011). So the time the cluster needs to expand to the tidal boundary is $\Delta t \simeq \tau_{\rm rh,0}\mathcal{R}_0^{1/2}$. Using also the initial values in the expression for $\dot{M}_{\rm BH}$, we thus find $\Delta M_{\rm BH} \propto M_0\mathcal{R}_0^{1/2}$ or $\Delta f_{\rm BH,0} \propto \mathcal{R}_0^{1/2}$. So the reduction of $f_{\rm BH}$ due to dynamical ejections depends only on the initial density, relative to the tidal density. The data in the bottom row of Fig. 2 show that indeed that the drop in $f_{\rm BH}$ is larger for the higher initial densities. We also note that there is a small dependence on M_0 for $\mathcal{R}_0 = 10 - 30$, with the drop in $M_{\rm BH}$ being (relatively) large for the low-mass clusters. This is in the regime where only a handful of BHs are left and therefore the theory of Breen & Heggie no longer holds and we will not attempt to capture this.

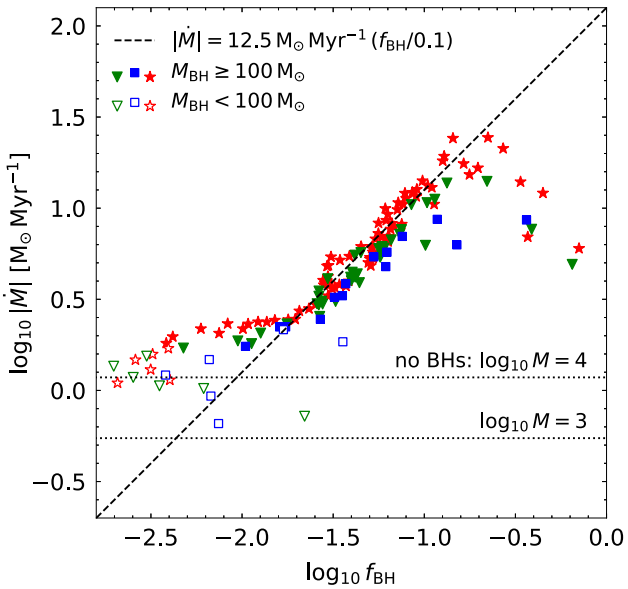

Figure 3. Mass-loss rate at different $f_{\rm BH}$ for all models with [Fe/H] $= -1.4$. We plot the models with $M_{\rm BH} < 100\,{\rm M}_\odot$ as open symbols, because these only have a few BHs, and these clusters are not expected to behave in the same way as the cluster with a population of BHs. The dashed line is a linear approximation for the relation $\dot{M}(f_{\rm BH})$.

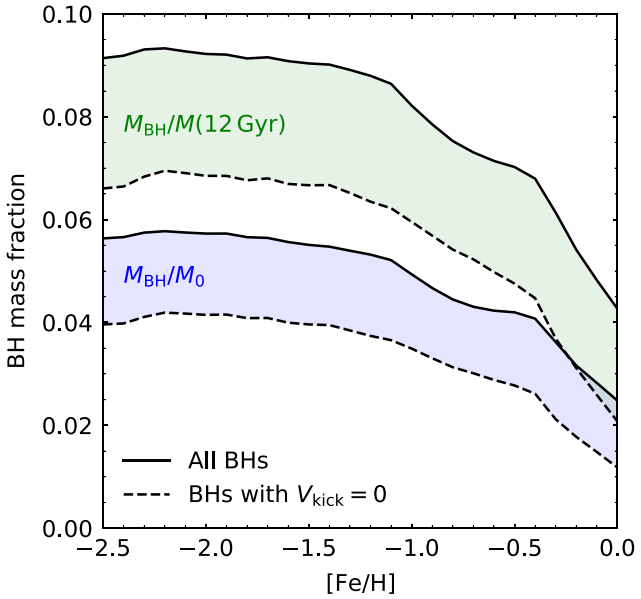

Figure 4. Mass fraction in stellar-mass BHs as a function of metallicity, from SSE (Hurley, Pols & Tout 2000; Banerjee et al. 2020). BHs that receive no kick are shown by dashed lines, the sum of BHs with and without kick is shown by solid lines.

There exists a critical density $\mathcal{R}_{\rm cr}$ between the models with $\mathcal{R}_0 = 30$ and $\mathcal{R}_0 = 100$ above which all BHs are ejected. We propose a simple relation for the drop in $f_{\rm BH}$ because of the dynamical ejection of the form

$$\Delta f_{\rm BH} \propto \left(\frac{\mathcal{R}_0}{\mathcal{R}_{\rm cr}}\right)^{1/2}, \tag{3}$$

for $\mathcal{R}_0 < \mathcal{R}_{\rm cr} \simeq 50$. Metal-rich clusters have lower $f_{\rm BH,0}$, and because the constant of proportionality in equation (3) does not depend on $f_{\rm BH,0}$, the critical density is found from equating $\Delta f_{\rm BH} = f_{\rm BH,0}$ such that $\mathcal{R}_{\rm cr} \propto f_{\rm BH,0}^2$. We discuss the consequences for metallicity next.

2.3 Metallicity

Stars of higher metallicity have stronger winds, which results in lower remnant masses. To quantify this effect on the masses of BHs, we adopt a Kroupa, Aarseth & Hurley (2001) IMF in the range $0.1 - 100\,{\rm M}_\odot$ and evolve the stars to an age of 12 Gyr with the single stellar evolution model SSE by Hurley et al. (2000), with the recent update for massive star winds from Banerjee et al. (2020). We then compute the mass fraction in BHs and show it in Fig. 4, both in terms of the initial mass ($f_{\rm BH,0} = M_{\rm BH}/M_0$, blue shaded region) and in terms of the total mass at 12 Gyr ($f_{\rm BH} = M_{\rm BH}/M$, green shaded region), which is roughly a factor $1/\mu_{\rm sev} \simeq 1.8$ higher. Since some of the BHs receive natal kicks (see Section 2.1), we show the two extreme cases where all BHs that are kicked are either lost (dashed lines) or retained (full lines). As can be seen, $f_{\rm BH}$ is roughly constant below [Fe/H] $\simeq -1.5$ (typical metal-poor GC) and decreases approximately by a factor of two going to [Fe/H] $\simeq -0.5$ (typical metal-rich GC).

To illustrate the effect of [Fe/H] and the resulting $f_{\rm BH}$ on the evolution of the cluster, we show here the results of two N-body models with higher metallicity $Z = 0.006$ ([Fe/H] $\simeq -0.4$) and $Z = 0.017$ ([Fe/H] $\simeq 0.1$) on the same orbit as the other models, and with

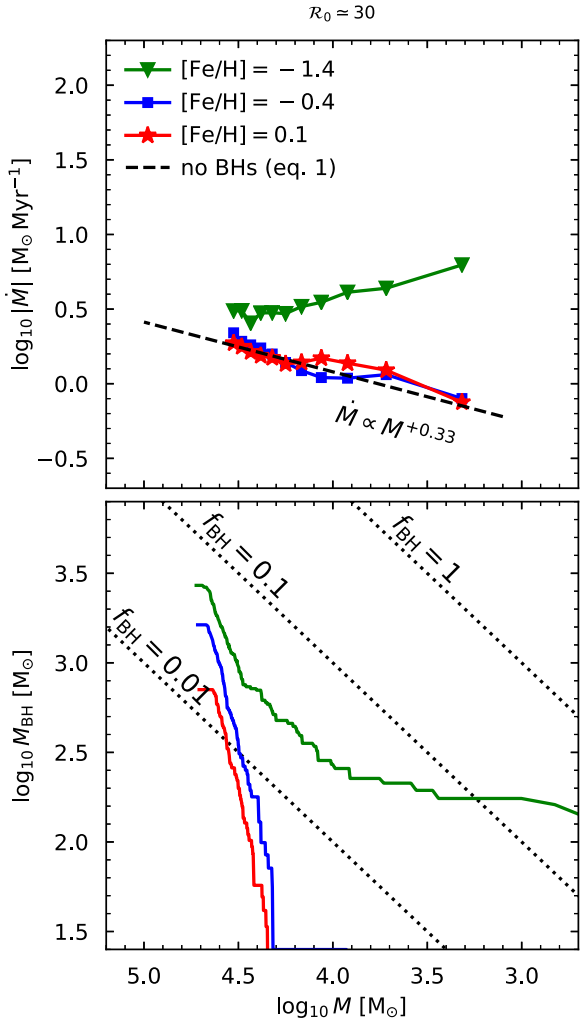

Figure 5. Results of N-body models with different initial metallicities. The model with [Fe/H] $= -1.4$ is the same as shown in Fig. 2. The models with higher [Fe/H] have lower initial $f_{\rm BH}$, leading to complete ejection of all BHs and mass-loss rates comparable to models without BHs (dashed line).

$N = 10^5$ and $\rho_{\rm h,0} = 300\,{\rm M}_\odot\,{\rm pc}^{-3}$ ($\mathcal{R}_0 \simeq 30$). For the metal-poor models shown in Fig. 2, sufficient number of BHs were retained for them to have a noticeable effect on $\dot{M}$ at this density. The effect of higher metallicity on $f_{\rm BH}$ and $\dot{M}$ in N-body models is shown in Fig. 5. From this plot we see that the lower initial value of $f_{\rm BH}$ results in all BHs being dynamically ejected and their $\dot{M}$ following the results of models without BHs (equation 1). Because $f_{\rm BH,0}$ is a factor of $\sim$2 lower at these higher metallicities, the critical initial density for ejection of all BHs is a factor of $\sim$4 lower (see the text below equation 3), so $\mathcal{R}_{\rm cr} \simeq 13$ instead of 50. This $\mathcal{R}_{\rm cr}$ is now lower than the density of these models ($\mathcal{R}_0 \simeq 30$) and explains why all BHs are ejected. We will use this metallicity dependence of $\dot{M}$ in the population model (Section 4).

3 A NOVEL PARAMETRIZATION OF THE CLUSTER MASS-LOSS RATE

In this section, we consider a simple analytical model for the evolution of the GCMF that results from adding the effects of BHs.

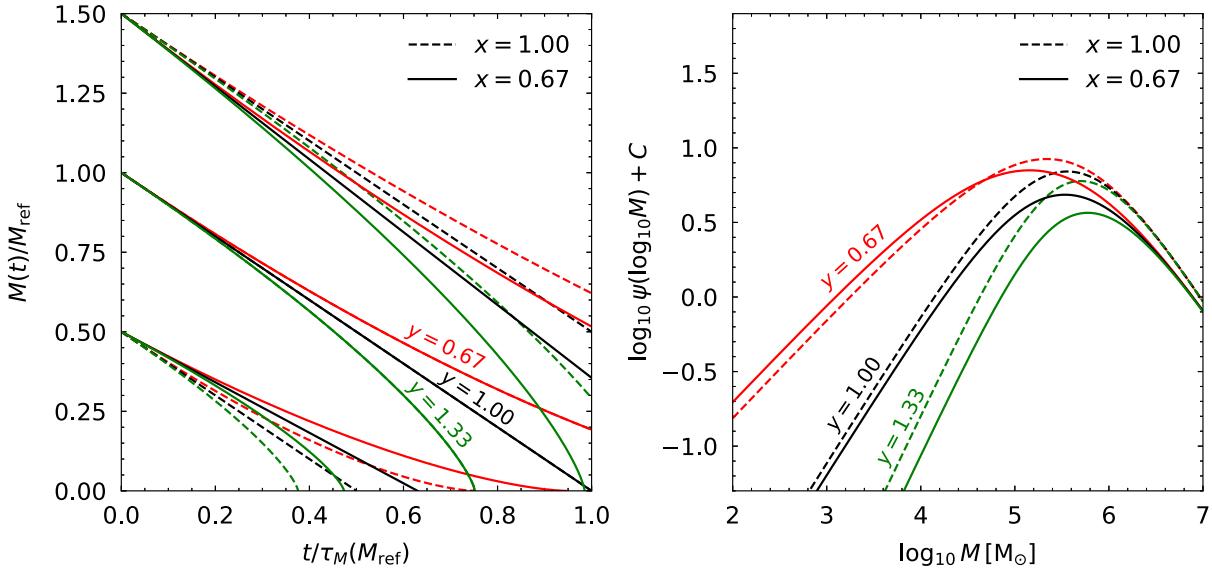

Figure 6. Left-hand panel: Evolution of cluster mass for different M_0, x and y. Here $M_{\rm ref}$ is a constant reference mass. Right-hand panel: Evolved cluster mass function for the same set of x and y as in the left-hand panel and for $\dot{M}_{\rm ref} = -30\,{\rm M}_\odot\,{\rm Myr}^{-1}$ and $\Omega_{\rm tid} = 0.32\,{\rm Myr}^{-1}$.

3.1 Mass-loss rate and mass evolution

If we assume clusters to be spherically symmetric and moving on circular orbits in a static galactic potential, the evaporation mass-loss rate would be independent of both cluster mass and time (for example, Hénon 1961; Fall & Zhang 2001). In this simple picture, the mass-loss time-scale is always proportional to the current mass $M(t)$: $\tau_M \equiv -M/\dot{M} \propto M$. The total disruption time for a cluster with initial mass $M_{\rm i}$ (that is, the mass after stellar evolution, equation 2) is then $t_{\rm dis} \propto M_{\rm i}$. Escape of stars from a more realistic, anisotropic Roche volume around the cluster is delayed relative to a spherical one (Fukushige & Heggie 2000), which leads to a modified scaling of the form $t_{\rm dis} \propto M_{\rm i}^x$, with $x \simeq 2/3$ (Baumgardt 2001). For clusters without BHs, the mass-loss time-scale is also $\tau_M \propto M^x$, such that $\dot{M} = -M/\tau_M \propto M^{1-x}$ (Lamers, Baumgardt & Gieles 2010), and this scaling is often used in GC population studies (Lamers et al. 2005; Pfeffer et al. 2018; Choksi & Gnedin 2019a; Chen & Gnedin 2022). The positive correlation between $|\dot{M}|$ and M leads to a reduced mass-loss rate as the cluster approaches dissolution, that is, a gentle 'skiing' $M(t)$ evolution. This allows proportionally more low-mass clusters to survive until the present, leading to GCMF with a lower $M_{\rm TO}$ and a larger dispersion (Gieles 2009) than for the case of $x = 1$.

As we have shown in Section 2, the opposite regime of $x > 1$ is needed to mimic the effect of BHs and obtain steep 'jumping' $M(t)$ curves. However, with the simple scaling above this would result in a superlinear scaling between $t_{\rm dis}$ and $M_{\rm i}$, which is not found in these N-body simulations (see Fig. 1).

To unify a 'jumping' $M(t)$ with a sublinear dependence of $t_{\rm dis}$ on the initial mass, we write the mass-loss rate generally as

$$\dot{M} = \dot{M}_{\rm ref} \left(\frac{M}{M_{\rm i}}\right)^{1-y} \left(\frac{M_{\rm i}}{2\times10^5\,{\rm M}_\odot}\right)^{1-x} \frac{\Omega_{\rm tid}}{0.32\,{\rm Myr}^{-1}}, \tag{4}$$

where $x > 0$ is a parameter that controls the relation between $t_{\rm dis}$ and $M_{\rm i}$, and $y > 0$ is a parameter that controls the shape of $M(t)$. Here $\dot{M}_{\rm ref} < 0$ is the mass-loss rate at a fixed reference mass $M = 2\times10^5\,{\rm M}_\odot$ and the same reference $\Omega_{\rm tid}$ as in equation (1), that is, for $V_{\rm c} = 220\,{\rm km\,s}^{-1}$ and $R_{\rm eff} = 1\,{\rm kpc}$.

The functional form of equation (4) recovers the simple constant mass-loss rate for $x = y = 1$, and the 'skiing' $M(t)$ evolution with the delayed escape for $x = y \simeq 2/3$. However, x and y do not need to be the same, and equation (4) allows for both 'skiing' ($y < 1$) and 'jumping' ($y > 1$) evolution, for any scaling between $t_{\rm dis}$ and initial mass via the parameter x. A physical explanation for why x and y can be different lies in the history of disruption of clusters with the same M but different $M_{\rm i}$. Clusters with the same M, but different $M/M_{\rm i}$ have different mass fractions in BHs, which has a large effect on $\dot{M}$, as we have seen in Section 2. The different dependence of $\dot{M}$ on M and on $M_{\rm i}$ was proposed previously by Muratov & Gnedin (2010), based on independent arguments.

Integrating equation (4) over time we find the mass evolution

$$M(M_{\rm i}, \Omega_{\rm tid}, t) = M_{\rm i}\left(1 - \frac{t}{t_{\rm dis}(M_{\rm i}, \Omega_{\rm tid})}\right)^{1/y}, \tag{5}$$

for $t \le t_{\rm dis}(M_{\rm i}, \Omega_{\rm tid})$, where

$$t_{\rm dis} = 10\,{\rm Gyr}\,\frac{2/3}{y}\,\frac{30\,{\rm M}_\odot/{\rm Myr}}{|\dot{M}_{\rm ref}|}\,\frac{0.32\,{\rm Myr}^{-1}}{\Omega_{\rm tid}}\left(\frac{M_{\rm i}}{2\times10^5\,{\rm M}_\odot}\right)^x \tag{6}$$

is the total lifetime, that is, the time for the cluster mass to reach zero. The mass evolution of a cluster is defined by the four parameters x, y, $\dot{M}_{\rm ref}$, and $\Omega_{\rm tid}$.

In the left-hand panel of Fig. 6 we provide a synopsis of the mass evolution for different choices of these parameters. The values of x and y also affect the shape of the GCMF, which we discuss next.

3.2 Cluster mass function

We define the mass function ψ at time t as the number of GCs in the mass range $[M, M + {\rm d}M]$, that is, $\psi(M, \Omega_{\rm tid}, t) = {\rm d}N/{\rm d}M(M, \Omega_{\rm tid}, t)$. We can relate the mass function to the initial mass function $\psi_0(M_{\rm i})$ as

$$\psi(M, \Omega_{\rm tid}, t) = \psi_0[M_{\rm i}(M)]\left|\frac{\partial M_{\rm i}}{\partial M}(M, \Omega_{\rm tid}, t)\right|. \tag{7}$$

An analytical expression for the dependence $M_{\rm i}(M)$ (that is, the inverse of equation (5) for $M(M_{\rm i})$) can only be found for $x = y$

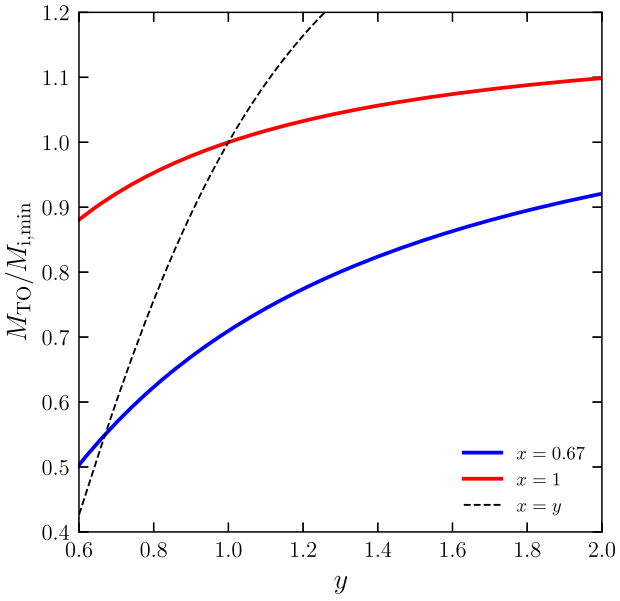

Figure 7. Turn-over mass $M_{\rm TO}$ of GCMF found numerically as a root of equation (10). The special case $x = y$ allows an analytic solution $M_{\rm TO}/M_{\rm i,\,min} = y^{1/y}$, shown by the dashed line.

(Gieles 2009), so we cannot write the mass function analytically in a general case $x \neq y$. Instead, we find $\partial M/\partial M_{\rm i}$ from the relation given by equation (5):

$$\frac{\partial \ln M}{\partial \ln M_{\rm i}} = 1 - \frac{x}{y} + \frac{x}{y}\left(\frac{M}{M_{\rm i}}\right)^{-y}. \tag{8}$$

By interpolating both this relation and $M(M_{\rm i})$ from equation (5), we find $M_{\rm i}(M)$ and $\partial M_{\rm i}/\partial M(M)$ to evaluate $\psi(M, \Omega_{\rm tid}, t)$.

The right-hand panel of Fig. 6 shows cluster mass functions evolved from the initial $\psi_0(M_{\rm i}) \propto M_{\rm i}^{-2}$, for different x and y. The logarithmic slope at low masses equals y, independent of x. This can be understood from considering the behaviour of $\partial M/\partial M_{\rm i}$ in the limit $M \ll M_{\rm i}$. In this case $\psi(M) \approx \psi_0(M_{\rm i})\,\frac{y}{x}\,(M/M_{\rm i})^{y-1}$ from equations (7) and (8). The logarithmic mass function $\psi(\log M) \propto M\psi(M) \propto M^y$. This power-law holds for masses below the minimum initial mass of clusters still surviving at the time of observation t, which can be found from equation (6):

$$M_{\rm i,min}(t) = 2\times10^5\,{\rm M}_\odot\left(\frac{y\,t\,|\dot{M}_{\rm ref}|}{2\times10^5\,{\rm M}_\odot}\right)^{1/x}. \tag{9}$$

For a given $\dot{M}_{\rm ref}t$, $M_{\rm TO}$ increases with increasing y. It can be calculated by setting ${\rm d}\ln\psi/{\rm d}\ln M = -1$ (Gieles 2009). For the initial $\psi_0(M_{\rm i}) \propto M_{\rm i}^{-2}$, after some algebra this gives the equation

$$x\,(x\mu - 1)\,(\mu - 1) = y, \tag{10}$$

where $\mu \equiv (M/M_{\rm i})^{-y}$. The root of this equation can be found numerically after we specify the relation between the initial cluster mass $M_{\rm i}$ and mass $M(t)$ after time t. That relation is given by equation (5) and can be expressed in both masses normalized by the minimum survival mass: $(M/M_{\rm i,\,min})^y = (M_{\rm i}/M_{\rm i,\,min})^y - (M_{\rm i}/M_{\rm i,\,min})^{y-x}$. The latter relation allows us to convert the turnover mass $M_{\rm TO}$ from equation (10) to a ratio $M_{\rm TO}/M_{\rm i,\,min}$. The numerical solutions for two cases $x = 1$ and $x = 0.67$ are shown in Fig. 7. They reproduce the values in Table 2 found from the evolved models.

There are three contributing factors to the higher $M_{\rm TO}$ when BHs are considered: (1) a larger y and (2) a higher $|\dot{M}_{\rm ref}|$ both increase $M_{\rm i,\,min}$ (equation 9) and (3) a larger y also results in a larger $M_{\rm TO}/M_{\rm i,\,min}$ (Fig. 7). As an illustration, we highlight the difference in $M_{\rm TO}$ for the parameter we will use in the next section: for the models with BHs ($y = 1.33$, $\dot{M}_{\rm ref} = -45\,{\rm M}_\odot\,{\rm Myr}^{-1}$), $M_{\rm TO}$ is a factor of $2.8 \times 1.45 \times 1.5 = 6.1$ higher than for the model without BHs ($y = 0.67$, $\dot{M}_{\rm ref} = -30\,{\rm M}_\odot\,{\rm Myr}^{-1}$), for the same $x = 0.67$ and the same $\Omega_{\rm tid}$.

Table 2. Turnover mass ($M_{\rm TO}$) and (logarithmic) mean ($\mu_{\log_{10}M}$) and dispersion ($\sigma_{\log_{10}M}$) for the mass functions shown in the right-hand panel of Fig. 6 and two additional values of y ($y = 1.75$ and $y = 2$).

x	y	$M_{\rm TO}/M_{\rm i,\,min}$	$\log_{10}M_{\rm TO}$	$\mu_{\log_{10}M}$ (dex)	$\sigma_{\log_{10}M}$
0.67	0.67	0.55	5.16	4.99	0.93
0.67	1.00	0.70	5.53	5.48	0.73
0.67	1.33	0.80	5.78	5.78	0.60
0.67	1.75	0.88	5.99	6.03	0.49
0.67	2.00	0.92	6.10	6.14	0.45
1.00	0.67	0.91	5.34	5.13	0.89
1.00	1.00	1.00	5.56	5.49	0.71
1.00	1.33	1.05	5.70	5.71	0.59
1.00	1.75	1.08	5.83	5.89	0.51
1.00	2.00	1.10	5.90	5.97	0.47

This simple model illustrates the general behaviour of the mass function under the new parametrization of the cluster mass-loss. While many of its assumptions are not valid in the real Milky Way galaxy, the analytical expressions help us understand main effects of including BHs on the accelerated disruption. In the next section, we present a more realistic model that accounts for the orbit distribution in the galaxy and matches the observations of Galactic GCs.

4 POPULATION MODEL

In this section, we present a detailed model for the evolution of a GC population in a Milky Way-like galaxy. Our approach is similar to that of Fall & Zhang (2001) who start with a distribution function (DF) and then evolve the GCMF due to various disruption mechanisms which all have a dependence on the orbit. Their DF depends on isolating integrals (specific energy and angular momentum), but we express the DF directly in terms of mass, position, and velocity. We adopt an SIS for the Galaxy, which has a potential

$$\Phi(R) = V_{\rm c}^2\,\ln(R), \tag{11}$$

and we assume the circular velocity of $V_{\rm c} = 220\,{\rm km\,s^{-1}}$.

4.1 Initial conditions

We write the initial DF, that is, the phase-space density of clusters in the galaxy, as a function of $M_{\rm i}$, galactic position ($\boldsymbol{R}$), and galactic velocity ($\boldsymbol{V}$)

$$f_0(M_{\rm i}, \boldsymbol{R}, \boldsymbol{V}) = \psi_0(M_{\rm i})\,n_0(R)\,F(R, V_{\rm r}, V_{\rm t}). \tag{12}$$

Here $\psi_0(M_{\rm i})$ describes the initial GCMF, $n_0(R)$ the radial number density profile, and $F(R, V_{\rm r}, V_{\rm t})$ the radius-dependent velocity distribution. Because we will consider the effect of radially biased velocity anisotropy for the GC orbits, we define the velocity distribution in terms of the radial velocity ($V_{\rm r}$) and the tangential velocity ($V_{\rm t}$). The phase-space density is normalized such that $\int_0^\infty {\rm d}^3\boldsymbol{V}\int_{R_{\rm lo}}^{R_{\rm up}}{\rm d}^3\boldsymbol{R}\int_{M_{\rm lo}}^{M_{\rm up}}{\rm d}M_{\rm i}\,f_0(M_{\rm i}, \boldsymbol{R}, \boldsymbol{V}) = 1$, where we adopt the following boundary values: $R_{\rm lo} = 1\,{\rm kpc}$, $R_{\rm up} = 100\,{\rm kpc}$, $M_{\rm lo} = 10^4\,{\rm M}_\odot$, and $M_{\rm up} = 10^8\,{\rm M}_\odot$. We discuss the effect of varying $M_{\rm lo}$

in Section 6 and next we discuss the functional forms for each contribution to f_0.

For $\psi_0(M_\mathrm{i})$ we adopt a power law with an exponential truncation

$$\psi_0(M_\mathrm{i}) \propto M_\mathrm{i}^{-\alpha} \exp\left(-\frac{M_\mathrm{i}}{M_\mathrm{c}}\right), \tag{13}$$

where M_c is the truncation mass, which we take either $M_\mathrm{c} \to \infty$ (the 'power-law' model considered in the previous section), or $M_\mathrm{c} \simeq 10^6\,\mathrm{M}_\odot$ (the 'Schechter' model). We fix the power-law index $\alpha = 2$ in all models, as it is a common value found for young massive clusters in nearby galaxies. Although it cannot be ruled out that GCs formed with a more peaked initial GCMF, here we adopt the hypothesis that massive star clusters form with a universal GCMF at all redshifts. The constant of proportionality is found from the requirement that $\int_{M_\mathrm{lo}}^{M_\mathrm{up}} \psi_0\,\mathrm{d}M_\mathrm{i} = 1$.

The velocity distribution is assumed to be Gaussian in all three components, such that

$$F(R, V_\mathrm{r}, V_\mathrm{t}) = \frac{\exp\left[-V_\mathrm{r}^2/(2\sigma_\mathrm{r}^2(R))\right]}{\sqrt{2\pi}\sigma_\mathrm{r}(R)} \frac{\exp\left[-V_\mathrm{t}^2/\sigma_\mathrm{t}^2(R)\right]}{\pi\sigma_\mathrm{t}^2(R)}. \tag{14}$$

Here $\sigma_\mathrm{t}(R) = \langle V_\mathrm{t}^2(R)\rangle^{1/2}$ is the root-mean-square tangential velocity at radius R and $\sigma_\mathrm{r}(R) = \langle V_\mathrm{r}^2(R)\rangle^{1/2}$ is the root-mean-square radial velocity at radius R. For an isotropic velocity distribution $\sigma_\mathrm{t}^2 = 2\sigma_\mathrm{r}^2$. It satisfies $\int \mathrm{d}^3\boldsymbol{V}\, F(R, V_\mathrm{r}, V_\mathrm{t}) = 2\pi \int_0^\infty V_\mathrm{t}\mathrm{d}V_\mathrm{t} \int_{-\infty}^{\infty} \mathrm{d}V_\mathrm{r}\, F(R, V_\mathrm{r}, V_\mathrm{t}) = 1$. We adopt an anisotropy profile of the GC system of the form

$$\begin{aligned}\beta(R) &\equiv 1 - \frac{\sigma_\mathrm{t}^2(R)}{2\sigma_\mathrm{r}^2(R)} \\ &= \frac{1}{1+(R_\mathrm{ani}/R)^\delta},\end{aligned} \tag{15}$$

where $\delta > 0$. This profile results in isotropy for $R \ll R_\mathrm{ani}$ and radial orbits for $R \gg R_\mathrm{ani}$, with the parameter δ determining how quickly $\beta(R)$ rises. DFs that include radial anisotropy with an $\exp(-J^2)$ term, where J is the specific angular momentum, result in a $\beta(R)$ profile as in equation (15) with $\delta = 2$ (Eddington 1915; Michie 1963; Osipkov 1979; Merritt 1985). The present-day anisotropy profile of Milky Way GCs as derived from line-of-sight velocities and *Gaia* proper motions (Vasiliev 2019) is better described by $\delta \simeq 1$. Here we will vary δ, together with R_ani, to match the β profile of the observed clusters (see the description of Model (3) in Section 4.3).

To find $\sigma_\mathrm{r}(R)$ and $\sigma_\mathrm{t}(R)$ we need to define $n_0(R)$ and solve the radially anisotropic Jeans equation (equation 4.215 in Binney & Tremaine 2008) with GCs as tracer particles in the Galactic potential (equation 11). A convenient choice for $n_0(R)$ is

$$n_0(R) \propto \frac{R^{-\gamma}}{\left[1+(R/R_\mathrm{ani})^\delta\right]^{2/\delta}}, \tag{16}$$

because it results in a constant radial dispersion $\sigma_\mathrm{r} = V_\mathrm{c}/\gamma^{1/2}$. For $R \ll R_\mathrm{ani}$, the profile is $n_0(R) \propto R^{-\gamma}$, and for $R \gg R_\mathrm{ani}$ it is $n_0(R) \propto R^{-\gamma-2}$. Fully isotropic models ($R_\mathrm{ani} \to \infty$) have a single power-law $n_0(R) \propto R^{-\gamma}$. The constant of proportionality is found from the requirement that $\int n_0(R)\,\mathrm{d}^3\boldsymbol{R} = 4\pi \int_{R_\mathrm{lo}}^{R_\mathrm{up}} n_0(R) R^2 \mathrm{d}R = 1$. Combined with the expression for $\beta(R)$, we find that $\sigma_\mathrm{t}^2(R) = 2\sigma_\mathrm{r}^2[1+(R/R_\mathrm{ani})^\delta]^{-1}$. We now have a fully analytic form for $f_0(M_\mathrm{i}, \boldsymbol{R}, \boldsymbol{V})$ (equation 12) and describe next how we evolve it to the present age of GCs.

4.2 Evolving the GC population

To evolve the mass function as a function of R, we need to obtain the effective tidal field strength $\Omega_\mathrm{tid} = V_\mathrm{c}[R_\mathrm{p}(1+\epsilon)]^{-1}$ from V_r, V_t and R. For the case of a SIS, $R_\mathrm{p}(R, V_\mathrm{r}, V_\mathrm{t})$ and $R_\mathrm{a}(R, V_\mathrm{r}, V_\mathrm{t})$ are the radii where $V_\mathrm{r} = 0$, which are roots that need to solved numerically from the orbital energy and angular momentum (see, for example, Section 2.1 of van den Bosch et al. 1999), which then provides $\epsilon = (R_\mathrm{a} - R_\mathrm{p})/(R_\mathrm{a} + R_\mathrm{p})$. Because of the scale-free nature of the SIS, we do this once for 100 eccentricities between 0 and 1 and then use interpolation to find $\Omega_\mathrm{tid}(R, V_\mathrm{r}, V_\mathrm{t})$.

We first find the present-day phase-space density $f(M, \boldsymbol{R}, \boldsymbol{V}, t) = \psi(M, \Omega_\mathrm{tid}, t)\, n_0(R)\, F(R, V_\mathrm{r}, V_\mathrm{t})$, where $\psi(M, \Omega_\mathrm{tid}, t)$ is given by equation (7) and we recall that $\Omega_\mathrm{tid} = \Omega_\mathrm{tid}(R, V_\mathrm{r}, V_\mathrm{t})$. We then integrate over all velocities to obtain the mass function as a function of R

$$\begin{aligned}\psi(M, R, t) &= \int f(M, \boldsymbol{R}, \boldsymbol{V}, t)\,\mathrm{d}^3\boldsymbol{V}, \\ &= 4\pi R^2 n_0(R) \int \mathrm{d}^3\boldsymbol{V}\, \psi(M, \Omega_\mathrm{tid}, t) F(R, V_\mathrm{r}, V_\mathrm{t}).\end{aligned} \tag{17}$$

The fraction of surviving clusters is then given by $f_\mathrm{surv} = \int_{M_\mathrm{lo}}^{M_\mathrm{up}} \mathrm{d}M \int_{R_\mathrm{lo}}^{R_\mathrm{up}} \mathrm{d}R\, \psi(M, R, t)$. To compare the model to the observations we multiply $\psi(M, R, t)$ by $N_\mathrm{GC}/f_\mathrm{surv}$, with $N_\mathrm{GC} = 156$ being the total number of GCs in the Milky Way for which a luminosity and R are available (Harris 1996, 2010). We summarize the various definitions of the DF and the mass function in Table 3.

Table 3. Overview of the functions used in the population model.

Function	Definition	Description
$f_0(M_\mathrm{i}, \boldsymbol{R}, \boldsymbol{V})$	$\frac{\mathrm{d}^7N}{\mathrm{d}M_\mathrm{i}\,\mathrm{d}^3\boldsymbol{R}\,\mathrm{d}^3\boldsymbol{V}}$	Initial DF (after stellar evolution)
$f(M, \boldsymbol{R}, \boldsymbol{V}, t)$	$\frac{\mathrm{d}^7N}{\mathrm{d}M\,\mathrm{d}^3\boldsymbol{R}\,\mathrm{d}^3\boldsymbol{V}}$	DF (after evaporation)
$\psi_0(M_\mathrm{i})$	$\frac{\mathrm{d}N}{\mathrm{d}M_\mathrm{i}}$	Initial GCMF (after stellar evolution)
$\psi(M, \Omega_\mathrm{tid}, t)$	$\frac{\mathrm{d}N}{\mathrm{d}M}$	Evolved GCMF for a single Ω_tid
$\psi(M, R, t)$	$\frac{\mathrm{d}^2N}{\mathrm{d}M\mathrm{d}R}$	Evolved GCMF at R (equation 17)
$\psi(M, t)$	$\frac{\mathrm{d}N}{\mathrm{d}M}$	$\psi(M, R, t)$ integrated over R

Table 4. Overview of the nine parameters of the GC population model.

Description	Parameter
Galaxy:	$V_\mathrm{c} = 220\,\mathrm{km\,s^{-1}}$
Mass-loss:	$\dot{M}_\mathrm{ref} = -30$ or $-45\,\mathrm{M}_\odot\,\mathrm{Myr}^{-1}$
	$x = 2/3$
	$y = 2/3$ or $4/3$
Initial GCMF:	$\alpha = 2$
	$M_\mathrm{c} = 10^6\,\mathrm{M}_\odot$ or ∞
$n_0(R)$:	$\gamma = 3.5$ or 4.5
$\beta(R)$ and $n_0(R)$:	$R_\mathrm{ani} = 5\,\mathrm{kpc}$ or ∞
	$\delta = 0.5$

4.3 Model parameters

In Table 4, we summarize the nine parameters of the GC population model we described above, including the adopted values. The parameters with a single mentioned value are fixed in all models. For some parameters we adopt two values, in order to study their effect on the resulting GCMF. We use these parameters to solve eight models summarized in Table 5 and described below.

(i) **Model (1), no BHs:** This model serves as a starting point and defines the magnitude of the GCMF problem by considering the simplest case: an isotropic velocity distribution $\beta(R) = 0$. Integrating

Table 5. Different population models shown in Figs 9 and 10. In all models $x = 0.67$, $\alpha = 2$, $V_c = 220\,\mathrm{km\,s^{-1}}$.

Model	Description					Parameters					
	Name	Anisotropy	ICMF	[Fe/H] gradient	Past tidal	$R_{\rm ani}$ (kpc)	δ	γ	M_c ($M_\odot$)	$\dot{M}_{\rm ref}$ ($M_\odot$/Myr)	y
(1)	no BHs	no	power law	no	no	∞	–	4.5	∞	−30	0.67
(2)	BHs	no	power law	no	no	∞	–	4.5	∞	−45	1.33
(3)	BHs + A (Anisotropy)	yes	power law	no	no	5	0.5	3.5	∞	−45	1.33
(4)	BHs + S (Schechter)	no	Schechter	no	no	∞	–	4.5	10^6	−45	1.33
(5)	BHs + F ([Fe/H] gradient)	no	power law	yes	no	∞	–	4.5	∞	−45	1.33
(6)	BHs + P (Past evolution)	no	power law	no	yes	∞	–	4.5	∞	−45	1.33
(7)	BHs + A + S	yes	Schechter	no	no	5	0.5	3.5	10^6	-45	1.33
(8)	BHs + A+S+F + P	yes	Schechter	yes	yes	5	0.5	3.5	10^6	−45	1.33

over all orbits, we find that the average mass-loss rate at R is a factor of ∼3.1 higher than that of the circular orbit at that R. The initial GCMF is a power law and for the mass-loss parameters we adopt the values found in models without BHs (equation 1), that is, $\dot{M}_{\rm ref} = -30\,\mathrm{M_\odot\,Myr^{-1}}$ and $x = y = 2/3$. This model roughly describes the contribution of evaporation in the hierarchical models of Pfeffer et al. (2018) and Choksi & Gnedin (2019b) at low redshift.

(ii) **Model (2), BHs:** Here we study the effect of higher $|\dot{M}|$ and y due to BHs. We base the values on the N-body models with $\mathcal{R}_0 \simeq 30$ from Section 2, which can be described by $\dot{M}_{\rm ref} = -45\,\mathrm{M_\odot\,Myr^{-1}}$ and $y = 4/3$. In these models the effect of BHs is moderate compared to the two sets of N-body models with lower $\mathcal{R}_0$ shown in Fig. 2, which have higher $|\dot{M}_{\rm ref}|$ and y. We assume the same $\mathcal{R}_0 \simeq 30$ and $r_{\rm h,\,0}/r_{\rm J,\,eff} \simeq 0.05$ (Table 1) for all clusters. This is of course not realistic, because real clusters have a spread in these parameters and the orbits evolve in time, but it serves as an approximation for the average filling factor of clusters. We discuss this point more in Section 6.4.

In the next four models, we add a single physical effect to the BHs, which each reduce the gradient of $M_{\rm TO}(R)$:

(iii) **Model (3), BHs + A (Anisotropy):** Here we add radially biased anisotropy by choosing $R_{\rm ani} = 5\,\mathrm{kpc}$ with a relatively slowly rising $\beta(R)$ ($\delta = 0.5$). These values were chosen such that in the final Model (8) the anisotropy profile of the surviving clusters is similar to the observed profile. Anisotropy increases $|\dot{M}|$ at large R, thereby reducing the gradient of $M_{\rm TO}(R)$.

(iv) **Model (4), BHs + S (Schechter):** Here we add a Schechter truncation mass of $M_c = 10^6\,\mathrm{M_\odot}$ as found by Jordán et al. (2007) from fits of 'evolved Schechter functions' of the Milky Way GCMF. If the amount of mass lost is comparable to M_c, the turnover mass only increases slowly for any additional mass-loss (Jordán et al. 2007; Gieles 2009), so this truncation mass reduces the gradient of $M_{\rm TO}(R)$ in the inner galaxy.

(v) **Model (5), BHs + F ([Fe/H] gradient):** Here we consider the effect of the GC metallicity gradient in the galaxy. In Section 2.3 we showed that more metal-rich clusters ([Fe/H] $\gtrsim -0.5$), with the same initial density ($\mathcal{R}_0 \simeq 30$) eject all BHs and evolve similarly to clusters without BHs. This means that in the inner galaxy, $|\dot{M}|$ is lower than in Model (2). Fig. 8 shows that in the range $0 \lesssim \log_{10}(R/\mathrm{kpc}) \lesssim 1$, [Fe/H] of Milky Way GCs decreases from approximately −0.5 to −1.5. We mimic the effect of a [Fe/H]-gradient by adopting R-dependent relations for $\dot{M}$ and y for $R \leq 10\,\mathrm{kpc}$:

$$\dot{M}_{\rm ref}(R) = -30\,\mathrm{M_\odot\,Myr^{-1}}\,\left(1 + \tfrac{1}{2}\log_{10}(R)\right) \quad (18)$$

$$y(R) = \tfrac{2}{3} + \tfrac{2}{3}\log_{10}(R). \quad (19)$$

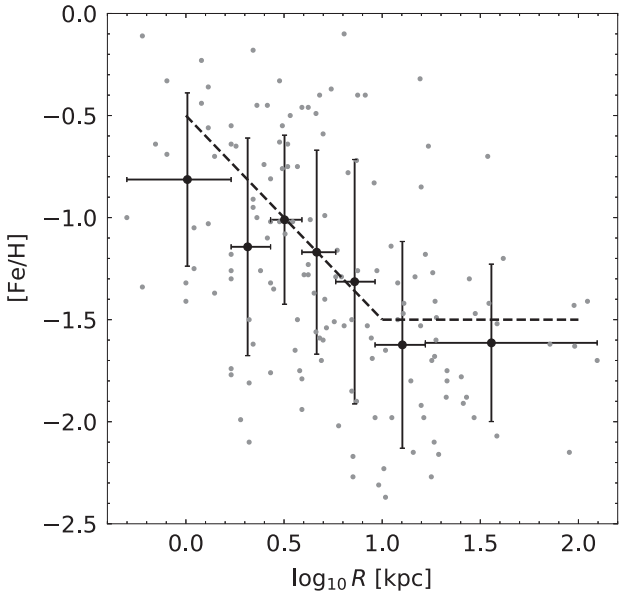

Figure 8. Metallicities of Milky Way GCs, as individual points (grey symbols) and in radial bins. The dashed line shows a simple analytic approximation, which for $R < 10\,\mathrm{kpc}$ has a logarithmic slope of −1: [Fe/H]$(R) = -0.5 - \log_{10}R$.

At $R = 10\,\mathrm{kpc}$ these relations result in the same values as Model (2) ($\dot{M}_{\rm ref} = -45\,\mathrm{M_\odot\,Myr^{-1}}$, $y = 4/3$), and at $R = 1\,\mathrm{kpc}$ they give the values found for clusters without BHs ($\dot{M}_{\rm ref} = -30\,\mathrm{M_\odot\,Myr^{-1}}$, $y = 2/3$, equation 1), thereby reducing the effect of BHs on $\dot{M}$ and therefore the gradient of $M_{\rm TO}(R)$.

(vi) **Model (6), BHs + P (Past evolution):** Here we include an approximate correction for the past tidal evolution of clusters from a full hierarchical model, described in Appendix A. As a result we multiply $\dot{M}$ by $(R_{\rm eff}/4)^{1/2}$ at $R_{\rm eff} > 4$ kpc.

(vii) **Model (7), BHs + A+S:** In this model, we combine the effect of anisotropy and the Schechter cutoff mass.

(viii) **Model (8), BHs + A+S+F+P:** Here we include all effects described in Models (2)–(6). This model represents a realistic way of modelling cluster evolution.

5 RESULTS OF POPULATION MODELS

In this section, we discuss the results of the eight models summarized in Table 5. We compare the models to 156 GCs with luminosities and R in the Harris catalogue (Harris 1996, 2010). We adopt a mass-to-light ratio $M/L_V = 1.8$ from Baumgardt, Sollima & Hilker (2020)

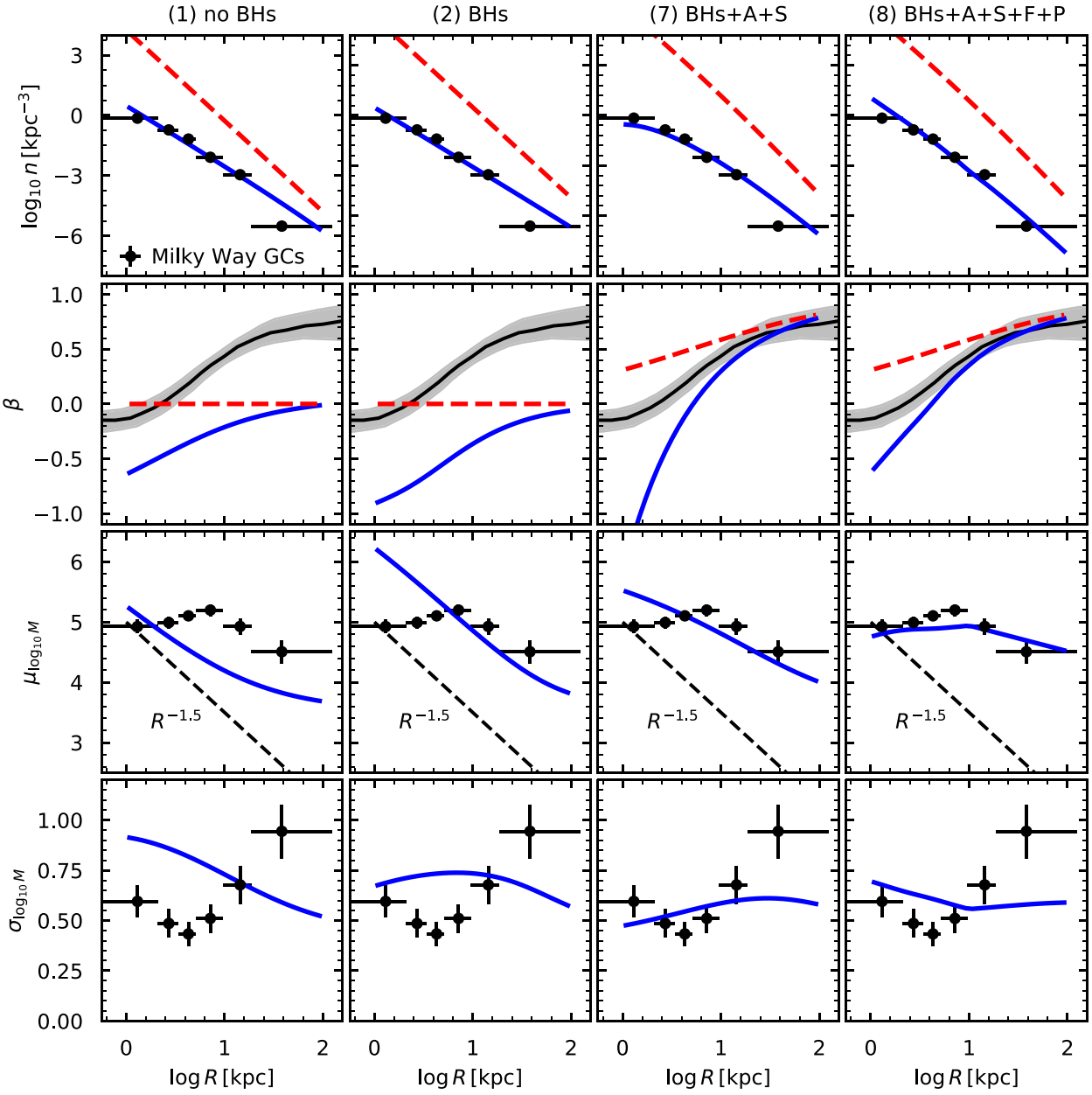

Figure 9. Different ingredients in the GCMF model compared. **A** stands for Anisotropy, **S** for Schechter, **F** for [Fe/H] gradient, **P** for past evolution. Red dashed lines in the top panels are the initial distributions, solid blue lines are the final distributions, black data points with error bars, and black solid line with grey shaded regions are observations.

and quantify the shape of the GCMF at different radii by $\mu_{\log_{10}M} \equiv \langle \log_{10}(M/\mathrm{M}_\odot)\rangle$ and the dispersion of the logarithmic mass distribution, $\sigma_{\log_{10}M}$, for $M \geq 10^3\,\mathrm{M}_\odot$. The mean of the logarithm of mass is a reasonable approximation to $M_{\rm TO}$ (see Table 2). For the anisotropy, we consider clusters with $M > 10^5\,\mathrm{M}_\odot$ and compare to the results from *Gaia* DR2 by Vasiliev (2019). All models are displayed in Figs 9 and 10.

5.1 Global properties

(i) **Model (1), no BHs:** This model quantifies the magnitude of the GCMF problem: $\mu_{\log_{10}M}$ reproduces the data only at $R \simeq 1$ kpc and then it declines as $R^{-1/x} = R^{-3/2}$, which is a known scaling for $M_{\rm TO}(R)$ for an initial GCMF that is a power law with logarithmic slope of -2 and $t_{\rm dis} \propto M_0^x$ (Gieles 2009). At $R \simeq 100$ kpc, $\mu_{\log_{10}M}$ is a factor of $\sim$10 too low. The model also underestimates β at all radii, because preferentially radial orbits are removed at small radii while the initial β is too low at large R and evolves very little. The power law initial GCMF and low value for x also result in a very wide GCMF ($\sigma_{\log_{10}M} \simeq 0.75$) compared to the observed width ($\sigma_{\log_{10}M} \simeq 0.5$). The model width is also decreasing with R, while the data show an increase. From this it is clear that evaporation of clusters without BHs is not able to explain the shape of the GCMF. Adding radial anisotropy within the constraints of the observed $\beta(R)$ increases $\mu_{\log_{10}M}$ by only ~ 0.3 dex at large R (not shown), and is therefore not sufficient.

(ii) **Model (2), BHs:** Here we only change $\dot{M}_{\rm ref}$ and y to mimic the effect of BHs on $\dot{M}$. This model almost reproduces $\mu_{\log_{10}M}$ at

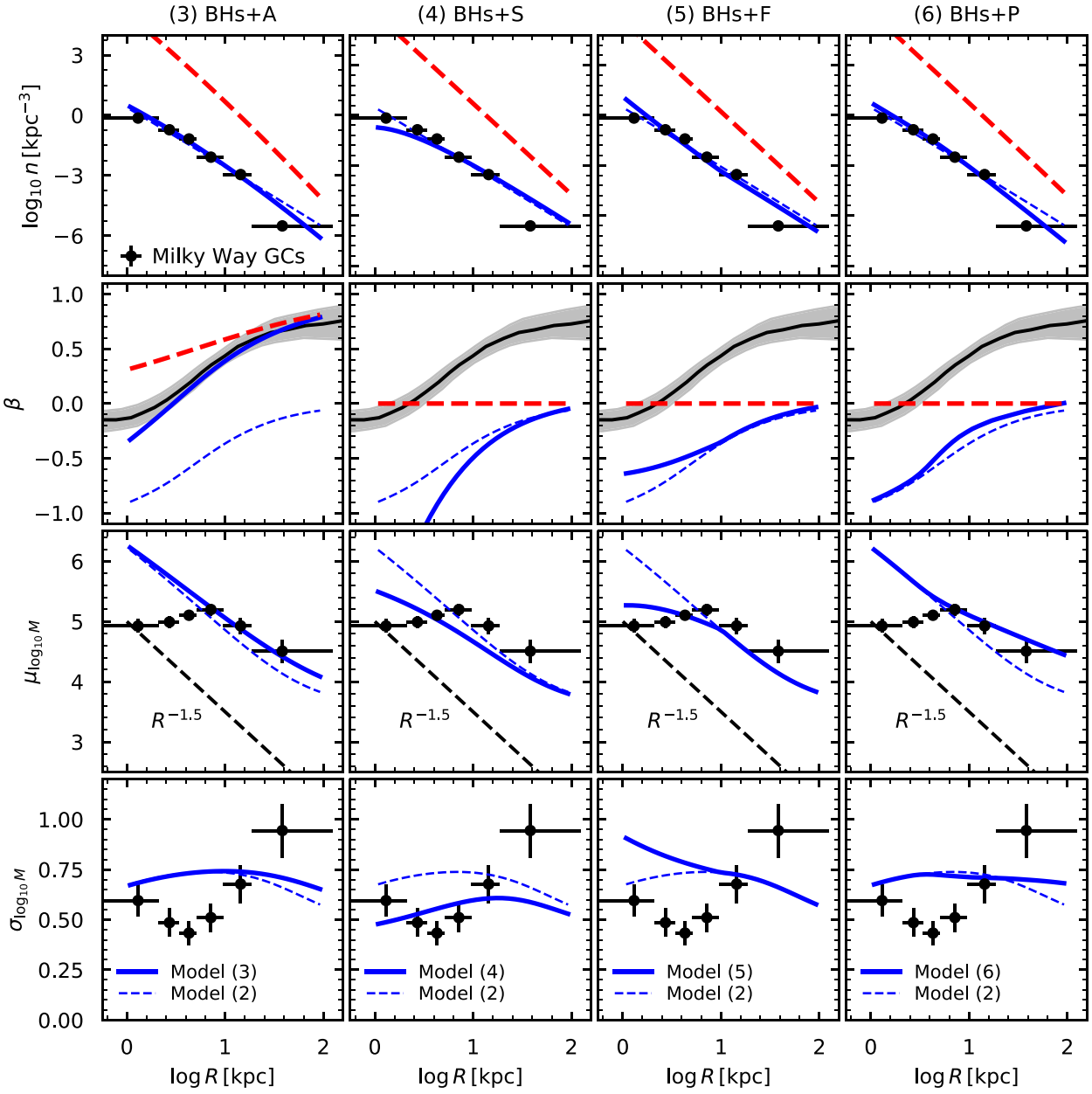

Figure 10. As Fig. 9, but now for Models (3–6), showing in each column the effect BHs and one additional ingredient (**A** = Anisotropy; **S** = Schechter; **F** = [Fe/H] gradient; **P** = past evolution).

$R \gtrsim 10$ kpc, suggesting that the effect of BHs alleviates a large part of the GCMF problem. This model gives rise to the same scaling $\mu_{\log_{10}M} \propto R^{-3/2}$, so when adding BHs the problem is that $M_{\rm TO}$ is too high in the inner galaxy ($\lesssim 10$ kpc).

Models (3)–(6) present four additional ingredients that all reduce the (logarithmic) slope of the $\mu_{\log_{10}M}(R)$ relation. The individual effects can be seen in Fig. 10.

(iii) **Model (3), BHs + A (Anisotropy):** In this model we add radial anisotropy to the effect of BHs. We find that for $R_{\rm ani} = 5$ kpc and $\delta = 0.5$ the observed $\beta(R)$ profile is well reproduced. Compared to Model (2), radially biased anisotropy increases $\mu_{\log_{10}M}$ by ~ 0.3 dex at $R \gtrsim 10$ kpc, improving agreement with the observations at those large radii.

(iv) **Model (4), BHs + S (Schechter):** The addition of the exponential truncation in the initial GCMF reduces $\mu_{\log_{10}M}$ at $R \simeq 1$ kpc by nearly an order of magnitude. It also narrows the GCMF to approximately the correct width $\sigma_{\log_{10}M} \simeq 0.5$.

(v) **Model (5), BHs + F ([Fe/H] gradient):** The inclusion of the metallicity gradient reduces $\dot{M}$ mainly at $R \lesssim 3$ kpc, decreasing $\mu_{\log_{10}M}$ and increasing $\sigma_{\log_{10}M}$ there.

(vi) **Model (6), BHs + P (Past evolution):** The past tidal evolution reduces the gradient of $\mu_{\log_{10}M}(R)$ from $R^{-3/2}$ to approximately $R^{-0.9}$.

(vii) **Model (7), BHs + A+S:** The combined effect of anisotropy and a Schechter cutoff reduces the decline of $\mu_{\log_{10}M}$, but still the disruption at small (large) R is overestimated (underestimated) slightly.

(viii) **Model (8), BHs + A+S+F+P:** This model combines all four effects (A+S+F + P) in addition to the BHs. We note that because of the relatively small $\delta = 0.5$, the logarithmic slope of the initial density profile is steeper than -3.5 at all radii: it reduces from -4.1 at 1 kpc to -5.1 at 100 kpc, and at $R_{\rm ani} = 5$ kpc the slope is -4.5. In this model the gradient of $\mu_{\log_{10}M}$ is substantially reduced compared to the power-law Model (2) and closely follows the observed $\mu_{\log_{10}M}(R)$. This model reproduces well the number density profile, the anisotropy profile, and the shape of the GCMF at different Galactic radii. The model orbits are a bit too tangentially biased ($\beta < 0$) at $R \lesssim 3$ kpc. We interpret this as an artefact of our assumption of a static potential in which orbits do not isotropize due to interactions with the galactic bar, other GCs and infalling satellite galaxies. We expect that these effects in a real galaxy would lead to an isotropic velocity distribution of inner GCs.

5.2 Mass function

In Figs 9 and 10, we used the (logarithmic) mean and dispersion of the GCMF as measures of the GCMF shape. It is also instructive to look at the full GCMF in different Galactocentric radius bins. The GCMF for all GCs is obtained from $\psi(M,t) = \int_{R_{\rm lo}}^{R_{\rm up}} \psi(M,R,t){\rm d}R$, with $\psi(M,R,t)$ from equation (17). The integration boundaries can be changed to obtain the GCMF in different radial intervals. Fig. 11 shows the total mass function for Model (8) and in three radial bins. For the observational data, we used radial bins with equal number of GCs (52) and the histograms were constructed with equal number of GCs: eight in the total sample and four in the three R bins. For the model we also defined the radial bins to contain exactly 1/3 of the total number of GCs, which results in slightly different bin edges than for the observations because the number density profile of the model is slightly steeper. We do this because we are mostly interested in comparing the shapes, rather then the vertical scaling. The model GCMF shows good resemblance to the observed one, reproducing the slightly lower $M_{\rm TO}$ and larger width at high R.

6 DISCUSSION

6.1 Contribution to field stars

The vast majority of GCs in our model do not survive, so an important check is to compare the contribution of dissolved clusters to the field stars. The total mass lost from star clusters with $M_{\rm i} > 10^4\,{\rm M}_\odot$ in Model (8) is $\Delta M \simeq 5\times10^8\,{\rm M}_\odot$ and here we discuss the implications. The total mass of the Galactic halo is $1.4\times10^9\,{\rm M}_\odot$ (Deason, Belokurov & Sanders 2019), so if all lost mass ended up in the halo, then roughly one-third of the stellar halo would be made out of disrupted star clusters. If we adopt a lower limit of $10^2\,{\rm M}_\odot$ this fraction approximately doubles. This appears in tension with the results of Deason, Belokurov & Weisz (2015) who find that the ratio of the number of blue stragglers over blue horizontal branch stars in the halo is more similar to that in dwarf galaxies than in surviving GCs. It must be noted that this ratio in low-mass clusters is closer to what is found in the field than in massive clusters (Deason et al. 2015). Also, the contribution of stars from GCs to the halo is more important in the inner Galaxy where tides are stronger. To quantify this, we determined the radial density profile of mass lost from GCs, $\rho_\star(R)$. We find $\rho_\star(R)$ by subtracting the mass in surviving clusters from the initial mass of the GC population as a function of R

$$\rho_\star(R) = \frac{N_{\rm GC}}{f_{\rm surv}}\left[\int_{M_{\rm lo}}^{M_{\rm up}} M_{\rm i}\psi_0(M_{\rm i})n_0(R)\,{\rm d}M_{\rm i} - \frac{1}{4\pi R^2}\int_{M_{\rm lo}}^{M_{\rm up}} M\psi(M,R,12\,{\rm Gyr})\,{\rm d}M\right] \quad (20)$$

Here $\psi(M,R,t)$ is the present-day GCMF (equation 17) and $\psi_0(M_{\rm i})$ and $n_0(R)$ are the initial GCMF and initial number density profile from equation (12). Note that this expression is only approximate, because it assumes that the mass is lost at R while in reality the escaped stars follow a distribution between the pericentre and apocentre distance of the orbit. Nevertheless, equation (20) provides a useful estimate that can be compared to observational data.

In Fig. 12 we compare $\rho_\star(R)$ to the stellar halo density profile from APOGEE (Horta et al. 2021). Because the APOGEE data only include stars with $-2.5 <$ [Fe/H] < -1, we multiply our model predictions by a correction factor $f_{\rm MP}(R) \leq 1$ that approximates the fraction of GCs with [Fe/H] < -1 as a function of R. From the Harris catalogue, we find that this fraction for GCs today is well described by $f_{\rm MP} = [1.5 + \log_{10}(R/{\rm kpc})]/3.5$ for $1 < R/{\rm kpc} < 100$. It increases from $f_{\rm MP}(1\,{\rm kpc}) \simeq 0.4$ to $f_{\rm MP}(100\,{\rm kpc}) = 1$. At the smallest R in the APOGEE data (1.5 kpc), the contribution of mass lost from GCs to the total halo mass is nearly 70 per cent. This seems extreme, but we note that at 15 kpc the fraction drops to ~ 10 per cent, so there is no tension with the conclusion of Deason et al. (2015), because their sample considered stars at $R \gtrsim 10$ kpc. Also, we note that GC may form in a disc-like configuration (Kravtsov & Gnedin 2005; Meng & Gnedin 2021) and later scatter into the halo. The fast disrupting low-mass clusters therefore may contribute more to the thick disc and/or the bulge, which are 5–10 times more massive than the halo, respectively.

6.2 Nitrogen-rich stars

Most mass is lost from now-disrupted low-mass clusters, and it is challenging to identify these stars as having originated from a cluster, because their streams will have phase mixed long ago. However, massive clusters ($\gtrsim 10^5\,{\rm M}_\odot$) have anomalous light-element abundances, manifesting as anticorrelations in N–C, Na–O, and sometimes Al–Mg (Bastian & Lardo 2018), and these chemical imprints are preserved when stars are lost from the cluster. Stars with such abundances have also been found in the (inner) halo (Martell et al. 2016; Schiavon et al. 2017; Horta et al. 2021; Belokurov & Kravtsov 2022). As another test, we compare $\rho_\star(R)$ of stars originating from massive clusters to the density profile of N-rich stars found in APOGEE by Horta et al. (2021). We assume that 2/3 of GC stars with $M_{\rm i} \geq 10^5\,{\rm M}_\odot$ ($\geq \mu_{\rm sev}10^5\,{\rm M}_\odot$ after stellar mass-loss) have anomalous abundances. In Fig. 12 we show that the predicted $\rho_\star(R)$ of N-rich stars matches the observed profile from APOGEE very well. This suggests that these N-rich halo stars have a GC origin. A follow-up test is to look for clustering of stars in energy and angular momentum space (or action-angles), because a GC origin predicts that the N-rich stars are more clustered than the rest of the halo stars, as they originated from more massive clusters, which disrupted more recently. Additionally, at a given R, most mass is lost from the GCs that have the most radial orbits, so we predict that the N-rich stars are preferentially on radial orbit ($\beta \sim 0.5$). Indeed, N-rich stars are on highly eccentric orbits (Fernández-Trincado et al. 2019), but not more eccentric than normal metal-poor stars (Tang et al. 2020).

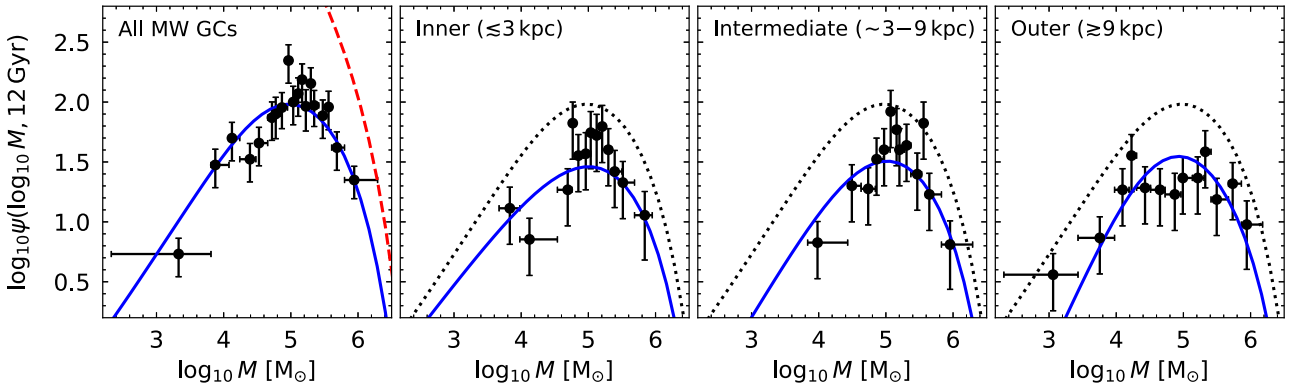

Figure 11. Comparison between the GCMF of Model (8) and the observed MF of Galactic GCs from the Harris catalogue for all 156 GCs (left-hand panel) and at three Galactocentric radius bins containing equal number of GCs (52) each. The dashed (red) line shows the initial GCMFs. The dotted (black) lines in the three right-hand panels show the total GCMF model for reference.

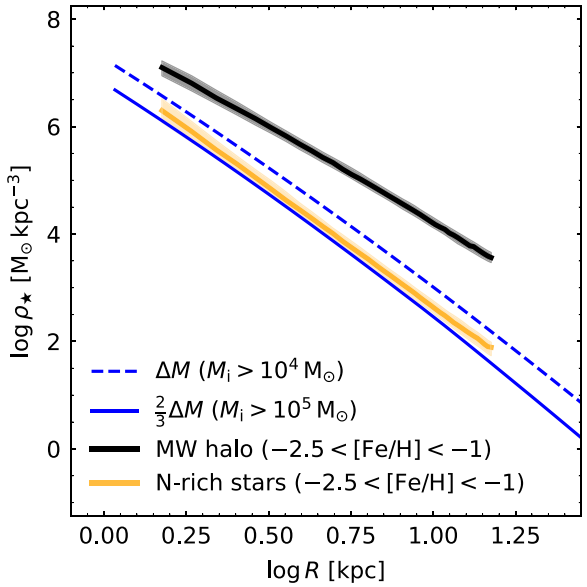

Figure 12. Comparison between the density profiles of stellar mass stripped from GCs (ΔM) and the N-rich stars in the Galactic halo (Horta et al. 2021), which are expected to originate from massive GCs. We compare this to ΔM from GCs with initial masses $M_{\rm i} > 10^5\,{\rm M}_\odot$ and assume that 2/3 of their stars are N rich. The APOGEE data do not include stars with [Fe/H] ≥ -1, so we multiply our model prediction (grey dashed lines) by an approximation for the fraction of metal-poor GCs as a function of radius $f(R) = (1.5 + \log(R))/3.5$, for $R \leq 100$ kpc, which describes the present-day fraction of Milky Way GCs with [Fe/H] < -1.

6.3 Specific frequency as a function of metallicity

Another aspect to consider is GC disruption as a function of [Fe/H]. In our model, metal-poor GCs have a higher $|\dot{M}|$ than metal-rich GCs for the same $\Omega_{\rm tid}$. If we assume that all star formation happens in clusters (that is, the cluster formation efficiency is 100 per cent), then in the absence of a metallicity gradient and for a metallicity-independent $\beta(R)$ we expect a lower GC specific frequency for metal-poor GCs. Here we define specific frequency as the fraction of the stellar mass that is in clusters at present. However, the observed specific frequency is a factor of $\sim$5 higher at [Fe/H] $\simeq -1.5$ compared to [Fe/H] $\simeq -0.5$ (in NGC 5128; Harris & Harris 2002). Because of the metallicity gradient, metal-rich clusters are closer to the Galaxy centre where the tides are stronger, resulting in more mass-loss for metal-rich clusters. Depending on the metallicity gradient, this effect may overcome the lower $|\dot{M}|$ at high [Fe/H]. We quantify this with our adopted implementation of [Fe/H]-gradient: we assume metal-rich GCs are within $R \lesssim 3$ kpc and metal-poor GCs at $R \gtrsim 3$ kpc (see Fig. 8). From this we find that the specific frequency of metal-poor (metal-rich) GCs is $\sim$ 4 per cent (2 per cent). So we recover the same sign as the observation, but the slope is not as steep (factor of $\sim$2 versus $\sim$5). Note that this is computed from the ratio of surviving clusters over disrupted clusters, not considering the actual field stars in the Milky Way.

6.4 Initial density

Our model assumes a fixed $\mathcal{R}_0 = \rho_{\rm h,0}/\rho_{\rm h,f} = 30$ (or $r_{\rm h,\,0}/r_{\rm J,\,eff} \simeq 0.05$, see Table 1). Here we discuss the implication for the distribution of initial densities of the GC population. Because $\rho_{\rm h,\,f}$ depends on the tidal field, from the orbit distribution we can derive the implied initial distribution of $\rho_{\rm h,\,0}$. For the SIS, $\rho_{\rm h,\,f}$ can be found from $\Omega_{\rm tid}$ as $\rho_{\rm h,f} \simeq 78.3\Omega_{\rm tid}^2/G$. To obtain a well-sampled density distribution, we draw 2×10^7 initial masses and orbits from $f_0(M_{\rm i}, \boldsymbol{R}, \boldsymbol{V})$ (equation 12) with parameters of Model (8), restricted to $M_{\rm i} > 10^4\,{\rm M}_\odot$, because lower mass GCs almost all dissolve. We evolve the initial masses to present-day masses with equation (5), which results in $\sim 2.5 \times 10^5$ surviving clusters. We note that although our model only evolves initial masses after stellar evolution (that is, $M_{\rm i}$), the $\dot{M}$ parameters that we use are based on initial densities before stellar evolution so we can find $\rho_{\rm h,\,0}$ for each GC from $\rho_{\rm h,0} = 30\,\rho_{\rm h,f}(R, \boldsymbol{V})$.

In Fig. 13, we plot the initial half-mass density distribution of all clusters and of the surviving clusters. Including clusters in the range $10^2 - 10^4\,{\rm M}_\odot$ would increase the distribution of all GCs by a factor of $\sim$110. This shows a peak at $\sim 10^{4.5}\,{\rm M}_\odot\,{\rm pc}^{-3}$. This is roughly an order of magnitude higher than young massive clusters in the Local Universe (Portegies Zwart et al. 2010; Brown & Gnedin 2021), but it is expected that GCs at a redshift of $z \simeq 4$ form denser because galaxies had higher gas fractions and velocity dispersion leading to higher pressure. Interestingly, it was recently shown (Antonini et al. 2023) that an initial density of $\gtrsim 10^4\,{\rm M}_\odot\,{\rm pc}^{-3}$ is what is needed to create sufficient numbers of (hierarchical) BH mergers to explain the gravitational wave sources with large primary masses ($\gtrsim 20\,{\rm M}_\odot$).

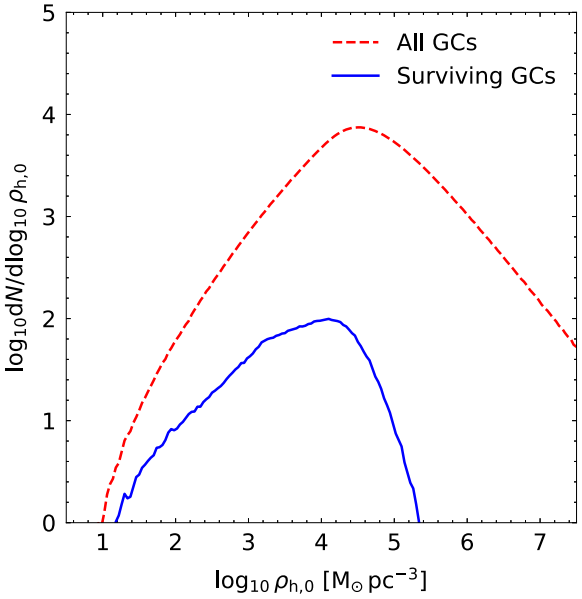

Figure 13. Initial density distribution for all clusters in Model (8) (red, dashed line) and for the surviving clusters (blue, solid line).

Fig. 13 shows that the surviving clusters have slightly lower initial densities, because the densest clusters are typically located near the Galactic centre where tidal disruption is most efficient.

We also compare these densities to present-day (half-mass) densities of Milky Way GCs. These can be estimated by assuming that mass follows light and give the median (mean) $\rho_{\rm h} \sim 300(1000)\,{\rm M}_\odot\,{\rm pc}^{-3}$, that is, more than an order of magnitude lower than the peak density in our model after stellar evolution. This is encouraging because it is expected that the densities of clusters decrease further after stellar mass-loss as the result of two-body relaxation (Hénon 1965; Gieles et al. 2011).

6.5 Black holes in present-day GCs

In our model, all GCs undergo a BH-dominated phase towards the end of their evolution. This is the result of our assumption of a constant $\mathcal{R}_0$ for all clusters. As already discussed in Section 6.4, this assumption is intended to describe the average GC. We can also estimate the distribution of the BH mass fraction ($f_{\rm BH}$) in our model. Towards the end of the evolution, $M_{\rm BH}$ remains approximately constant (Fig. 2). Assuming a constant mass-loss rate in time ($y = 1$), the distribution $dN/df_{\rm BH} = (dN/dM)\,|\partial M/\partial f_{\rm BH}|$ and because $dN/dM \simeq$ constant at low masses and $f_{\rm BH} \propto 1/M$ for $M_{\rm BH} \simeq$ constant, we have $dN/df_{\rm BH} \propto f_{\rm BH}^{-2}$, so we expect the majority of clusters to have $f_{\rm BH}$ close to the minimum $f_{\rm BH} \simeq 0.02$ for clusters with $\mathcal{R}_0 = 30$ (see Fig. 2).

Various studies have pointed out that populations of stellar-mass BHs may be present in GCs, based on their large core radii (Mackey et al. 2007, 2008); the absence of mass segregation of stars in some GCs (Alessandrini et al. 2016; Peuten et al. 2016; Weatherford et al. 2020); the central mass-to-light ratio (for the cases of Omega Centauri and 47 Tucanae; see Baumgardt et al. 2019; Zocchi, Gieles & Hénault-Brunet 2019; Hénault-Brunet et al. 2019); the core over half-light radius (Askar, Arca Sedda & Giersz 2018; Kremer et al. 2020) and the presence of tidal tails (see Gieles et al. 2021, for the case of Palomar 5). Breen & Heggie (2013) suggest that all GCs apart from the ones that are classified as 'core collapsed' possess BHs, which implies that 80 per cent of Milky Way GCs still contains BHs. Several studies that try to quantify $f_{\rm BH}$ for larger numbers of GCs with different methods have recently become available (Askar et al. 2018; Weatherford et al. 2020; Dickson et al. 2023). There is generally poor agreement for individual GCs, but all studies find typical $f_{\rm BH} \simeq 0.01$, with exceptions like Omega Centauri ($f_{\rm BH} \simeq 0.05$; Baumgardt et al. 2019; Zocchi et al. 2019) and Pal 5 ($f_{\rm BH} \simeq 0.2$; Gieles et al. 2021).

Observationally inferred $f_{\rm BH}$ are very uncertain and with the available data we can only say at this moment that observations support our assumption that the majority of GCs retained some BHs until the present day. In a future modelling exercise that includes also the evolution of cluster radii it would be interesting to see how a spread in $\mathcal{R}_0$ affects the final distribution of $f_{\rm BH}$ to check, for example, whether the fraction of core collapsed clusters can be reproduced.

6.6 Tidal perturbations and black holes

Several studies have invoked tidal perturbations with gas clouds in the early Universe to explain the shape of the GCMF (Elmegreen 2010; Kruijssen 2015; Pfeffer et al. 2018; Reina-Campos et al. 2018). The magnitude of this disruption mechanism is highly uncertain because it relies on poorly understood conditions in the early Universe (initial cluster densities, gas properties, etc.). The population models of Pfeffer et al. (2018) and Reina-Campos et al. (2018) adopt initial radii of ~ 3 pc (after stellar evolution), implying densities of $\rho_{\rm h,0} \sim 10^3\,{\rm M}_\odot{\rm pc}^{-3}$, i.e. more than an order of magnitudes lower than the peak of the initial density distribution in our models (Fig. 13). Because the disruption time-scale due to tidal shocks is directly proportional to the initial density, a higher initial density would decrease the disruption rate in their models. Similarly, including tidal shocks in our model would have only a small effect on our results.

The interplay between BH heating and tidal shocks is likely non-linear, but we can speculate what would happen if both effects play a role. The BHs sink to the cluster centre on a time-scale of ~ 10 Myr, while interactions with gas clouds can be important for up to ~ 1 Gyr, hence tidal shocks will predominantly remove stars from the cluster and the BH population is shielded, thereby increasing $f_{\rm BH}$. So mass-loss as a result of tidal shocks amplifies the effect of BHs at later times because of an increase in $f_{\rm BH}$.

6.7 Application to other galaxies

We applied our modelling to the Milky Way GC system, so it is interesting to consider to what extent our results apply to other galaxies. For a power-law initial GCMF with index -2, relaxation-driven evaporation predicts a correlation between the turnover mass, $M_{\rm TO}$, and the average tidal field strength experienced by the GCs, $\langle\Omega_{\rm tid}\rangle$. One may therefore expect $M_{\rm TO}$ to depend on the galaxy mass/luminosity. Jordán et al. (2007) fit 'evolved Schechter functions'[1] to luminosity functions of GC systems in early-type galaxies in the Virgo Cluster. They find that $M_{\rm TO}$ is within a factor of $\sim$2 constant over two orders of magnitude of galaxy luminosity ($L_{\rm gal}$), with the faintest galaxies having on average a lower $M_{\rm TO}$. Using galaxy scaling relations and assuming that the half-light radius of a galaxy is a proxy for the typical orbital radii of GCs, Jordán et al. (2007) show that for galaxies with $M_B > -18$ (that is, approximately 1 mag fainter than the Milky Way), the average tidal field strength

[1]These functions correspond to our $\psi(M, \Omega_{\rm tid}, t)$, for $x = y = 1$.

depends on galaxy luminosity as $\langle\Omega_{\rm tid}\rangle \propto L_{\rm Gal}^{0.35}$, while for brighter galaxies it goes as $\langle\Omega_{\rm tid}\rangle \propto L_{\rm Gal}^{-0.5}$. They argue that this 'peaked' relation between $\langle\Omega_{\rm tid}\rangle$ and $L_{\rm Gal}$ is one of the explanations for the near constant $M_{\rm TO}$. The authors also show that in addition to $\langle\Omega_{\rm tid}\rangle$, variations of $M_{\rm c}$ with $L_{\rm gal}$ affect the relation $M_{\rm TO}(L_{\rm gal})$. For bright galaxies they find a correlation between $M_{\rm c}$ and $L_{\rm gal}$, which offsets the anticorrelation between $\langle\Omega_{\rm tid}\rangle$ and $L_{\rm gal}$, leading to a near constant $M_{\rm TO}$ and a correlation between the width of the GCMF and $L_{\rm gal}$, as is observed. These arguments apply to our suggested mass-loss recipe, with the additional effects of y and metallicity discussed in this work. This helps in reaching relatively high $M_{\rm TO}$ even in the smallest galaxies.

In the Local Group we have even fainter galaxies, with GC systems that resemble the Milky Way GCs. For example, the Fornax dSph galaxy has five old GCs with an average logarithmic mass $\mu_{\log_{10}M} \simeq 5.0 \pm 0.2$ and dispersion $\sigma_{\log_{10}M} \simeq 0.5 \pm 0.2$ (based on the luminosities from Larsen, Strader & Brodie 2012, and a mass-to-light ratio of 1.8). Four of these GCs are metal-poor ([Fe/H] $\lesssim -2$), so BHs are expected to be important for their evolution and $\dot{M}$. Adopting $V_{\rm c} = 20\,{\rm km\,s^{-1}}$ for Fornax dSph, and $R_{\rm eff} = 1$ kpc for the GC orbits and the same model parameters as for our population model (that is, $x = 0.67$, $y = 1.33$, $\dot{M}_{\rm ref} = -45\,{\rm M_\odot/Myr}$), we find $\mu_{\log_{10}M} = 4.8$ and $\sigma_{\log_{10}M} = 0.5$, in satisfactory agreement with the observed GCMF. An important constraint for GC evolution models comes from the field stars. Larsen et al. (2012) find that about 20 per cent–30 per cent of all metal-poor stars ([Fe/H] $\lesssim -2$) in the galaxy resides in the four metal-poor GCs. For the simple model GCMF discussed here, and the assumption that all stars formed in GCs, we find this fraction to be $\sim$ 20 per cent(40 per cent) for $M_{\rm lo} = 10^2\,{\rm M_\odot}(10^4\,{\rm M_\odot})$ (see also Chen & Gnedin 2023, who reach a similar conclusion). We therefore conclude that even in the faintest galaxies our proposed mass-loss model can reproduce the shape of the GCMF.

7 CONCLUSIONS

We find that two-body relaxation in a static tidal field can be the dominant disruption process in shaping the GCMF if GCs retain some of their BHs. Earlier studies on the effect of evaporation on the shape of the GCMF showed that models of clusters without BHs cannot reproduce the observed shape of the GCMF and its insensitivity to Galactocentric radius (Baumgardt 1998; Vesperini 1998). In particular, these models find a turnover mass that is too low at large Galactocentric radii. Using N-body models of clusters with BHs, we show that the initial density is a critical parameter in setting the dynamical retention of BHs, and that high density clusters (relative to the tidal density) eject all their BHs and have similar $\dot{M}$ to clusters without BH. In models where dynamical BH retention is modest, the resulting $|\dot{M}|$ is still an order of magnitude higher than for models without BHs (Fig. 3) and the resulting mass evolution is sufficient to explain the observed turnover mass of $\sim 10^5\,{\rm M_\odot}$ at $R \gtrsim 10$ kpc (Fig. 11).

Several additional ingredients are needed to reduce the turnover mass in the inner galaxy to a similar value. We show that the anisotropy profile of GC orbits, a Schechter-like truncation in the initial GCMF, the metallicity gradient of GCs, and the effect of the past tidal evolution all reduce the decline of the turnover mass with R, with the combined effect providing a satisfactory match to the properties of Milky Way GCs (Figs 9 and 10). The proposed solution to the GCMF problem implies that the turnover mass gradually decreases with redshift, which is different from models that rely on early disruption mechanisms which leads to a redshift independent turnover mass. The difference may be observable with future thirty-metre class telescope and/or the *JWST* (Kruijssen 2015).

We present a modified analytical model for the cluster disruption rate that accounts for the effect of BHs. It is given by equation (4) and for our parameters it reads

$$\dot{M} = -45\,{\rm M_\odot\,Myr^{-1}} \left(\frac{M}{M_{\rm i}}\right)^{-1/3} \left(\frac{M_{\rm i}}{2\times10^5\,{\rm M_\odot}}\right)^{1/3} \frac{\Omega_{\rm tid}}{0.32\,{\rm Myr^{-1}}}. \quad (21)$$

The scaling with the initial mass is the same as of clusters without BHs (equation 1, Fig. 1). The scaling with the remaining mass fraction depends on the cluster density, which sets the dynamical BH retention, and the index can be between 1/3 (high density, all BHs ejected) and -1 (low density, almost all BHs retained, Fig. 2 and Table 1). For negative indices, we obtain the 'jumping' evolution of mass with time, where the mass-loss rate accelerates near cluster dissolution (Fig. 6).

Although we have focused on the Milky Way system, the physical ingredients of our model are found in all galaxies and we therefore expect that the model presented here can also explain the near universality of the GCMF among different galaxies. More work is needed to confirm this.

ACKNOWLEDGEMENTS

We thank the referee, Douglas Heggie, for carefully reading the manuscript and providing useful feedback. We thank Yingtian Chen for providing results of the hierarchical model described in Appendix A and Nate Bastian for helpful comments on the manuscript. MG acknowledges support from the Ministry of Science and Innovation (EUR2020-112157, PID2021-125485NB-C22, CEX2019-000918-M funded by MCIN/AEI/10.13039/501100011033) and from AGAUR (SGR-2021-01069). OG was supported in part by the U.S. National Science Foundation through grant AST-1909063 and by NASA through contract NAS5-26555 for STScI program HST-AR-16614. Most of the processing of the results has been done using the PYTHON programming language and the following open source modules: NUMPY,[2] SCIPY,[3] MATPLOTLIB.[4]

DATA AVAILABILITY

N-body data are available from MG upon reasonable request. All observational data of Milky Way GCs are from Harris (2010). The PYTHON code EVGCMF to evolve the GCMF is available from https://github.com/mgieles/evgcmf. The data used in the Appendix are available from OG upon reasonable request.

REFERENCES

Aarseth S. J., 2003, Gravitational N-Body Simulations. Cambridge Univ. Press, Cambridge
Alessandrini E., Lanzoni B., Ferraro F. R., Miocchi P., Vesperini E., 2016, ApJ, 833, 252
Antonini F., Gieles M., Dosopoulou F., Chattopadhyay D., 2023, MNRAS, 522, 466
Askar A., Arca Sedda M., Giersz M., 2018, MNRAS, 478, 1844
Asplund M., Amarsi A. M., Grevesse N., 2021, A&A, 653, A141

[2] http://www.numpy.org
[3] http://www.scipy.org
[4] http://matplotlib.sourceforge.net

Banerjee S., Kroupa P., 2011, ApJ, 741, L12
Banerjee S., Belczynski K., Fryer C. L., Berczik P., Hurley J. R., Spurzem R., Wang L., 2020, A&A, 639, A41
Barnard R., Garcia M., Murray S. S., 2012, ApJ, 757, 40
Bastian N., Lardo C., 2018, ARA&A, 56, 83
Baumgardt H. et al., 2019, MNRAS, 488, 5340
Baumgardt H., 1998, A&A, 330, 480
Baumgardt H., 2001, MNRAS, 325, 1323
Baumgardt H., Makino J., 2003, MNRAS, 340, 227
Baumgardt H., Kroupa P., Parmentier G., 2008, MNRAS, 384, 1231
Baumgardt H., Sollima A., Hilker M., 2020, Publ. Astron. Soc. Austr., 37, e046
Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689
Binney J., Tremaine S., 2008, Galactic Dynamics: Second Edition. Princeton Univ. Press, Princeton
Breen P. G., Heggie D. C., 2013, MNRAS, 432, 2779
Bromm V., Clarke C. J., 2002, ApJ, 566, L1
Brown G., Gnedin O. Y., 2021, MNRAS, 508, 5935
Cai M. X., Gieles M., Heggie D. C., Varri A. L., 2016, MNRAS, 455, 596
Chatterjee S., Rodriguez C. L., Rasio F. A., 2017, ApJ, 834, 68
Chen Y., Gnedin O. Y., 2022, MNRAS, 514, 4736
Chen Y., Gnedin O. Y., 2023, MNRAS, preprint (arXiv:2301.08218)
Chernoff D. F., Weinberg M. D., 1990, ApJ, 351, 121
Choksi N., Gnedin O. Y., 2019a, MNRAS, 486, 331
Choksi N., Gnedin O. Y., 2019b, MNRAS, 488, 5409
Choksi N., Gnedin O. Y., Li H., 2018, MNRAS, 480, 2343
Chomiuk L., Strader J., Maccarone T. J., Miller-Jones J. C. A., Heinke C., Noyola E., Seth A. C., Ransom S., 2013, ApJ, 777, 69
Contenta F., Varri A. L., Heggie D. C., 2015, MNRAS, 449, L100
Deason A. J., Belokurov V., Weisz D. R., 2015, MNRAS, 448, L77
Deason A. J., Belokurov V., Sanders J. L., 2019, MNRAS, 490, 3426
Dickson N., Hénault-Brunet V., Baumgardt H., Gieles M., Smith P., 2023, MNRAS, preprint (arXiv:2303.01637)
Eddington A. S., 1915, MNRAS, 75, 366
Elmegreen B. G., 2010, ApJ, 712, L184
Elmegreen B. G., Efremov Y. N., 1997, ApJ, 480, 235
Fall S. M., Rees M. J., 1985, ApJ, 298, 18
Fall S. M., Zhang Q., 2001, ApJ, 561, 751
Fernández-Trincado J. G., Beers T. C., Tang B., Moreno E., Pérez-Villegas A., Ortigoza-Urdaneta M., 2019, MNRAS, 488, 2864
Forbes D. A. et al., 2018, Proc. R. Soc. A, 474, 20170616
Fryer C. L., Belczynski K., Wiktorowicz G., Dominik M., Kalogera V., Holz D. E., 2012, ApJ, 749, 91
Fukushige T., Heggie D. C., 2000, MNRAS, 318, 753
Gieles M., 2009, MNRAS, 394, 2113
Gieles M., Baumgardt H., 2008, MNRAS, 389, L28
Gieles M., Renaud F., 2016, MNRAS, 463, L103
Gieles M., Heggie D. C., Zhao H., 2011, MNRAS, 413, 2509
Gieles M., Erkal D., Antonini F., Balbinot E., Peñarrubia J., 2021, Nat. Astron., 5, 957
Giersz M., Askar A., Wang L., Hypki A., Leveque A., Spurzem R., 2019, MNRAS, 487, 2412
Giesers B. et al., 2018, MNRAS, 475, L15
Harris W. E., 1996, AJ, 112, 1487
Harris W. E., 2001, in Labhardt L., Binggeli B., eds, Saas-Fee Advanced Course 28: Star Clusters. Springer, Berlin, p. 223
Harris W. E., 2010, preprint (arXiv:1012.3224)
Harris W. E., Harris G. L. H., 2002, AJ, 123, 3108
Harris W. E., Pudritz R. E., 1994, ApJ, 429, 177
Hénault-Brunet V., Gieles M., Sollima A., Watkins L. L., Zocchi A., Claydon I., Pancino E., Baumgardt H., 2019, MNRAS, 483, 1400
Hénon M., 1961, Ann. Astrophys., 24, 369
Hénon M., 1965, Ann. Astrophys., 28, 62
Horta D. et al., 2021, MNRAS, 500, 5462
Hurley J. R., Pols O. R., Tout C. A., 2000, MNRAS, 315, 543
Jordán A. et al., 2007, ApJS, 171, 101
Kamann S. et al., 2020, A&A, 635, A65
Kimm T., Cen R., Rosdahl J., Yi S. K., 2016, ApJ, 823, 52
Kravtsov A. V., Gnedin O. Y., 2005, ApJ, 623, 650
Kremer K. et al., 2020, ApJS, 247, 48
Kroupa P., 2001, MNRAS, 322, 231
Kroupa P., Aarseth S., Hurley J., 2001, MNRAS, 321, 699
Kruijssen J. M. D., 2015, MNRAS, 454, 1658
Kruijssen J. M. D., Pfeffer J. L., Crain R. A., Bastian N., 2019, MNRAS, 486, 3134
Krumholz M. R., McKee C. F., Bland-Hawthorn J., 2019, ARA&A, 57, 227
Lamers H. J. G. L. M., Gieles M., Bastian N., Baumgardt H., Kharchenko N. V., Portegies Zwart S., 2005, A&A, 441, 117
Lamers H. J. G. L. M., Baumgardt H., Gieles M., 2010, MNRAS, 409, 305
Lamers H. J. G. L. M., Baumgardt H., Gieles M., 2013, MNRAS, 433, 1378
Larsen S. S., Strader J., Brodie J. P., 2012, A&A, 544, L14
Lee H. M., Ostriker J. P., 1987, ApJ, 322, 123
Maccarone T. J., Kundu A., Zepf S. E., Rhode K. L., 2007, Nature, 445, 183
Maccarone T. J., Kundu A., Zepf S. E., Rhode K. L., 2011, MNRAS, 410, 1655
Mackey A. D., Wilkinson M. I., Davies M. B., Gilmore G. F., 2007, MNRAS, 379, L40
Mackey A. D., Wilkinson M. I., Davies M. B., Gilmore G. F., 2008, MNRAS, 386, 65
Martell S. L. et al., 2016, ApJ, 825, 146
McLaughlin D. E., Fall S. M., 2008, ApJ, 679, 1272
Meng X., Gnedin O. Y., 2021, MNRAS, 502, 1433
Meng X., Gnedin O. Y., 2022, MNRAS, 515, 1065
Merritt D., 1985, AJ, 90, 1027
Michie R. W., 1963, MNRAS, 125, 127
Miller-Jones J. C. A. et al., 2015, MNRAS, 453, 3918
Muratov A. L., Gnedin O. Y., 2010, ApJ, 718, 1266
Okazaki T., Tosa M., 1995, MNRAS, 274, 48
Osipkov L. P., 1979, Pisma Astron. Zh., 5, 77
Ostriker J. P., Spitzer L. J., Chevalier R. A., 1972, ApJ, 176, L51
Peebles P. J. E., Dicke R. H., 1968, ApJ, 154, 891
Peuten M., Zocchi A., Gieles M., Gualandris A., Hénault-Brunet V., 2016, MNRAS, 462, 2333
Peuten M., Zocchi A., Gieles M., Hénault-Brunet V., 2017, MNRAS, 470, 2736
Pfeffer J., Kruijssen J. M. D., Crain R. A., Bastian N., 2018, MNRAS, 475, 4309
Portegies Zwart S. F., McMillan S. L. W., Gieles M., 2010, ARA&A, 48, 431
Prieto J. L., Gnedin O. Y., 2008, ApJ, 689, 919
Reina-Campos M., Kruijssen J. M. D., Pfeffer J., Bastian N., Crain R. A., 2018, MNRAS, 481, 2851
Renaud F., Gieles M., Boily C. M., 2011, MNRAS, 418, 759
Rodriguez C. L., Hafen Z., Grudić M. Y., Lamberts A., Sharma K., Faucher-Giguère C.-A., Wetzel A., 2023, MNRAS, 521, 124
Saracino S. et al., 2022, MNRAS, 511, 2914
Schiavon R. P. et al., 2017, MNRAS, 465, 501
Spitzer L. J., 1958, ApJ, 127, 17
Strader J., Chomiuk L., Maccarone T. J., Miller-Jones J. C. A., Seth A. C., 2012, Nature, 490, 71
Tang B., Fernández-Trincado J. G., Liu C., Yu J., Yan H., Gao Q., Shi J., Geisler D., 2020, ApJ, 891, 28
van den Bosch F. C., Lewis G. F., Lake G., Stadel J., 1999, ApJ, 515, 50
Vasiliev E., 2019, MNRAS, 484, 2832
Vesperini E., 1998, MNRAS, 299, 1019
Vesperini E., 2000, MNRAS, 318, 841
Vesperini E., Zepf S. E., 2003, ApJ, 587, L97
Vesperini E., Zepf S. E., Kundu A., Ashman K. M., 2003, ApJ, 593, 760
Wang L., 2020, MNRAS, 491, 2413
Wang L., Spurzem R., Aarseth S., Nitadori K., Berczik P., Kouwenhoven M. B. N., Naab T., 2015, MNRAS, 450, 4070
Weatherford N. C., Chatterjee S., Kremer K., Rasio F. A., 2020, ApJ, 898, 162
Zocchi A., Gieles M., Hénault-Brunet V., 2019, MNRAS, 482, 4713

APPENDIX A: CLUSTER PROPERTIES FROM A FULL FORMATION MODEL

Several recent studies focused on modelling GCs from their formation in high-redshift galaxies through evolution until the present in the context of hierarchical galaxy formation (Choksi et al. 2018; Pfeffer et al. 2018; Kruijssen et al. 2019; Rodriguez et al. 2023). These studies assume that proto-GCs form in giant molecular clouds with the same initial cluster mass function as young clusters in the nearby Universe. This formation process continues as long as the specific adopted criteria are satisfied and results in a range of GC formation times, typically $1-4$ Gyr after the Big Bang. Then cluster mass-loss is calculated using various combinations of two-body relaxation, tidal shocks, and stellar mass loss. Despite differences in detailed implementation of all these processes, these studies reach similar conclusions that the resulting present-day GC populations can match observed properties of GC systems such as the age–metallicity distribution and the spatial and kinematic distributions. However, the resulting GCMF is usually skewed towards lower masses than observed. Given the successes in reproducing other GC properties, which support the main assumptions of the hierarchical models, the issues with the GCMF are likely to be due to inaccuracies in modelling cluster disruption. Possible inaccuracies can arise from insufficient resolution of the tidal field along cluster trajectories or using simplified mass-loss prescriptions.

Our work can help improve the modelling of cluster evolution in hierarchical models. And in turn, hierarchical models can help test some of the assumptions made in this work. Chen & Gnedin (2022) presented the latest version of a GC formation and evolution model, which matches most observed properties of the Galactic GC system. Here we use the predicted properties of GCs from the model to validate our adopted initial conditions.

We use the catalogue of model GC properties available online at https://github.com/ognedin/gc_model_mw. The catalogue contains three systems chosen specifically to be analogous to the Milky Way in its present halo and stellar mass and in its history of the mass assembly. We use all three systems to represent a range of possible initial conditions of the Galactic GC system. Fig. A1 shows the number density profile of all clusters that formed in the model as a function of the distance to the main galaxy center. This includes clusters formed throughout the cosmic time, although the middle half of them formed at the cosmic times between 11.1 and 12.5 Gyr, which is close to the assumed fixed age of 12 Gyr in our population model.

The assumed initial number density profile in our population model (equation 16) is a good match to the range of profiles shown in red. The hierarchical model predicts a slightly shallower slope at large radii ($R > 30$ kpc) but those radii correspond to the locations of satellite galaxies in which outer GCs formed. The satellites may bring their GC systems closer to the main galaxy by dynamical friction and the eventual radii of these halo GCs would be smaller. In the population model, we do not include changes of orbits due to dynamical friction, and therefore survived clusters would retain their initial radii. Thus we conclude that the hierarchical model provides support for our assumed $n_0(R)$.

To check consistency with observations, in the lower blue shaded region we show the range of number density profiles of survived clusters. Here the slope is more noticeably shallower than in our population model, but the difference is expected because the Chen & Gnedin (2022) model used a different GC disruption prescription with $x = y = 2/3$ and $\dot{M}_{\rm ref} = -45\,{\rm M}_\odot\,{\rm Myr}^{-1}$ (in our notation). The smaller y (compared to $y = 4/3$ in our Model 8) leads to slower disruption of low-mass clusters and allows them to survive longer at large radii where the disruption time is longer than the age. In our population model such clusters are more easily disrupted and the outer density profile steepens.

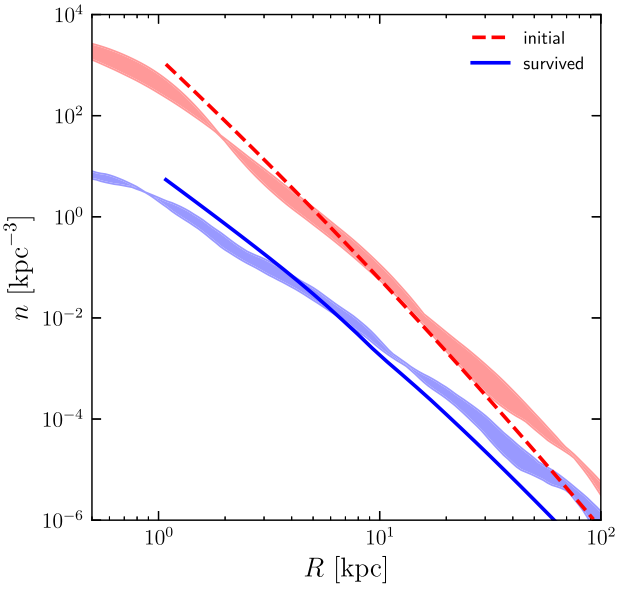

Figure A1. Number density profile of GCs from the mock catalogue of the hierarchical formation model (Chen & Gnedin 2022). Upper red shaded region shows the range of initial densities in the three model realizations of Milky Way-like systems. Lower blue shaded region shows the range of the density of survived clusters at present. Dashed red and solid blue lines show the corresponding densities in our population Model (8), also plotted in the right-hand column of Fig. 9.

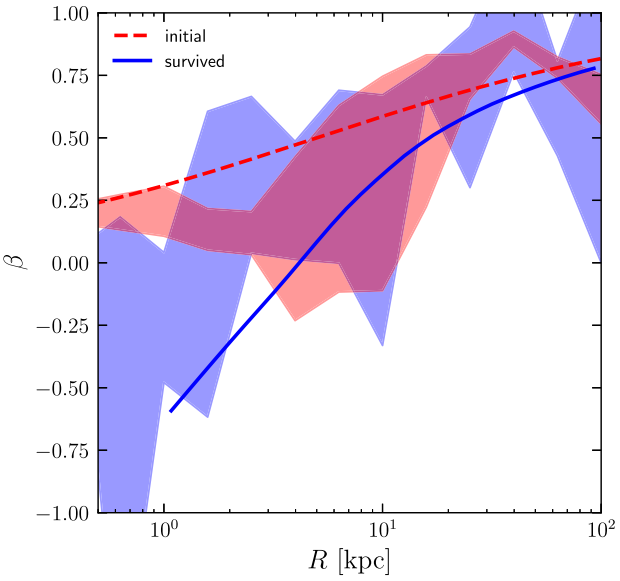

Figure A2. Velocity anisotropy profile of GCs from the mock catalogue of the hierarchical formation model. Shaded regions show the range covered by the three realizations: red for initial, blue for final. The dashed red and solid blue lines show the corresponding densities in our population Model (8), also plotted in right-hand column of Fig. 9.

Fig. A2 compares the velocity anisotropy profiles of the two models. Variations among the three hierarchical model realizations are large but the overall trend of initial β increasing with radius is in

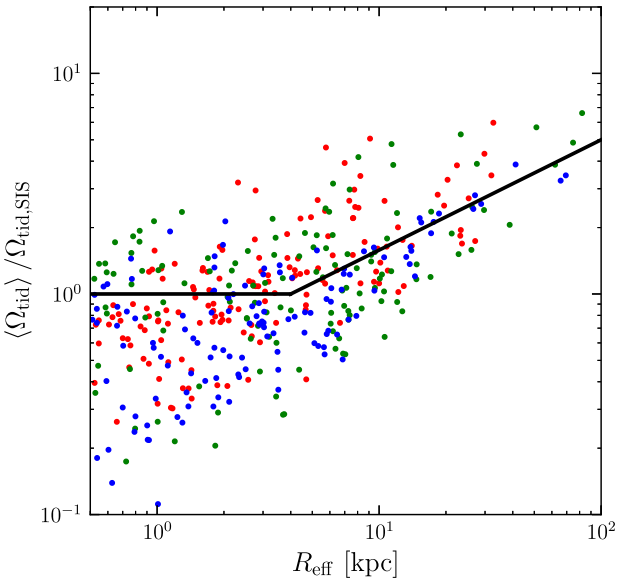

Figure A3. The effective strength of tidal field averaged over cluster history in the hierarchical formation model, relative to that of the best-fitting fixed SIS potential. Points show $\Omega_{\rm tid}$ averaged over all simulation outputs for three Milky Way analogue systems indicated by different colour. The solid line shows our modified expression $\Omega_{\rm tid,\,1}$ at $R_{\rm eff} > 4$ kpc.

reasonable agreement with the assumed form in equation (15). For the survived clusters, the β-profiles are even closer and both clearly predict a mildly tangential anisotropy in the inner few kpc. Thus we can conclude that the results of the full hierarchical formation model support our assumed initial conditions.

The hierarchical model also allows us to investigate the evolution of the tidal field along the trajectories of model GCs. For example, Meng & Gnedin (2022) showed that a typical effective strength of tidal field $\Omega_{\rm tid}$ was a factor of 10 higher in the first ∼300 Myr after cluster formation compared to the values inferred from the present-day potential. Young clusters experience stronger tides because they are still surrounded by dense gaseous and stellar structure. The tidal field can also vary rapidly in time depending on the GC trajectories. A higher fraction of GCs migrate outward from the galaxy centre than inward, also leading to the weaker tidal field at present. To account for this 'past evolution' of the tidal field, we calculate the time averaged $\Omega_{\rm tid}$ experienced by survived clusters in the Chen & Gnedin (2022) model and compare it with the value in our assumed potential. Fig. A3 shows the ratio of the two for model clusters as a function of their effective radius $R_{\rm eff} = R_{\rm p}(1+\epsilon)$, where $R_{\rm p}$ is the pericentre distance of the orbit near the present. The ratio is based on the following calculation.

Chen & Gnedin (2022) calculated the tidal strength via a combination of the highest and lowest eigenvalues of the tidal tensor that accounts for the tidal and centrifugal forces: $\Omega^2_{\rm tid} \simeq \lambda_1 - \lambda_3$. Typically, $\lambda_1 > 0$ and $\lambda_3 < 0$. For a SIS, $\lambda_1 = -\lambda_3 = V_{\rm c}^2/R^2$, and therefore

$$\Omega^2_{\rm tid,SIS}(R) = \frac{2V_{\rm c}^2}{R^2}. \tag{A1}$$

For a general power-law density distribution $\rho \propto R^{-n}$ with $0 < n < 3$

$$\Omega^2_{{\rm tid},n}(R) = 4\pi G\rho(R)\,\frac{n}{3-n} = \frac{n\,V_{\rm c}(R)^2}{R^2}, \tag{A2}$$

where $V_{\rm c}(R)^2 \equiv GM(R)/R$ and $n = 2$ for the SIS.

The SIS model is a good approximation to the total mass density in the three model realizations in the range of radii from 1 to 100 kpc; in the inner 1 kpc the density profile approaches a core. In the middle part of the galaxy, at $R_{\rm eff} \lesssim 4$ kpc, the SIS potential gives overall correct scaling of $<\Omega_{\rm tid}>$ with radius, however the scatter of individual points is significant. At larger radii, the present-day $\Omega_{\rm tid,\,SIS}$ visibly underestimates the past tidal strength. We can approximately correct this underestimate by switching to $\Omega_{{\rm tid},\,n}$ corresponding to shallower distribution with $n \approx 1$. This results in a stronger tidal field at $R_{\rm eff} > 4$ kpc relative to our SIS model by a factor

$$\frac{\Omega^2_{\rm tid,1}(R_{\rm eff})}{\Omega^2_{\rm tid,2}(R_{\rm eff})} = \frac{R_{\rm eff}}{4\,{\rm kpc}}. \tag{A3}$$

This adopted modified expression for $<\Omega_{\rm tid}>$ is shown by the broken line in Fig. A3 and is used in Models (6) and (8) to include the effect of the past evolution of GCs.

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