Star clusters: basic galactic building blocks
Proceedings IAU Symposium No. 266, 2009
R. de Grijs & J. R. D. Lépine, eds.
© International Astronomical Union 2010
doi:10.1017/S1743921309991116

On the origin of globular cluster bimodality

Oleg Y. Gnedin

Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA
ognedin@umich.edu

Abstract. Globular cluster systems in most large galaxies display bimodal color and metallicity distributions, which are frequently interpreted as indicating two distinct modes of cluster formation. The metal-rich (red) and metal-poor (blue) clusters have systematically different locations and kinematics in their host galaxies. However, the red and blue clusters have similar internal properties, such as their masses, sizes, and ages. It is therefore interesting to explore whether both metal-rich and metal-poor clusters could form by a common mechanism and still be consistent with the bimodal distribution. We show that if all globular clusters form only during mergers of massive, gas-rich protogalactic disks, their metallicity distribution could be statistically consistent with that of the Galactic globulars. We take the galaxy assembly history from cosmological dark-matter simulations and couple it with the observed scaling relations for the amount of cold gas available for star formation. In the best-fitting model, early mergers of smaller hosts create exclusively blue clusters, while subsequent mergers of progenitor galaxies with a range of masses create both red and blue clusters. Thus, bimodality arises naturally as the result of a small number of late, massive merger events. We calculate cluster mass loss, including the effects of two-body scattering and stellar evolution, and find that more blue than red clusters are disrupted by the present time because of their lower initial masses and older ages. The present-day mass function in the best-fitting model is consistent with the Galactic distribution. However, the spatial distribution of model clusters is much more extended than observed and is independent of the parameters of our model.

Keywords. galaxies: formation, galaxies: star clusters, globular clusters: general

1. Are the red and blue clusters really that different?

A self-consistent description of the formation of globular clusters remains a challenge to theorists. A particularly puzzling observation is the apparent bimodality, or even multimodality, of the color distribution of globular cluster systems in galaxies ranging from dwarf disks to giant ellipticals. This color bimodality likely translates into a bimodal distribution of the abundances of heavy elements such as iron. We know this to be the case in the Galaxy as well as in M31, where relatively accurate spectral measurements exist for a large fraction of the clusters. The two most frequently encountered modes are commonly called *blue* (metal poor) and *red* (metal rich).

Bimodality in the globular cluster metallicity distribution of luminous elliptical galaxies was proposed by Zepf & Ashman (1993), following a theoretical model by Ashman & Zepf (1992). The concept of cluster bimodality became universally accepted because the two populations also differ in other observed characteristics. The system of red clusters has a significant rotational velocity similar to the disk stars, while blue clusters have little rotational support in the three disk galaxies observed in detail, the Milky Way, M31, and M33. In elliptical galaxies, blue clusters have a larger velocity dispersion than red clusters, due to both a lack of rotation and their more extended spatial distribution. Red clusters are usually more spatially concentrated than blue clusters (Brodie & Strader 2006). All of these differences, however, are in the external properties (location and kinematics), which reflect *where* the clusters formed, but not *how*. The internal properties of the red and

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

blue clusters are similar, i.e., their masses, sizes, and ages, with only slight differences. Even the metallicities themselves differ typically by a factor of 10 between the two modes, which is not enough to affect the dynamics of molecular clouds from which these clusters formed. Could it be, then, that both red and blue clusters form in a similar way on small scales, such as in giant molecular clouds, while the differences in their metallicities and spatial distributions reflect when and where such clouds assemble?

All scenarios proposed in the literature assume different formation mechanisms for the red and blue clusters, and most scenarios envision the stellar population of one mode to be tightly linked to that of the host galaxy (e.g., Forbes *et al.* 1997; Côté *et al.* 1998; Strader *et al.* 2005). The other mode is assumed to have formed differently, in some unspecified 'primordial' way. This assumption only pushed the problem back in time, but did not solve it. For example, Beasley *et al.* (2002) used a semi-analytic model of galaxy formation to study bimodality in luminous elliptical galaxies and needed two separate prescriptions for the blue and red clusters. In their model, red clusters formed in gas-rich mergers with a fixed efficiency of 0.007 relative to field stars, while blue clusters formed in quiescent disks with a different efficiency of 0.002. The formation of blue clusters also had to be artificially truncated at $z = 5$. Strader *et al.* (2005) and Rhode *et al.* (2005) suggested that the blue clusters could instead have formed in very small halos at $z > 10$, before cosmic reionization removed cold gas from such halos. This scenario requires high efficiency of cluster formation in the small halos and also places stringent constraints on the age spread of blue clusters, of less than 0.5 Gyr. Unfortunately, even the most recent measurements of the relative cluster ages in the Galaxy (De Angeli *et al.* 2005; Marín–Franch *et al.* 2009) cannot detect age differences smaller than 9%, or about 1 Gyr, and therefore cannot support or falsify the reionization scenario.

2. Globular clusters could form in protogalactic disks

We set out to test whether a common mechanism could explain the formation of both red and blue modes and produce an entire metallicity distribution consistent with the observations. We begin with a premise of hierarchical galaxy formation in a ΛCDM (cold dark matter) universe. Primordial density fluctuations in the early Universe, probed directly by anisotropies in the cosmic microwave background, set the seeds for structure formation. Cosmological numerical simulations study the growth of the fluctuations through gravitational instability and show us a history of galaxy assembly. The simulations begin with tiny deviations from the Hubble flow, whose amplitudes are set by the measured power spectrum, while the phases are assigned randomly. Therefore, each particular simulation provides only a statistical description of a representative part of the Universe, although current models successfully reproduce major features of observed galaxies.

Hubble Space Telescope observations have convincingly demonstrated one of the likely formation routes for massive star clusters today: in the mergers of gas-rich galaxies (e.g., Holtzman *et al.* 1992; O'Connell *et al.* 1995; Whitmore *et al.* 1999; Zepf *et al.* 1999). We adopt this single formation mechanism and assume that globular clusters form only during massive gas-rich mergers. We follow the merging process of progenitor galaxies in a Galaxy-sized environment using a set of cosmological N-body simulations from Kravtsov *et al.* (2004). We need to decide what type and how many clusters will form in each merger event. For this purpose, we use observed scaling relations to assign each dark-matter subhalo a certain amount of cold gas that will be available for star formation throughout cosmic time and an average metallicity of that gas. To keep the model transparent, we choose a parametrization of the cold gas mass that is as simple as

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

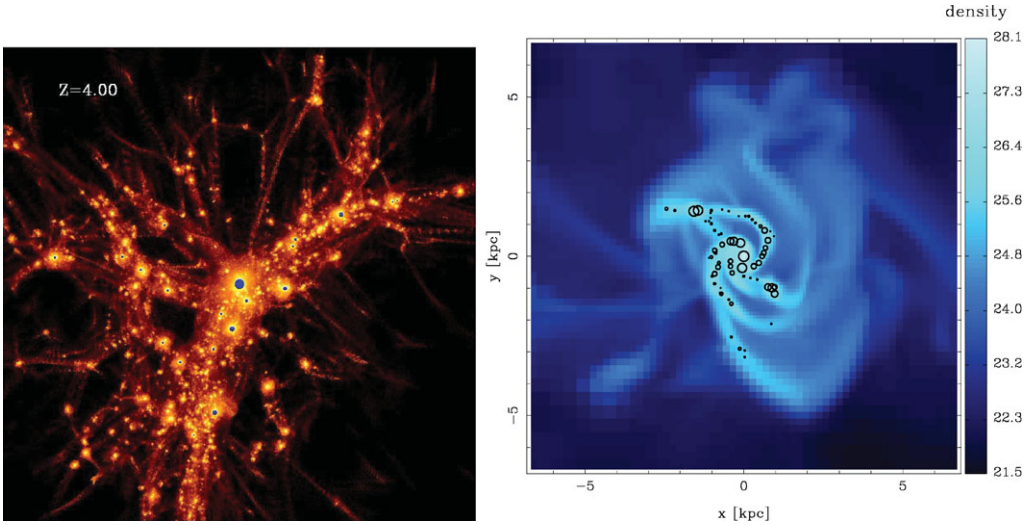

Figure 1. *(left)* Hierarchical buildup of a Milky Way-type galaxy in an adaptive-mesh simulation at $z = 4$. The view is centered on the largest progenitor galaxy in the simulation and shows a $1h^{-1}$ Mpc^2 region (comoving). Dark circles show protogalactic disks in the central regions of dark-matter halos. *(right)* A massive, gaseous disk with prominent spiral arms, seen face-on, in the process of active merging at $z = 4$. In Kravtsov & Gnedin (2005), model star clusters form in giant gas clouds, shown by circles with sizes corresponding to the cluster masses.

possible. Finally, we make the simplest assumption that the mass of all globular clusters formed in the merger is linearly proportional to the mass of this cold gas, $M_{\rm GC} \propto M_{\rm g}$. We discuss our results in Section 3.

Although such a model appears extremely simplistic, we have some confidence that it may capture the main elements of the formation of massive clusters. Kravtsov & Gnedin (2005) used a hydrodynamic simulation of the Galactic environment at high redshifts ($z > 3$) and found dense, massive gas clouds within the protogalactic clumps. These clouds assemble during gas-rich mergers of progenitor galaxies, when the cold gas forms a thin, self-gravitating disk. The disk develops strong spiral arms, which further fragment into separate molecular clouds located along the arms as beads on a string (see Figure 1). If the high-density regions of these clouds formed star clusters, the resulting distributions of cluster mass, size, and metallicity are consistent with those of Galactic metal-poor clusters. The high stellar density of Galactic clusters restricts their parent clouds to be in relatively massive progenitors, with a total mass of $M_{\rm h} > 10^9$ M$_\odot$. The mass of the molecular clouds increases with cosmic time, but the merger rate declines steadily. Therefore, the cluster formation efficiency peaks during an extended epoch, $5 < z < 3$, when the Universe is less than 2 Gyr old. The molecular clouds are massive enough to be shielded from extragalactic ultraviolet radiation, so that globular cluster formation is unaffected by the reionization of cosmic hydrogen. The mass function of model clusters is consistent with a power law $\mathrm{d}N/\mathrm{d}M \propto M^{-\alpha}$, where $\alpha = 2.0 \pm 0.1$, similar to that of the local young star clusters. The total mass of clusters formed in each progenitor is roughly proportional to the available gas supply and the total mass, $M_{\rm GC} \propto M_{\rm g} \propto M_{\rm h}$.

In Prieto & Gnedin (2008) we showed that subsequent mergers of the progenitor galaxies ensure that the present distribution of the globular cluster system is spheroidal, as observed. Since the hydrodynamic simulation was stopped at $z \approx 3.3$, we used the Kravtsov *et al.* (2004) N-body simulation to calculate cluster orbits to $z = 0$. We used the evolving properties of all progenitor halos, from the outputs with a time resolution of $\sim 10^8$ yr, to

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

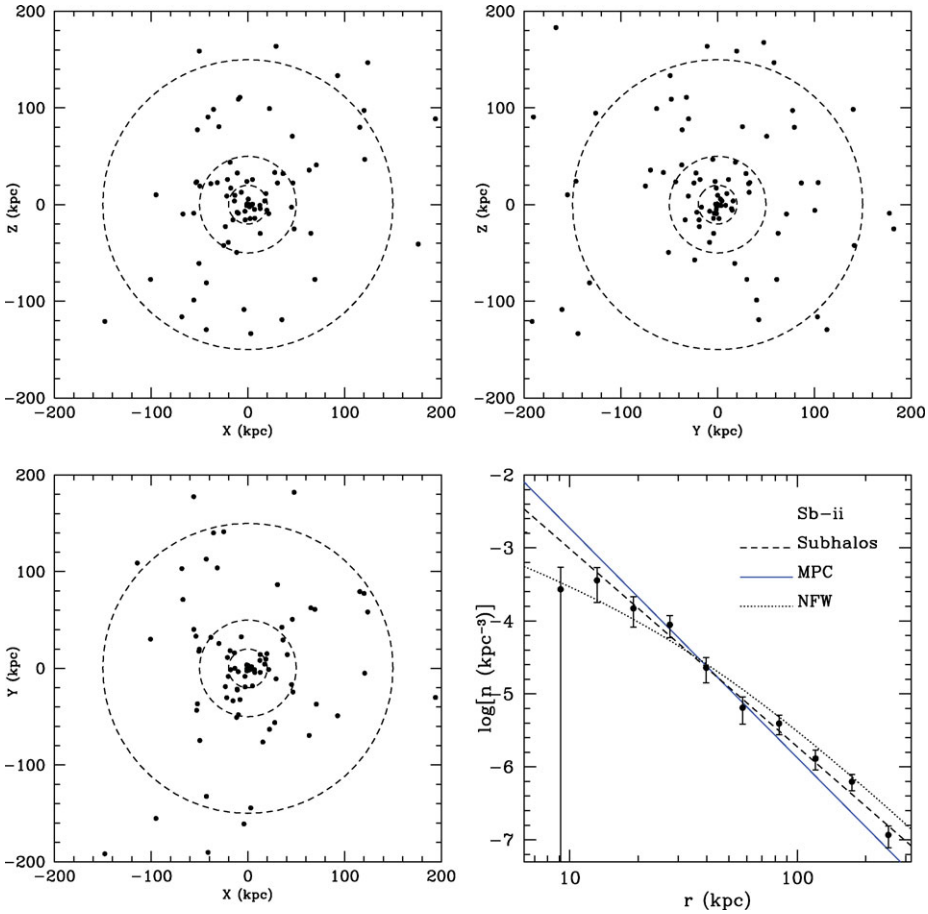

Figure 2. Spatial distribution of surviving model globular clusters, calculated by Prieto & Gnedin (2008). Dashed circles illustrate projected radii of 20, 50, and 150 kpc. The number–density profile *(bottom right)* can be fit by a power law, $n(r) \propto r^{-2.7}$. The distribution of model clusters is similar to that of surviving satellite halos (dashed line) and smooth dark matter (dotted line). It is also consistent with the observed slope of the metal-poor globular clusters in the Galaxy (solid line), plotted using data from the catalog of Harris (1996).

derive the gravitational potential in the entire computational volume at all epochs. We calculated the orbits of globular clusters in this potential from the time when their host galaxies accrete onto the main (most massive) galaxy. Using these orbits, we calculated the dynamical evolution of the model clusters, including the effects of stellar mass loss, two-body relaxation, tidal truncation, and tidal shocks.

In this model, all clusters form on nearly circular orbits within the protogalactic disks. Depending on the subsequent trajectories of the hosts, clusters form three main subsystems at the present time. *Disk clusters* formed in the most massive progenitor that eventually hosts the present Galactic disk. These clusters, found within the inner 10 kpc, do not actually stay on circular orbits but are instead scattered to eccentric orbits by perturbations from accreted galactic satellites. *Inner halo clusters*, found between 10 and 60 kpc, came from the now-disrupted satellite galaxies. Their orbits are inclined with respect to the Galactic disk and are fairly isotropic. *Outer halo clusters*, beyond 60 kpc from the center, are either still associated with the surviving satellite galaxies, or

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

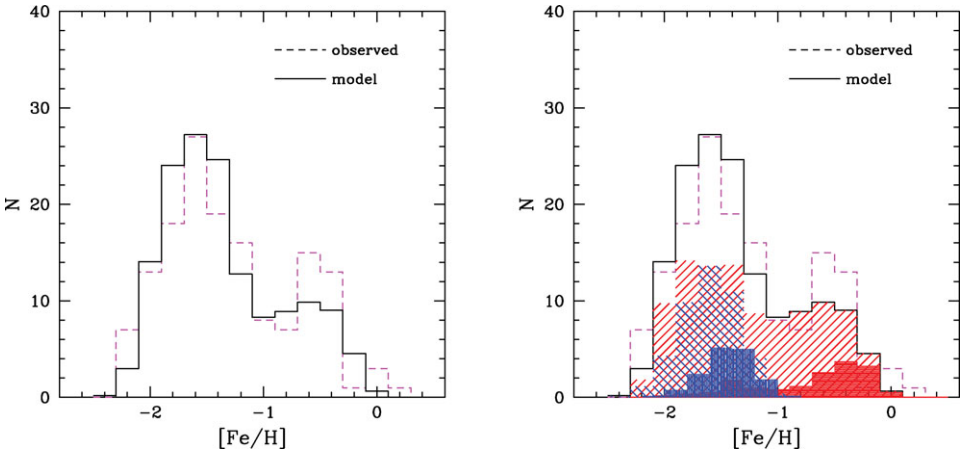

Figure 3. *(left)* Metallicities of model clusters at $z = 0$ (solid histogram) compared to the observed distribution of Galactic globular clusters (shaded histogram). *(right)* Contributions of major mergers (red) and early mergers (blue) to the total model count. Filled histograms are for the main Galactic disk. (From Muratov & Gnedin, in prep.)

were scattered away from their hosts during close encounters with other satellites and consequently appear isolated.

The azimuthally averaged space density of globular clusters is consistent with a power law, $n(r) \propto r^{-\gamma}$, with a slope $\gamma \approx 2.7$ (see Figure 2). Since all of the distant clusters originate in progenitor galaxies and share similar orbits with their hosts, the distribution of the clusters is almost identical to that of the surviving satellite halos. This power law is similar to the observed slope of the metal-poor ([Fe/H] < -1 dex) globular clusters in the Galaxy. However, the model clusters have a more extended spatial distribution (larger median distance) than observed. In the model it is largely determined by the orbits of the progenitor galaxies and the epoch of formation. Moore *et al.* (2006) showed that early-forming halos are more spatially concentrated and to match the Galactic distribution, globular clusters would need to form at $z \sim 12$. However, such an early formation is inconsistent with the requirement of high mass and high density of the parent molecular clouds. At present, we do not have an acceptable solution to this problem.

3. Metallicity bimodality as a natural outcome of hierarchical galaxy formation

Following the scenario outlined above, A. Muratov & Gnedin (in prep.) developed a semi-analytic model that aims to reproduce statistically the metallicity distribution of the Galactic globular clusters, as compiled by Harris (1996). The formation of clusters is triggered during a merger of gas-rich protogalaxies with a mass ratio of 1:5 or higher, and during very early mergers with any mass ratio when the cold-gas fraction in the progenitors is close to 100%. These criteria are applied at every time step of the simulation, every $\sim 10^8$ yr. In the best-fitting model, the mass of globular clusters formed in each event is

$$M_{\rm GC} \approx 7 \times 10^{-4}\ M_{\rm g} \approx 10^{-4}\ M_{\rm h}. \tag{3.1}$$

This imposes the minimum mass of a halo capable of forming a globular cluster. Based on dynamical-disruption arguments, we track only the clusters that are more massive

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

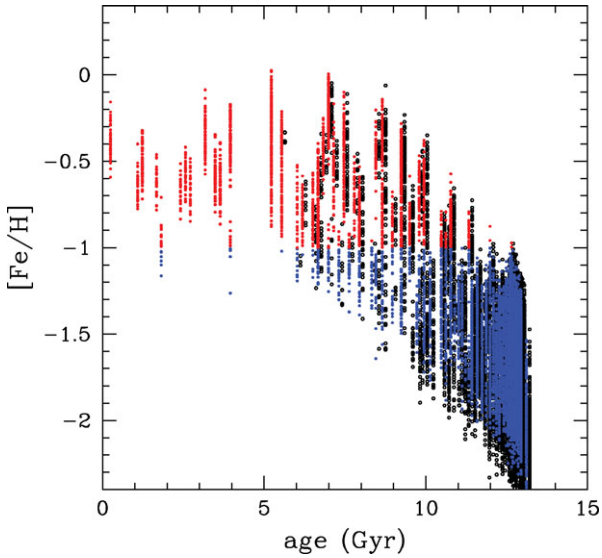

Figure 4. Age–metallicity relation of model clusters in 64 realizations of the Galactic sample. Filled (colored) circles show surviving clusters, open (black) circles show disrupted clusters. The buildup of massive halos drives the steep slope of this relation at early epochs. Notice an order-of-magnitude spread in metallicities of clusters forming at a given epoch.

than $M_{\rm min} = 10^5$ M$_\odot$. Therefore, to form even a single cluster with the minimum mass, the halo needs to be more massive than $\sim 10^9$ M$_\odot$. Individual cluster masses are drawn randomly from the assumed initial cluster mass function, $\mathrm{d}N/\mathrm{d}M \propto M^{-2}$, normalized to $M_{\rm GC}$. The clusters are assigned the mean metallicity [Fe/H] of their host protogalaxies, using the observed galaxy stellar-mass/metallicity relation. Overall, the model has five free parameters relating to the normalization of the cluster-formation rate and the merger mass ratio.

Figure 3 shows the metallicity distribution in the best-fitting model. The red peak is not as pronounced as in the observations but is noticeable. Interestingly, major mergers contribute both to the red and blue modes, in about equal proportions. Early mergers of low-mass progenitors contribute only blue clusters.

Note that the model has the same formation criteria for all clusters, without explicitly differentiating between the two modes. The only variables are the gradually changing amount of cold gas, the growth of protogalactic disks, and the merger rate. Yet, the model produces two peaks in the metallicity distribution, centered at [Fe/H] ≈ -1.6 and ≈ -0.6 dex, matching those of the Galactic globular clusters. The Kolmogorov–Smirnov probability of the model being consistent with the data is $P_{\rm KS} = 80\%$.

Our prescription links cluster metallicity to the average galaxy metallicity in a one-to-one relation, albeit with random scatter. Since the average galaxy metallicity grows monotonically with time, the cluster metallicity also grows with time. The model thus encodes an age–metallicity relation, in the sense that metal-rich clusters are younger than their metal-poor counterparts by several Gyr. However, Figure 4 shows that clusters of the same age may differ in metallicity by as much as a factor of 10, as they formed in progenitors of different mass. Observations of Galactic globular clusters do not show a clear age–metallicity relation, but instead indicate an age spread that increases with metallicity (De Angeli *et al.* 2005; Marín–Franch *et al.* 2009). Our model does not appear to obviously contradict this trend.

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

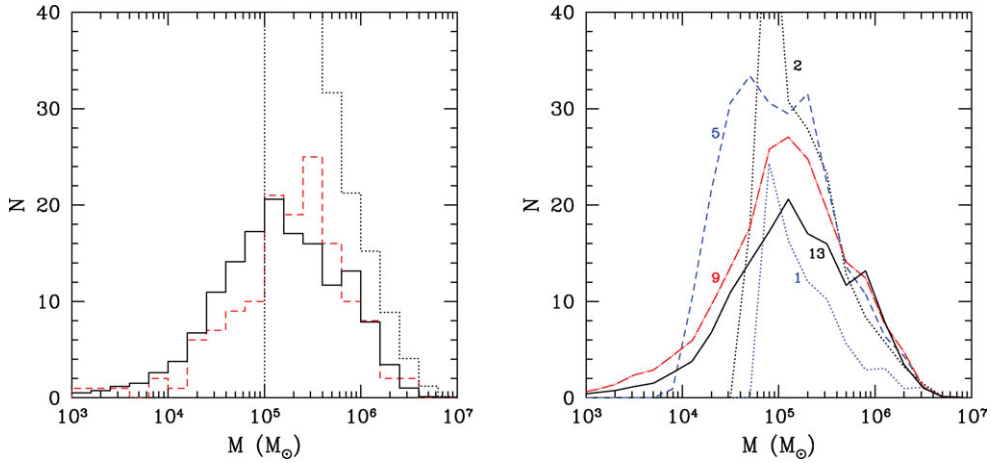

Figure 5. *(left)* Mass function of model clusters at $z = 0$ (solid histogram) versus Galactic clusters (dashed histogram). The dotted histogram shows the combined initial masses of the model clusters formed at all epochs, including those that did not survive until the present. We do not follow clusters with initial masses below 10^5 $M_\odot$. *(right)* Actual mass function at cosmic times of 1 Gyr ($z \approx 5.7$; dotted), 2 Gyr ($z \approx 3.2$; dotted), 5 Gyr ($z \approx 1.3$; dashed), 9 Gyr ($z \approx 0.5$; dot-dashed), and 13 Gyr ($z \approx 0$; solid).

Some of the old and low-mass clusters will be disrupted by the gradual escape of stars and will not appear in the observed sample. We calculated the effects of the dynamical evolution of the model clusters, including stellar mass loss and two-body evaporation, but ignored tidal shocks for simplicity. Figure 5 compares the resulting model mass function at $z = 0$ with the observed Galactic distribution. Since the model parameters were tuned to reproduce the metallicity distribution, the mass functions do not match as well but are still consistent at the level of $P_{\rm KS} = 7\%$. The majority of the disrupted clusters were blue clusters that formed in early low-mass progenitors.

The right panel of Figure 5 illustrates the evolution of the mass function as an interplay between the continuous buildup of massive clusters ($M > 10^5$ $M_\odot$) and the dynamical erosion of the low-mass clusters ($M < 10^5$ $M_\odot$). Expecting that the clusters below $M_{\rm min}$ would eventually be disrupted, we did not track their formation in the model. Instead, the low end of the mass function was built by the gradual evaporation of more massive clusters. Note that most of the clusters were not formed until the Universe was 2 Gyr old, corresponding to $z \approx 3$. The fraction of clusters formed before $z \approx 6$, when cosmic hydrogen was reionized, is small.

In calculating the rate of two-body relaxation, we assumed a standard result for the evaporation time, $t_{\rm d}(M) \approx 10^{10}(M/2\times10^5$ $M_\odot)$ yr. We also used an alternative rate with the weaker mass dependence, $t_{\rm d}(M) \approx 10^{10}(M/2\times10^5$ $M_\odot)$ yr, suggested by the recent results of Gieles & Baumgardt (2008). While this prescription leads to slower disruption of the low-mass clusters, we find that the resulting distribution at $z = 0$ is still consistent with observations at the level of $P_{\rm KS}$ of a few percent.

In our scenario, bimodality results from the history of galaxy assembly (rate of mergers) and the amount of cold gas in protogalactic disks. Early mergers are very frequent but involve relatively low-mass protogalaxies, which produce preferentially blue clusters. Late mergers are infrequent but typically involve more massive galaxies. As the number of clusters formed in each merger increases with progenitor mass, just a few late, supermassive mergers can produce a significant number of red clusters. The concurrent growth

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press

of the average metallicity of galaxies between the late mergers leads to an apparent 'gap' between the red and blue clusters.

We expect that our formalism could be applied to other galactic environments, such as those of elliptical galaxies with larger samples of globular clusters. For example, Peng *et al.* (2008) showed that the fraction of red clusters increases from 10 to 50% with increasing luminosity of elliptical galaxies in the Virgo cluster. Compared to the Galaxy, giant ellipticals are expected to experience more massive mergers, which would produce similar numbers of red and blue clusters according to Figure 3. Thus, a red-cluster fraction peaking at ∼50% is a natural outcome of hierarchical formation.

References

Ashman, K. M. & Zepf, S. E. 1992, *ApJ*, 384, 50

Beasley, M. A., Baugh, C. M., Forbes, D. A., Sharples, R. M., & Frenk, C. S. 2002, *MNRAS*, 333, 383

Brodie, J. P. & Strader, J. 2006, *ARA&A*, 44, 193

Côté, P., Marzke, R. O., & West, M. J. 1998, *ApJ*, 501, 554

De Angeli, F., Piotto, G., Cassisi, S., Busso, G., Recio–Blanco, A., Salaris, M., Aparicio, A., & Rosenberg, A. 2005, *AJ*, 130, 116

Forbes, D. A., Brodie, J. P., & Grillmair, C. J. 1997, *AJ*, 113, 1652

Gieles, M. & Baumgardt, H. 2008, *MNRAS* (Letters), 389, L28

Harris, W. E. 1996, *AJ*, 112, 1487

Holtzman, J. A., Faber, S. M., Shaya, E. J., Lauer, T. R., Groth, J., Hunter, D. A., Baum, W. A., Ewald, S. P., Hester, J. J., Light, R. M., Lynds, C. R., O'Neil, E. J., & Westphal, J. A. 1992, *AJ*, 103, 691

Kravtsov, A. V. & Gnedin, O. Y. 2005, *ApJ*, 623, 650

Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, *ApJ*, 609, 482

Marín–Franch, A., Aparicio, A., Piotto, G., Rosenberg, A., Chaboyer, B., Sarajedini, A., Siegel, M., Anderson, J., Bedin, L. R., Dotter, A., Hempel, M., King, I., Majewski, S., Milone, A. P., Paust, N., & Reid, I. N. 2009, *ApJ*, 694, 1498

Moore, B., Diemand, J., Madau, P., Zemp, M., & Stadel, J. 2006, *MNRAS*, 368, 563

O'Connell, R. W., Gallagher, J. S., Hunter, D. A., & Colley, W. N. 1995, *ApJ* (Letters), 446, L1

Peng, E. W., Jordán, A., Côté, P., Takamiya, M., West, M. J., Blakeslee, J. P., Chen, C.-W., Ferrarese, L., Mei, S., Tonry, J. L., & West, A. A. 2008, *ApJ*, 681, 197

Prieto, J. L. & Gnedin, O. Y. 2008, *ApJ*, 689, 919

Rhode, K. L., Zepf, S. E., & Santos, M. R. 2005, *ApJ* (Letters), 630, L21

Strader, J., Brodie, J. P., Cenarro, A. J., Beasley, M. A., & Forbes, D. A. 2005, *AJ*, 130, 1315

Whitmore, B. C., Zhang, Q., Leitherer, C., Fall, S. M., Schweizer, F., & Miller, B. W. 1999, *AJ*, 118, 1551

Zepf, S. E. & Ashman, K. M. 1993, *MNRAS*, 264, 611

Zepf, S. E., Ashman, K. M., English, J., Freeman, K. C., & Sharples, R. M. 1999, *AJ*, 118, 752

https://doi.org/10.1017/S1743921309991116 Published online by Cambridge University Press