STAR CLUSTER DEMOGRAPHICS. I. A GENERAL FRAMEWORK AND APPLICATION TO THE ANTENNAE GALAXIES

BRADLEY C. WHITMORE, RUPALI CHANDAR, AND S. MICHAEL FALL

Received 2006 February 13; accepted 2006 October 17

ABSTRACT

We present a framework for understanding the demographics of star cluster systems and develop a toy model that incorporates a universal initial power-law mass function, selected formation histories, selected disruption laws, and a convolution with common artifacts and selection effects found in observational data. A wide variety of observations can be explained by this simple model, including the observed correlation between the brightest young cluster in a galaxy and the total number of young clusters. The model confirms that this can be understood as a statistical size-of-sample effect rather than a difference in the physical process responsible for the formation of the clusters, suggesting that active mergers have the brightest clusters simply because they have the most clusters. A comparison is made of different cluster disruption laws, and it is shown that the break in the $dN/d\tau$ diagram used to determine the parameters in the Boutloukos & Lamers model may be produced by incompleteness near the break point. A model of the Antennae galaxies is developed and compared with the observational data; this extends the mass-age range for cluster comparison over previous studies. An important component of our model is the use of a “two-stage” disruption process, with a very high “infant mortality” rate for the clusters with ages less than $\sim 10^6$ yr (i.e., roughly 80%–90% are lost each factor of 10 in time, $\tau$, independent of mass), and two-body relaxation, which becomes the dominant disruption mechanism at older ages, preferentially removing the lower mass clusters. Hence, in our model, stars from the dissolved clusters form the field population. We note that a 90% infant mortality rate for each factor of 10 in $\tau$ (i.e., $dN/d\tau \propto \tau^{-1}$) is consistent with all measured young cluster populations, including those in the Antennae, the Small Magellanic Cloud, and the Milky Way. In fact, the population of clusters in the Antennae can be viewed as a scaled-up version of the Milky Way in many respects, with a scale factor of roughly 1000 times the Lada & Lada sample of embedded star clusters in the local Milky Way. We find no evidence for a truncation of the Antennae cluster mass function at the high-mass end.

Key words: galaxies: individual (NGC 4038/4039) — galaxies: interactions — galaxies: star clusters

1. INTRODUCTION

The discovery of young massive clusters (YMCs) in merging and starburst galaxies has been a major catalyst for the field of extragalactic star clusters in recent years (see Whitmore 2003 and Larsen 2006 for reviews). In particular, the fact that the most massive of these young clusters have all the attributes expected of young globular clusters means that we can study the formation of globular clusters in the local universe rather than trying to understand how they formed some 13 Gyr in the past.

A natural inference one might draw from the discovery of YMCs in merging and starburst galaxies is that some “special physics” is occurring in these environments: that normal quiescent spirals can only form low-mass open clusters. However, the discovery of a small number of YMCs in normal spiral galaxies by Larsen & Richtler (1999) showed that this inference is incorrect. The formation mechanism required to make YMCs appears to be more “universal” than originally thought. Early hints that this phenomenon might be quite widespread were the discovery of YMCs in barred galaxies (Barth et al. 1995), M33 (e.g., Chandar et al. 1999a, 1999b), and the central region of our own Milky Way (Figer et al. 1999). The fact that most studies of the luminosity functions of young compact clusters have found them to be power laws of the form $\phi(L)dL \propto L^\alpha dL$ with a value of $\alpha \approx -2$ (e.g., compilations in Whitmore [2003] and Larsen [2006, 2007]; however, see § 4.2 for alternative views) provides additional evidence for the universality hypothesis. The main difference appears to be the normalization of the power law, since young active mergers typically have hundreds of YMCs, older mergers and starbursts have hundreds of YMCs, and normal spiral galaxies have between 0 and 100 YMCs.

While cluster mass functions have been determined for only a small subset of these galaxies (due to the need for cluster age estimates, which are observationally expensive), a similar power-law form has been found for the mass and luminosity functions in the Antennae [i.e., $\psi(M)dM \propto M^{\beta}dM$ with a value of $\beta \approx -2$; Zhang & Fall 1999; Fall 2004; S. M. Fall et al. 2007, in preparation] and the Small and Large Magellanic Clouds (SMC and LMC) (Hunter et al. 2003), as well as a few other galaxies (see review by Larsen 2007).

Since the number of galaxies with sufficient data to form even a meaningful luminosity function is small, Whitmore (2003) enlarged his sample by comparing the luminosity of the brightest young cluster in a galaxy with the number of clusters in the system brighter than $M_V = -9$. Figure 1 shows an updated version of Whitmore’s figure, with the primary improvements being the addition of galaxies from Tables 1–5 of Larsen (2006) and updated cluster numbers for several galaxies. Although the galaxy sample is heterogeneous, this diagram shows no evidence...
for a discontinuity in the ability of quiescent spiral galaxies to form bright star clusters, when compared with mergers and starbursts. The observed slope is consistent with the hypothesis that all of the galaxies have the same universal cluster luminosity function: a power law with an index of $-2$. This suggests that active mergers have the brightest clusters only because they have the most clusters (i.e., a size-of-sample effect). It appears to be a matter of simple statistics rather than a difference in physical formation mechanisms. A more detailed discussion of Figure 1 is included in § 3.1.

While the luminosity and mass functions of young star clusters appear to be power laws, the luminosity and mass functions of old globular clusters are peaked and are generally characterized by a mean value of $M_V \approx -7.4$ (corresponding to a mean mass of $\approx 2 \times 10^5 M_\odot$) and width of $\sigma \approx 1.4$ mag (e.g., see compilation in Ashman & Zepf 1998). Several theoretical works (Fall & Rees 1977; Vesperini 1998; Fall & Zhang 2001) have studied the effects of two-body relaxation, tidal shocks, and stellar mass loss on cluster populations. These studies have suggested that a natural way to explain this apparent evolution is via the preferential destruction of the fainter, less massive clusters, due primarily to the effects of two-body relaxation in a tidal field. Modeling of this effect has shown preferential dissolution of lower mass clusters ($M \leq 10^4 M_\odot$) within the first Gyr. Over $\approx 10$ Gyr this can introduce a turnover in an initial power-law mass distribution (e.g., Fall & Zhang 2001), similar to what is observed for ancient globular cluster systems in a number of massive galaxies.

More recently, several studies (e.g., Fall 2004; Whitmore 2004; Fall et al. 2005; Bastian et al. 2005; Mengel et al. 2005) have suggested that the majority of stellar clusters are disrupted within the first $\approx 10$ Myr of life and hence are presumably not bound to begin with. More specifically, Fall et al. (2005) find $dN/d\tau \propto \tau^{-1}$ for clusters with ages of $\approx 100$ Myr and masses of $\geq 3 \times 10^4 M_\odot$ in the Antennae galaxies, implying that roughly 90% of the clusters are disrupted each decade (i.e., factor of 10) of time $\tau$. This rapid destruction, or “infant mortality,” is the dominant contributor to cluster demographics over this period of time. A similar effect has been noted for $10^2 - 10^3 M_\odot$ embedded clusters in the Milky Way (Lada & Lada 2003). In apparent contradiction to these results for the young cluster systems in the Antennae and Milky Way, Rafelski & Zaritsky (2005) suggest $dN/d\tau \propto \tau^{-2}$ for the cluster population in the SMC. However, a reanalysis of their data (Chandar et al. 2006a) indicates that $dN/d\tau \propto \tau^{-1}$ in the SMC as well (see § 4.3 and Fig. 15).

As outlined above, previous results on the Antennae cluster system have provided some of the motivation for the framework presented in this paper. These have revealed that the mass distribution is approximately independent of age (Fig. 3 of Zhang & Fall 1999) and that the age distribution is approximately independent of mass for YMCs (Fig. 2 of Fall et al. 2005). As in these previous works, we define clusters as the compact, barely resolved objects detected in galaxies such as the Antennae (e.g., Whitmore et al. 1999). Since we cannot directly assess whether a given aggregate of stars is bound or unbound based on the photometric data alone, we make no assumption about how long a cluster might survive. The development of the model presented here allows us to extrapolate outside of the ranges used previously to establish the cluster mass, age, and luminosity relations and to simultaneously compare model predictions with the observed properties of the host galaxy. To the degree that our models fit the integrated properties of a galaxy (e.g., the total stellar mass and total luminosity) we can have some confidence that these extrapolations are reasonable.

In summary, it appears plausible that most star clusters form from a universal initial cluster mass function, which is then modified by several disruption mechanisms. This paper describes a simple toy model designed to test this framework in detail. In order to facilitate direct comparison with observations, common selection effects and data reduction artifacts are included in the model. The results are then compared with observations of the Antennae cluster system that extend the mass-age domain explored previously. An earlier version of some of this work was described in Whitmore (2004). Paper II (R. Chandar & B. C. Whitmore 2007, in preparation) in this series will extend the comparison to a number of other nearby galaxies and introduce an objective, three-dimensional classification system that helps facilitate the comparison (see Whitmore 2004). We are therefore testing a specific view, that most star formation occurs in groups and clusters of stars according to a single power-law mass function. This distribution is subsequently modified, with disrupted clusters forming the field. An alternative view is that clusters form with a distinct break or even a turnover in the mass function (e.g., Fritze-v. Alvensleben 2005; Mengel et al. 2005; de Grijs et al. 2005). We discuss the evidence for and against this possibility in § 4.2.

This paper is organized as follows: Section 2 develops the framework for understanding the observed properties of star cluster populations; § 3 presents some results and predictions of the model, in particular a simulation of the $M_V$ (brightest) versus log $N$ diagram and a comparison with observations of the Antennae galaxies; and § 4 discusses broader implications and some potential anomalies for the universal framework. Finally, in § 5 we summarize the main conclusions of this work.

2. A UNIVERSAL FRAMEWORK FOR UNDERSTANDING STAR CLUSTER DEMOGRAPHICS

In this section we develop a framework for understanding the demographics of star clusters, following the theme outlined in § 1. A simple toy model has been created using the IDL programming
language to implement this framework. The key ingredients of the model are as follows:

Initial mass function.—The initial mass function is assumed to be a power law with an index of \( \beta = -2 \). The effects of using other indexes are briefly examined in §§ 3.1 and 3.2.

1. Power law \( \psi(M) \propto M^\beta dM \).

Various cluster formation histories.—Two cluster formation histories are included: a constant rate of formation and a Gaussian burst. Different components can be combined as needed to produce the model for a galaxy. For example, the five-component model for the Antennae cluster system discussed in § 3.2.3 consists of three constant-formation components and two Gaussian components. All star formation is assumed to occur in clusters. The field stars result from the disruption of the clusters. The Bruzual & Charlot (2003) models are used for the spectral energy distributions (SEDs):

1. Constant formation in time \( (dN/d\tau = \text{constant}) \).
2. Gaussian burst \( (dN/d\tau \propto \tau_0 e^{-\Delta \tau^2/2\tau^2}) \).

Various disruption laws.—Four disruption laws are included. It is assumed that the stars from the disrupted clusters become the field stars in the galaxy:

1. Constant mass loss \( (M = M_0 - \mu_{\text{ev}} \tau) \). This is the time dependence resulting from two-body relaxation for a cluster in the tidal field of the host galaxy (i.e., a value of \( \mu_{\text{ev}} \sim 2 \times 10^{-5} \, M_\odot \, \text{yr}^{-1} \) matches the detailed calculations in Fall & Zhang [2001]; see their Figs. 1 and 4 and their eq. [6]). We also note that observations of faint globular clusters in the Milky Way (see Fig. 3 of Fall & Zhang 2001) and M87 (Waters et al. 2006) agree with this functional form.
2. Constant number loss \( (dN/d\tau \propto \tau^\gamma) \). A value of \( \gamma = -1 \) results in 90\% loss each decade of \( \tau \); this is defined as 90\% infant mortality.
3. A two-stage disruption model, incorporating constant number loss (infant mortality) for the first 10^5 yr and constant mass loss (two-body relaxation) at all ages.
4. An empirical disruption law from Boutloukos & Lamers (2003, hereafter BL03) \( \tau_{\text{dis}}(M_\odot) = t_4^{\text{dis}} (M_\odot 10^4 M_\odot)^{0.8} \), where \( t_4^{\text{dis}} \) is the disruption time for a 10^4 M_\odot cluster and \( \gamma_{\text{BL}} \) is the mass dependence of the disruption timescale.

Convolution with common observational artifacts and selection effects.—Three artifacts and selection effects are currently incorporated into the model. Section 2.5 describes the following in more detail and outlines plans for the inclusion of additional selection effects in the future:

1. Magnitude threshold.
2. Reddening and extinction.
3. Artifacts from age-dating algorithms.

As briefly mentioned in § 1, the mass and age distributions for YMCs in the Antennae, which provided some of the motivation for this work, are independent of one another (Fall 2006). This means that the joint distribution \( g(M, \tau) \) must be the product of the mass and age distributions, \( \psi(M) \) and \( \chi(\tau) \), which roughly follow the relation \( g(M, \tau) \propto M^{-\beta} \tau^\gamma \). We have adopted this approach, hence simplifying our model.

We should note that these relationships are only meant to be first-order approximations to reality, designed to develop a simple framework that captures the dominant mechanisms required to understand the demographics of star clusters in galaxies. For example, S. M. Fall et al. (2007, in preparation) show that a single power law with an index of \( \beta \sim -2 \) is a very good fit to the data for the young (\( \leq 100 \) Myr) Antennae cluster system over the mass range 10^-4–10^0 M_\odot. While there are small deviations (such as slight curvature; Whitmore et al. 1999) from a power law, these deviations are at the \( \approx 2 \sigma \) level. A similar result is found for the luminosity function (S. M. Fall et al. 2007, in preparation). The framework presented here allows us to test whether a value of \( \beta \sim -2 \) beyond this range is consistent with other available data and our assumed model.

2.1. Initial Mass Function

The choice of a power law with an index of \( \beta = -2 \) for the initial distribution of mass has already been mentioned in § 1. We address recent claims that the initial cluster mass function in some nearby galaxies deviates significantly from a power law in § 4.2. One nice attribute of adopting \( \beta = -2 \) is that the number of clusters in each decade of decreasing mass increases by a factor of 10; hence, there is equal mass in each decade (e.g., the total mass of all the clusters in the range 10^-5–10^0 M_\odot is the same as the total mass of all the clusters in the range 10^1–10^4 M_\odot).

A word about boundary conditions is in order, since integrating a power-law mass function with an index of \( -2 \) over all mass ranges would lead to an infinite mass. The lower mass limit is controlled by the masses of individual stars, roughly 0.1–100 M_\odot. The highest mass cluster formed may be controlled by statistics, namely, the very low probability of producing a cluster with more than 10^7 M_\odot, or possibly by physics if there is a physical upper limit to the allowed mass for an individual cluster. Therefore, there are approximately five to eight decades (i.e., factors of 10) of cluster masses to be concerned with. Hence, if the majority of stars are born in star clusters, each decade in initial mass of the clusters should produce roughly 10%-20% of the total mass in stars for a given galaxy.

One may argue that the mass of stars in globular clusters in the Milky Way is only about 10^4 of the total mass of the Galaxy halo, rather than the \( \approx 40%–60% \) expected from the argument above and the fact that globular clusters span roughly three decades in mass (from \( \approx 10^4 \) to \( \approx 10^7 M_\odot \)). However, as we shall see, this is consistent with the dominating role that cluster disruption plays in determining the demographics of clusters. Hence, most of the stars initially formed in clusters eventually end up in the field.

We also note that our observations in galaxies like the Antennae only allow us to directly determine the initial mass function over about two decades of mass, from \( \approx 10^4 \) to \( 10^6 M_\odot \). However, to the degree that our models fit the integrated properties of a galaxy (e.g., the total stellar mass and total luminosity) we can have some confidence that extrapolations outside of these ranges are reasonable.

2.2. Cluster Formation Histories

Our model accommodates linear combinations of constant cluster formation and Gaussian bursts to reproduce observations of different galaxies. For example, typical spiral galaxies can be modeled by an initial burst of star formation to produce the ancient globular cluster population roughly 13 Gyr ago, combined with a constant rate of formation since that time.

We begin our examination of the cluster formation history with perhaps the simplest model: constant formation over the past 13 Gyr with no cluster disruption. Figure 2a shows the log \( (M/M_\odot) \) versus log \( \tau \) diagram for this model, using our standard initial mass function, which has a power law with an index of \( \beta = -2.0 \).

We note that, similar to our choice of \( \beta = -2 \) for the power-law index for the mass distribution, the choice of a constant formation rate simplifies the accounting, since there are a factor of 10
more clusters in each decade of time, by definition. Similarly, the boundary conditions in the temporal dimension are fixed, since clusters (and the brightest stars within them) require \( \approx 1 \) Myr to form. Hence, there are only about four decades in \( \tau \) to integrate over (i.e., \( \sim 1 \) Myr to \( \sim 10 \) Gyr).

Observations of cluster systems are limited by the quality of the data. While cluster detection is a complicated function of object brightness, local background level and complexity, and object size, the most basic observable parameter that sets the limit for a given data set is the brightness of an object (i.e., the magnitude threshold). Figure 2 shows the age vs. mass distribution, which has been drawn randomly from a power-law mass function that has an index of \( \beta = -2 \). The middle panel shows the age vs. mass diagram for the same model but with a \( F \) magnitude limit imposed to mimic observations. The bottom panel shows the \( V \)-band luminosity vs. the age distribution of the synthetic cluster population, including the magnitude limit.

Fig. 2.—Synthetic cluster population formed at a constant rate (and with no disruption) for clusters with mass \( > 10^3 M_\odot \). The top panel shows the age vs. mass distribution, which has been drawn randomly from a power-law mass function that has an index of \( \beta = -2 \). The middle panel shows the age vs. mass diagram for the same model but with a \( F \) magnitude limit imposed to mimic observations. The bottom panel shows the \( V \)-band luminosity vs. the age distribution of the synthetic cluster population, including the magnitude limit.

more directly related to the observations and serves as a reminder that the magnitude threshold (defined by the lower edge seen horizontally in Fig. 2c) translates to a diagonal line in Figure 2b due to the dimming of the clusters with age.

2.3. Inclusion of Various Cluster Disruption Laws

At present, four idealized disruption laws have been incorporated into the toy model, as outlined in § 2. In Figure 3 we plot age versus mass distributions for simulated clusters for three of the laws (the fourth, the two-stage disruption model, is a combination of two of the other laws and is discussed in more detail in § 3.2). Each model shown in Figure 3 begins with a constant rate of cluster formation, which is subsequently modified by the various disruption laws described below. Figure 4 shows the corresponding age distributions for the 12 models presented in Figure 3.

In the constant mass-loss models (Figs. 3 and 4, left columns) a fixed amount of mass is removed from each cluster during each unit of time. A value of \( \mu_{\text{loss}} \sim 10^{-5} M_\odot \text{yr}^{-1} \) results in reasonable values (e.g., a \( 10^4 M_\odot \) cluster is disrupted in \( 10^7 \) yr). The resulting mass distribution of surviving clusters after a Hubble time looks similar to a typical distribution of globular clusters, with a turnover around \( 10^5 M_\odot \), as shown in Figure 5. The constant-mass-loss disruption law was motivated by the detailed calculations made in Fall & Zhang (2001), which shows that two-body relaxation is the dominant destruction mechanism for clusters older than about \( 10^9 \) yr, and that the mass loss of a cluster via two-body relaxation is linear with age. The peak in the mass distribution will also increase linearly with time, as shown in the \( 10^{-5} M_\odot \text{yr}^{-1} \) panel of Figure 3 (which mirrors the result in Figs. 3–11 of Fall & Zhang 2001). We note that two-body relaxation alone destroys very few clusters in the first \( 10^8 \) yr (e.g., Baumgardt & Makino 2003), which is contrary to the important role infant mortality plays in the evolution of cluster systems.

Models of constant number loss, or infant mortality, reduce the existing population of clusters by a fixed percentage (e.g., 0%, 50%, 80%, or 90% \( \text{each decade in } \tau \), corresponding to slopes of \( \gamma = 0.0, -0.3, -0.7 \), and \( -1.0 \) for the age distribution). This type of (mass-independent) cluster disruption model is motivated by the age distribution of clusters observed originally in the Antennae (as described in detail in Fall et al. 2005) and further discussed in § 2.4. Figures 3 and 4 (middle columns) show illustrative examples for different values of constant number loss. Over the first \( \sim 10 \) Myr in the life of a cluster, photoionization, stellar winds, and supernovae can inject sufficient energy into the intercluster medium to remove much of the gas. This gas removal can unbind the cluster (e.g., Hills 1980; Lada et al. 1984; Boily & Kroupa 2003a, 2003b; Fall et al. 2005). Following this stage, stellar evolution will continue to significantly remove mass from the cluster for longer periods of time (e.g., Applegate 1986; Chernoff & Weinberg 1990; Fujishige & Heggie 1995).

Figure 3 (middle column) shows that as the fraction of the disrupted population is increased, the predicted ratio of young to old clusters increases dramatically. The age distributions for cluster populations affected by different levels of infant mortality are shown in Figure 4. The top middle panel shows that for a constant rate of cluster formation with no disruption, a mass-limited cluster sample would have a flat distribution in the log \( \text{dN/d} \tau \) versus log \( \tau \) diagram. Higher values of infant mortality would cause steeper age distributions, as shown by the filled circles in the other three middle panels in Figure 4. The predicted age distributions in Figure 4 are also compared with the empirical age distribution for the Antennae galaxies (presented in Fall et al. 2005). In this figure it is obvious that the 90% constant...
number loss (i.e., infant mortality–type disruption) gives the best match between models and observations, although the 80% model is also reasonable given the uncertainties (e.g., the data point for the youngest clusters may be artificially high due to a recent burst of star formation in the “overlap” region; see § 2.4).

The main shortcoming of this model is that we do not expect it to be relevant past a few $\times 10^8$ yr, after which two-body relaxation should begin to dominate (Fall et al. 2005). We also note that from an empirical standpoint, the 90% disruption rate has only been shown to be roughly linear out to about this age (i.e., Fig. 2 from Fall et al. 2005). After this point stellar contamination begins to affect the data set.

The third cluster disruption law (not shown) is a two-stage disruption process, which includes both the constant number loss (infant mortality) and the constant mass loss (two-body relaxation) models. While this two-stage process has been discussed previously (e.g., Fall et al. 2005), this work provides the first detailed application to cluster systems. We note that the resulting disruption law is therefore mass-independent for young ages and mass-dependent for older ages. This is our primary tool when building our model for the Antennae and hence is described in more detail in § 3.2.

The fourth cluster disruption law is described in detail in BL03 (also see Lamers et al. 2005a). They fit the formula $t_{\text{dis}}(M_\odot) = t_{\text{dis}}(M_\odot/10^4 M_\odot)^{0.6}$, where $t_{\text{dis}}$ is the disruption time for a $10^4 M_\odot$ cluster and $\gamma_{\text{BL}} \approx 0.6$. The value of $t_{\text{dis}}$ is derived based on apparent bends in the log $dN/d\tau$ versus log $\tau$ diagram that, as we shall see, vary dramatically from galaxy to galaxy.

Lamers et al. (2005b, hereafter LGPZ05) use this technique to determine a very long characteristic disruption time ($t_{\text{dis}} \approx 8 \times 10^7$ yr) for clusters in the SMC. Our simulations for LGPZ05 show dramatically from galaxy to galaxy.

The disruption time derived by LGPZ05 for M33 ($t_{\text{dis}} = 6 \times 10^8$ yr) is intermediate between the values given for the SMC, M33, and M51. We note that the estimated values of $t_{\text{dis}}$ decrease monotonically as a function of distance for the SMC, M33, and M51.
M51. This led us to investigate whether the bend in the age and mass diagrams that LGPZ05 use to derive $t_{\text{dis}}^4$ could be an artifact caused by the magnitude threshold (which is typically distance-dependent), combined with various selection effects.

In Figure 6 we show the resulting age distributions for an artificial population of clusters formed continuously and subject to a magnitude threshold, with three different mass limits assumed. If a sufficiently high mass limit is chosen, the resulting age distribution is flat, as expected, since it is not affected by completeness or disruption of clusters. The diagram also shows that as lower mass limits are used, the resulting age distributions become incomplete due to the magnitude threshold, resulting in an apparent break. Similar results are found using the infant mortality or two-stage disruption laws but with sloping $dN/d\tau$ diagrams.

One prediction of this analysis is that two galaxies at the same distance, observed using similar magnitude thresholds, should result in similar estimates of $t_{\text{dis}}^4$ (i.e., the observational artifacts would result in the same position of the bend). This appears to be the case, since de Grijs & Anders (2006) find the same value of $\log t_{\text{dis}}^4$ for the LMC ($9.9 \pm 0.1$) as LGPZ05 found for the SMC ($9.9 \pm 0.2$). We conclude that currently, there are insufficient data to construct a sample of nearby galaxies that cover a range of distances and local densities needed to definitively establish whether the derived $t_{\text{dis}}^4$ timescales result from biases or density.

To test our suspicion that this may be a magnitude threshold bias, we used our Antennae data set according to the prescription of BL03. When restricted to dynamic ranges of age and mass in which the data are approximately complete, no breaks are seen in the age and mass distributions (i.e., they are smooth power laws). If age and mass distributions are plotted over larger dynamic ranges in which we know completeness is an issue, breaks are seen that imply even shorter $t_{\text{dis}}^4$ timescales than found for M51. Hence, in both this case and in the artificial cluster simulation described above, we know that the apparent breaks are caused by selection effects and incompleteness. We predict that as age and mass distributions become available for other galaxies it will become evident that any apparent breaks are roughly dependent on the depth of the cluster survey (i.e., the distance) and hence caused by an artifact.

### 2.4. The Need for Rapid Cluster Disruption (i.e., Infant Mortality)

The age-luminosity distribution for the constant cluster formation model (with no disruption but with a magnitude threshold; Fig. 2c) looks much different than the observed distribution for the Antennae cluster system (Fig. 7). Instead of having the highest density of clusters per unit $\log \tau$ at older ages, the observations have the highest density of young clusters. This is demonstrated more clearly in Figure 2 from Fall et al. (2005). They find that a

---

**Figure 4**—Age distribution (with arbitrary offset) for the 12 simulated models shown in Fig. 3 (filled circles). For comparison, we show the completeness-corrected age distribution for clusters in the Antennae galaxies with masses $\geq 10^5 M_\odot$ (open circles).
plot of log (dN/dt) versus log t for the Antennae clusters has a slope that declines with a value of \( \approx -1 \) out to an age of a few \( 10^7 \) yr, rather than being flat, as would be the case for constant cluster formation with no disruption (see Figs. 3 and 4). Hence, if the Antennae has had a history of roughly constant cluster formation during this period, clusters must be disrupted at a rate of \( \approx 1 \). It is not possible with the current observations to determine whether the same relationship extends beyond a few \( 10^8 \) yr for masses \( M_\odot \), due to potential stellar contamination.

The same result is apparent from an examination of Figure 4. Only the 80% and 90% (i.e., slopes of \(-0.7\) and \(-1.0\) in the log (dN/dt) versus log t diagram) infant mortality laws agree with the data. None of the other disruption mechanisms come close to reproducing the large fraction of observed young clusters.

Perhaps the high density of young clusters is due to a recent burst of star formation in the Antennae? This is likely to be true, of course, since the galaxies are currently merging and regions like the overlap region have large numbers of very young clusters. However, if we look at each of the Wide Field Planetary Camera 2 (WFPC2) chips independently (Fig. 8), we find that the log luminosity versus log (t/yr) diagrams for all the chips look quite similar, with the largest numbers of clusters always having ages \( < 10^7 \) yr! Since it is not possible for the galaxy to synchronize a burst of star formation over its entire disk (i.e., assuming a “signal speed” of \( \approx 30 \) km s\(^{-1}\) [Whitmore et al. 1999] requires approximately 500 Myr for one side of the galaxy to communicate with the other side 10 kpc away), we can only conclude that the vast majority of the clusters are being disrupted almost as fast as they form, making it appear that the largest number of clusters is always the youngest.

If one looks closely at Figure 8, or at the corresponding log (dN/dt) versus log t diagram shown in Figure 9, there actually are small differences on the various WFPC2 CCDs due to variations in the local cluster formation histories. For example, the number of clusters at the youngest ages is especially high for
the overlap region (X pixel values < 440 on WF3, i.e., the dusty overlap region; see Fig. 5a from Whitmore et al. 1999), compared to regions where there appears to be less recent formation (e.g., X pixel values > 440 on WF3, which is the dust-free region containing several clusters with ages > 100 Myr). However, these are secondary effects; the primary correlation is the rapid decline in the number of clusters as a function of age for all of the CCDs. We return to this point in § 3.2.

This behavior is not confined to mergers and starburst galaxies, such as the Antennae. In typical spiral galaxies (e.g., M51, Fig. 10 of Bastian et al. 2005; M83, Harris et al. 2001; LMC, Hunter et al. 2003; NGC 6745, de Grijs et al. 2003a; M51 and M101, R. Chandar & B. C. Whitmore 2007, in preparation; M33, Chandar et al. 1999a, 1999b; SMC, Rafelski & Zaritsky 2005), the observational data almost universally show that the density of clusters at a given mass, per unit log \( \log \rho / C^2 \), is roughly constant. The fact that such apparently different galaxies have such similar cluster age distributions is one more demonstration of the universal nature of cluster demographics and highlights the dominant role cluster disruption appears to play in all galaxies.

The phenomenon of star cluster disruption is not surprising. We are familiar with the process both observationally (e.g., Wielen 1971; Bica et al. 2001; Rockosi et al. 2002) and theoretically (e.g., Fall & Zhang 2001; Vesperini & Zepf 2003; Baumgardt & Makino 2003). What is surprising is that massive, compact clusters, which might be expected to withstand many of the disruption mechanisms that are important for low-mass, diffuse clusters, appear to have disruption timescales that are comparable to those of low-mass clusters.

In particular, we note that this is dramatically different from the canonical picture for open clusters developed by Wielen (1971), which showed essentially no cluster disruption for the first \( \approx 100 \) Myr (see Fig. 13 of Wielen 1971). However, it does match the recent results from Lada & Lada (2003) for the young embedded clusters in the solar neighborhood, since they find that \( \sim 90\% \) of these clusters are disrupted in the first 10 Myr.

A natural explanation for this high infant mortality rate is that most of the YMCs may be unbound when they form (or rapidly become unbound) and dissolve on the order of a few crossing times (e.g., see discussion in Fall et al. 2005). Briefly, the energy and momentum input from massive young stars to a protocluster are roughly proportional to the number of massive stars and to its mass. Therefore, such an internal process would be expected to
operate roughly independent of cluster mass, as well as of host galaxy properties.

2.5. Selection Effects and Artifacts

A number of selection effects and artifacts must be taken into account before we are able to compare our synthetic data with the observations. The first selection effect is a magnitude threshold, as shown in Figure 2b. As discussed in §2.3, this magnitude threshold, when coupled with the dimming of clusters with age, can cause serious difficulties when attempting to determine mass and age distributions over wide dynamic ranges.

The second effect that we consider is extinction and reddening due to the presence of dust. Dust is most closely associated with very young clusters. This has been modeled by adopting the Mathis (1990) coefficients for the various filters and a time dependence defined by $A_V(\log \tau) = -2.5 \log \tau + 18$ (which results in a mean value of $A_V = 3$ mag at $\log \tau = 6.0$, $A_V = 1.0$ mag at $\log \tau = 6.8$, and $A_V = 0$ mag at $\log \tau \geq 7.2$). Noise is added by multiplying the $A_V$ value by a random number from 0 to 1 and then adding a mean value of $A_V = 0.2$ to represent patchy foreground dust. The time-dependent component is consistent with the extinction distribution $A_V(\log \tau)$ found for the Antennae by Whitmore & Zhang (2002) and Mengel et al. (2005). Using other time-dependent extinction laws (e.g., Charlot & Fall 2005), we find that there is larger scatter about this relation for young clusters that have strong Hα emission, and we introduce an additional age-dependent photometric uncertainty for simulated clusters younger than $\log \tau = 6.8$ yr in this filter. The resulting input age versus output (i.e., fitted) age is shown in Figure 10 for clusters more massive than $3 \times 10^4 M_\odot$. We find that regardless of the mass cut we use (up to $10^6 M_\odot$), the gap in the age range 10–20 Myr does not go away.

Note that this appearance contrasts with Figure 2 in Gieles et al. (2005), where the gap appears to be present only at lower masses but then fills in for higher mass clusters. We suggest that this is due to an underestimate in their models of the true photometric uncertainties at higher masses. As in this work, the Gieles et al. (2005) simulations use the SED models for both the input and output. Hence, if the photometric uncertainties are underestimated, the age-dating algorithm will always find the
“right” answer. In reality, there will be a mismatch between the models and the observations, resulting in a 10 Myr gap at all masses, as seen in the actual data.

Figure 10 allows one to see in more detail where the data from the “gap” are repositioned. While there is a clear, well-defined correlation between the input and output ages, the reason for the 10 Myr artifact can be seen as a repositioning of many of the clusters with ages between log τ = 7.0 and 7.2 to an age of log τ = 7.3.

A second prominent feature in Figure 10 is the region of filled circles between log τ of 6.0 and 6.4 yr. This feature implies that clusters with intrinsic ages between 1 and 2.5 Myr cannot be separated using integrated colors and narrowband photometry alone. This is not surprising, given that the Hα emission predicted by models “saturates” at these very young ages; hence, we see a large concentration there.

Other potentially important selection effects and artifacts are limited spatial resolution (i.e., confusion between clusters and stars), blending of objects, and stochasticity for very young low-mass clusters (e.g., below 10^4 M⊙, the chance of having a single O star becomes less than unity, which can affect the photometric colors dramatically; see, e.g., Cervino et al. 2002; Ubeda et al. 2007; Oey & Clarke 1998). These and other effects will be simulated in more detail in future versions of the model.

3. RESULTS

Two topics are examined in this paper in order to exercise the model described in the previous section. The first looks at whether the model can account for the $M_V$ (brightest) versus log $N$ diagram, as shown in Figure 1. The second topic is the development of a toy model for the Antennae galaxies. Paper II will extend the analysis to data sets for a number of other nearby galaxies (e.g., M101, M51, etc.).

3.1. A Monte Carlo Simulation of the $M_V$ (Brightest) versus log $N$ Diagram

As briefly described in § 1, the $M_V$ (brightest) versus log $N$ diagram (Whitmore 2003) provided much of the initial impetus for the development of the framework developed in the present paper, since it was consistent with the idea that mergers, starbursts, and normal spiral galaxies all share a common universal power-law luminosity function $\phi(L)$ for clusters. Larsen (2002) further developed this basic idea by performing Monte Carlo simulations of cluster populations. This study supports the conclusion that the magnitude of the brightest cluster is controlled by statistics rather than by physics. Larsen (2002) shows that the observed scatter in the relation can also be explained by statistics. Several other studies have also begun to address the important role this size-of-sample effect can play in the observed properties of star cluster systems (e.g., Billet et al. 2002; Hunter et al. 2003; Weidner et al. 2004).

Figure 1 shows an updated version of the database from Whitmore (2003). Several more galaxies have been added, primarily from the Larsen (2006) compilation, and the values have been updated for individual galaxies for which more recent work has become available.

Here we present a test of the power-law index $\alpha$ over absolute magnitudes ranging from $M_V$ of $-9$ to approximately $-16.5$. Figure 11 shows a comparison of the observational data presented in Figure 1 with sets of Monte Carlo simulations of 500 cluster populations, which are randomly drawn from a power-law luminosity function with values of the index $\alpha$ ranging from $-1.5$ to $-3.0$ and ages between 1 and 100 Myr (i.e., the brightest clusters are essentially always found in this age range). The 90%
infant mortality law is assumed for this particular simulation. Several other star formation histories have also been tested (e.g., constant with time, Gaussian bursts at various times) with similar results. An observational cutoff brighter than $M_V = -9$ has been imposed on both the observations and the simulations to approximate the stellar contamination limit.

The figure shows that values of $\alpha$ equal to $-1.5$ and $-3.0$ are clearly ruled out. Values of $\alpha$ between $-2.0$ and $-2.4$ all fit the observations to some degree, with values of $-2.0$ matching the slope but showing an offset of about 1 mag and values around $-2.4$ showing no offset but having a flatter slope than the observations. Values of the fits are included in Table 1. It is likely that part of the offset is due to many of the data sets not being complete to $M_V = -9$; hence, the value of the number of clusters in the sample, $N$, is underestimated. In principle, the slope $\alpha$ should only reflect the underlying luminosity function. However, different selection effects (such as the imposed cutoff of $M_V = -9$) can affect the measured slope. For this reason, we consider the $\alpha = -2.2$ fit as the best compromise between matching the slope and minimizing the offset, and we estimate an uncertainty of 0.2.

Figure 12 shows a comparison between the scatter in the observations and the simulation for our adopted $\alpha = -2.2$ model. Based on the similarities between the two, we conclude, as Larsen (2002) did, that the observational scatter can be explained just by statistics. There is no indication that any "special physics" is needed to explain the differences between, for example, the cluster systems of mergers and spiral galaxies.

We conclude that both the correlation between $M_V$ (brightest) and $\log N$ and the scatter in the correlation can be explained by statistics if all star-forming galaxies have the same universal luminosity function. The only difference is the normalization, with mergers having thousands of clusters brighter than $M_V = -9$ and spirals having tens of such clusters. One possibility for the much larger number of clusters in mergers is that the conditions for making YMCs are globally present, while in spiral galaxies they are only locally present (i.e., in the spiral arms; Whitmore 2003).

### 3.2. A Model of the Antennae

#### 3.2.1. Background

Because of its proximity (19.2 Mpc), large collection of YMCs, and extensive set of multiwavelength observations (e.g., Zhang et al. 2001), the prototypical merger NGC 4038/4039 (the Antennae) currently represents the best data set for testing the model described in this paper. Here we briefly summarize the observations and analysis from previous works.

The HST observations of the Antennae galaxies are described more fully by Whitmore et al. (1999). The images were taken in 1996 January by the WFPC2 with each of the broadband filters F336W ($U$), F439W ($B$), F555W ($V$), and F814W ($I$) and with the narrowband F658N (Hα) filter. Approximately 14,000 point-like objects (stars and clusters) were detected.

We estimate both the age ($\tau$) and extinction ($A_V$) of each cluster by comparing the observed magnitudes in the five bands with those from stellar population models, as described in Fall et al. (2005). More specifically, we use the Bruzual & Charlot (2003) models with solar metallicity and the Salpeter initial mass function. While this metallicity is appropriate for comparison with young stellar populations (≤1 Gyr), estimates of ages for older stellar populations should be viewed with caution, as they suffer from the well-known age-metallicity degeneracy and likely have abundances lower than the solar value. Because dust is prevalent and patchy, it is important to estimate the internal effects of dust individually for each cluster. However, for compact star clusters in dusty regions it is not obvious that either a simple foreground screen or an attenuation law that assumes gas and dust mixed with stars (e.g., Calzetti et al. 1994) is appropriate. We therefore experimented with both a Galactic-type extinction law (Fitzpatrick 1999) and the absorption curves determined from starburst galaxies by Calzetti et al. (1994) for the age-dating procedure. The main difference between the two is the value of $R_v$, which is 3.1 for Galactic extinction laws and 4.05 for the Calzetti absorption law. We found relatively minor differences in the results with little change in the derived cluster ages, and

---

**TABLE 1**

| Power-Law Index $\alpha$ | Slope $M_V$ (log $N = 1.5$) | St. Dev. of Residuals $\sigma$ |
|--------------------------|-----------------------------|-------------------------------|
| $-1.5$                   | $-1.9$                      | 16.0                          | 1.2                          |
| $-2.0$                   | $-2.4$                      | 13.6                          | 1.3                          |
| $-2.2$                   | $-2.2$                      | 12.9                          | 1.3                          |
| $-2.4$                   | $-1.9$                      | 12.4                          | 1.2                          |
| $-3.0$                   | $-1.3$                      | 11.4                          | 0.9                          |
| Data*                   | $-2.3$                      | 12.4                          | 1.0                          |

*a The data for 40 galaxies have been collected from the literature, as described in the text.
masses that were 10%–20% higher for clusters younger than 10^7 yr when the Calzetti law was assumed.

As briefly described in § 1, our earlier analysis of the Antennae data set provided some of the motivation for this paper. Of particular relevance for this work, Fall (2006) notes that the mass and age distributions over limited ranges for the Antennae cluster system are independent of one another. This means that the joint distribution g(M, τ) must be the product of the mass and age distributions, ψ(M) and χ(τ), which roughly follow the relation

\[ g(M, \tau) \propto M^3 \tau^{-1} M^{-2} \tau^{-1} \text{ for } \tau \lesssim 10^7 (M/10^4 M_\odot)^{1.3} \text{ yr,} \]

(1)
as used in § 2 to define the approximate mass-age distributions of the YMCs.

In general, the luminosity function is not independent of the mass and age distributions. However, under certain circumstances (e.g., when the mass distribution is a power law and is independent of the age distribution) the luminosity and mass functions can both be power laws with the same exponent. This appears to be the case for the Antennae, with ω(L)dl ≈ L^3dl; α ≈ -2 for -14 < MV < -9 (Whitmore & Schweizer 1995; Whitmore et al. 1999); and ω(M)dM ≈ M^2dM, with β ≈ -2 for 10^4–10^6 M_☉ (e.g., Zhang & Fall 1999; S. M. Fall et al. 2007, in preparation).

Our purpose here is not to revise the previous work but rather to extrapolate to both higher and lower mass ranges. Applying these relations, the model allows us to take the next step and move beyond the observed cluster system and compare predictions with the integrated properties of a galaxy (e.g., the total stellar mass and total luminosity). For example, to the degree that our models fit the integrated galaxy properties, we can have some confidence that extrapolations outside of the original age and mass ranges (as specified in eq. [1]) are reasonable.

In addition, we can test a variety of the other assumptions that go into the model. Does our assumption that all stars form in clusters, and that the field stars are the remnants of disrupted clusters, give values for the luminosities of the clusters and field that are consistent with observations? Can we predict other distributions (or different projections of the data, e.g., color-color and color-magnitude diagrams) that were not used in the original determination of α, β, and γ above? Does an extrapolation to the highest observed mass support the recent suggestion that the Antennae system has a fixed upper cluster mass? While there is already some evidence that parts of the model are consistent with observations for a few other nearby galaxies (see, e.g., § 1 and Figs. 1, 11, 12, and 15), in Paper II we will extend the analysis to a larger sample to better test the universality of the basic framework.

3.2.2. Four-Component Model of the Antennae

We begin by constructing a model for the Antennae that assumes four cluster formation components, which are derived from previous information (Whitmore et al. 1999). The four components are as follows: Component 1 is an initial Gaussian burst with an age of 13 Gyr and σ of 1 Gyr that formed the population of old globular clusters. Component 2 is a constant rate of cluster formation over the past 12 Gyr similar to the cluster formation rate in typical spiral galaxies. Component 3 is a Gaussian burst of cluster formation 500 ± 100 Myr ago that was triggered when the galaxies first encountered each other and the long tidal tails were ejected. Component 4 is an increased rate of cluster formation during the past 100 Myr responsible for most of the recent star/cluster formation. The last two components are also motivated by the simulations of the dynamical interaction between the two galaxies (Barnes 1988), which suggested that the galaxies first encountered each other a few hundred Myr ago and are now in their second encounter. Hence, the four-component model essentially has no free parameters. It is an attempt to construct a model using previous determinations of β, γ, the cluster formation history of the Antennae, and an assumption that the power laws continue beyond the ranges in which they have been observed directly.

The observations and derived properties of the Antennae are shown in Figure 13 (left column). In order to restrict the Antennae sample to actual star clusters and not individual stars, we make a cut in absolute luminosity at MV = -3.9 (see Whitmore et al. 1999). We also impose a mass limit (≥3 × 10^4 M_☉). Note that in the age versus mass figure of the Antennae clusters this mass limit begins to show incompleteness for clusters starting at ages of a few ×10^6 yr.

The model uses a value of β = -2.0 and the two-stage disruption law with 90% infant mortality, as mentioned above. This four-component Antennae model is shown in Figure 13 (middle column). We remind the reader that the apparent gap in the age distribution around 10 Myr (i.e., Figs. 7 and 10) is due to artifacts from the age-dating algorithm. While our modeling of this artifact (§ 2.5) is able to subjectively reproduce the major features, some of the small-scale remaining differences (e.g., the narrowness of the 10 Myr artifact) are still not very well represented.

There is reasonable agreement between the four-component model and many of the diagnostics and observations [e.g., the log (dN/dτ) versus log τ diagram] right out of the box (i.e., using previous determinations of β, γ, and the cluster formation history of the Antennae). However, a closer look reveals three issues: (1) the lack of intermediate-age clusters around 100–300 Myr (besides being evident in the luminosity-age and mass-age diagram, this deficit can also be seen in the color-color diagram as the slight misplacement of the small enhancement around U - B = 0.0 and V - I = 0.5 mag; see also Fig. 17 from Whitmore et al. 1999); (2) a slight deficit in the number of young clusters with ages of 1–10 Myr; and (3) a slight enhancement in the predicted number of very luminous and/or very massive clusters.

The first difference can be fixed if we move the 500 Myr burst to an age of about 200 Myr. There is independent evidence for this in the numerical modeling of the Antennae (Barnes 1988). The second difference is the need for an enhancement at very young ages (i.e., 1–10 Myr). This is probably related to the enhancement we see in the overlap region (on WF3, as noted in § 2.4). The third difference suggests either the possibility of an upper mass cutoff for the clusters or a slightly steeper initial mass function (i.e., β = -2.1 rather than -2.0; these possibilities are explored in § 3.3).

3.2.3. Five-Component Model of the Antennae

Our second attempt at making a model takes into account the shortcomings of the four-component model mentioned above. Because of the addition of a 0–10 Myr component (to address the second shortcoming), we now refer to this as the five-component model. In addition, we relax the constraint that only the canonical values of γ = -1 (i.e., the coefficient describing the infant mortality law) and β = -2.0 must be used.

The four-component model had a total galaxy luminosity that is considerably too high (i.e., MV_galaxy = -23.0, compared with the estimated MV_galaxy = -21.7 for the Antennae; Whitmore
et al. 1999). We therefore used an 80% infant mortality law (i.e., $\gamma = -0.7$) for all components of our five-component Antennae model, which results in a smaller total galaxy luminosity and mass. This is because the predicted cluster population is normalized to the observed Antennae cluster system, so that the lower rate of cluster disruption implies that fewer clusters were formed initially. This also creates a smaller number of very young clusters, which is accounted for by increasing the 1–10 Myr component slightly. From Figure 4 it appears that the 80% and 90% infant mortality laws both give reasonable fits. Finally, we change the power-law mass index $\beta$ from $-2.0$ to $-2.1$, as justified in § 3.3. These two changes reduce the total galaxy mass to $\sim 4 \times 10^{10} M_\odot$ and total $M_{V,\text{galaxy}}$ to $-22.2$ (much closer to the observed value of $-21.7$; see Table 2). The results from the final five-component model are shown in Figure 13 (right column).

As an additional consistency check, we compare the predicted UV luminosity in clusters from our five-component model with the observed luminosity. The model predicts that $\sim 7\%$ of the UV luminosity should come from clusters. This is quite similar to the
value of ~9% measured for the Antennae cluster system by Whitmore & Zhang (2002).

While we could have continued to tweak various parameters in our model to exactly match the observations of the Antennae, we do not think this approach is useful, since there are a large number of potential variables, many of them coupled. Instead, our approach has been to only make changes where there is an independent indication that a value should be changed (e.g., the 500 Myr component should be changed to 200 Myr; see previous section). The only exceptions are the variables $\beta$ and $\gamma$, which we allowed to vary (within the uncertainties) in order to address the discrepancy between the predicted and observed total galaxy mass and luminosity.

### 3.3. Is There an Upper Mass Cutoff for Clusters in the Antennae?

Is there a physical limit at which star clusters in the Antennae can form, or does the most massive cluster just result from

---

**TABLE 2**

**Final Results of Toy Model for the Antennae**

| Galaxy Property | Antennae Galaxy | Total | 1 (const) | 2 (GC) | 3 (200 Myr) | 4 (1--100 Myr) | 5 (1--10 Myr) |
|-----------------|-----------------|-------|-----------|---------|-------------|---------------|--------------|
| Stellar mass    |                 | 5E10  | 3.8E10    | 1.2E10  | 1.2E10      | 6.4E9         | 4.6E9        | 2.5E9        |
| Total $M/\odot$ |                 | −21.7 | −22.2     | −18.7   | −18.1       | −20.1         | −21.1        | −21.6        |

**Notes.**—Component 1: Continuous creation over the lifetime of the galaxy (1 Myr-13 Gyr); Component 2: Ancient Gaussian burst at 12.6 ± 1 Gyr (i.e., old globular cluster population); Component 3: 200 ± 100 Myr Gaussian burst (i.e., initial encounter); Component 4: 1–100 Myr continuous creation; Component 5: 1–10 Myr continuous creation (recent burst observed on CCD WF3).

---

**Fig. 14.**—Mass distributions for the first, third, and 10th most massive clusters produced in 1000 Monte Carlo simulations. The total number of clusters used in each simulation is matched to the observed mass distribution in the range $4.5 \leq \log M/\odot \leq 5.5$ for the young Antennae cluster population ($10^6 \leq \log \tau \leq 10^8$). The three columns show simulations for $\beta$ values of −2.0, −2.1, and −2.2, and the dashed lines show the measured mass for the Antennae cluster system. The match between simulations and measurements shows that a value of $\beta = -2.1$ provides good fits without the need for a physical cutoff to the allowed cluster masses in the Antennae.
One example of a potential anomaly, the very bright cluster in NGC 1569, has already been discussed extensively in the literature (e.g., Whitmore 2003; Billet et al. 2002; Larsen 2002). We first note that this cluster is no longer as discrepant as previously believed, since the number of detected clusters in the sample of Hunter et al. (2000; N = 18 with M_\text{F} < −9) is larger than in the original study by de Marchi et al. (1997; N = 7). In § 3.1 we find that the scatter predicted by statistics is consistent with the observed scatter. The position of NGC 1569 in Figure 1 using the new Hunter data is ≲2σ (predicted scatter) from the mean relation. Hence, the point is no longer very discrepant.

We should, however, keep in mind that subtle selection effects can alter the results for individual galaxies. For example, Billet et al. (2002) and Whitmore (2004) both point out that in a sample of star-forming dwarf galaxies, the ones with the brightest clusters (i.e., the 1–2σ high points) are more likely to have drawn attention to themselves and to have been observed than dwarf galaxies that have not formed bright clusters (i.e., the 1–2σ low points).

The second apparent anomaly we discuss is the lack of intermediate-age globular clusters in the Milky Way. If the majority of stars form in clusters, and the initial cluster mass function is universal, then the fact that there are many intermediate-age stars (e.g., the Sun) implies that there should be some intermediate-age (i.e., a few billion years old) globular clusters.

To some extent this may just be a matter of semantics. The tail of the distribution of old “open” clusters may simply be intermediate-age globular clusters, as suggested in Whitmore (2003). While there are only a handful of known old open clusters, we should keep in mind that the catalogs of open clusters are only complete within a distance of about 1–2 kpc of the Sun; hence, the volume for the entire Milky Way is roughly 100 times larger. There are certain to be many more clusters in this category hidden behind the dust in the plane of the Milky Way. Indeed, in neighboring spiral galaxies, where we are able to see the entire disk, intermediate-age globular clusters are now being found (e.g., M31 [see Puzia et al. 2005; Beasley et al. 2004; Burstein et al. 2004; Li & Burstein 2003; and Barmby et al. 2000; but see Cohen et al. 2005 concerning the inclusion of asterisms in some of these compilations] and M33 [see, e.g., Chandar et al. 2002, 2006b]).

We might also note that due to a combination of the various disruption mechanisms and the completeness limit, the number of observable intermediate-age clusters is expected to be relatively low. For example, in our five-component simulation of the Antennae, only ∼10 clusters are predicted in the range 1–10 Gyr with M_\text{F} > −9, and only 8 clusters are observed in this region. The enhancement of clusters with ages of ∼13 Gyr (i.e., old globular clusters) in the model and in the data is due to a strong initial burst, which is expected to produce roughly the same number of clusters as are produced in the next ∼10 Gyr.

The third “anomaly” that we address is the apparent difference between the metallicity distribution of Galactic field stars, with a broad peak around [Fe/H] ≈ −0.8, and the metallicity distribution of the globular clusters, which is bimodal with peaks around [Fe/H] ≈ −1.3 and −0.75 for the Galaxy (see Fig. 8 in Côté 1999). This is highlighted in Harris (2003) for M31 and by Beasley et al. (2003) for NGC 5128. In the simplest version of our model, we would predict that the metallicity distributions of the field stars and the clusters would be identical.

Harris (2003) points out that another way of thinking about this is that the specific frequency (number of clusters per galaxy luminosity) needs to be roughly 5 times larger for the blue (metal-poor) clusters than the red (metal-rich) clusters. He argues...
that the efficiency of formation of the blue clusters must somehow be higher. However, another possibility is that the difference lies in disruption rather than formation mechanisms. In the context of the current paper, this might simply be explained by changing the infant mortality rate from 90% for the current epoch to 80% during the epoch when the old (blue) globular cluster population formed. This would result in ~4 times more blue clusters surviving infant mortality in the first 100 Myr. Hence, a very small change in the survival rate for the blue, metal-poor clusters would be enough to change the value of $S_N$ by a factor of ~5 without significantly changing the overall framework of the model. We refer the reader to Harris et al. (2006) for a discussion of what might cause the blue, metal-poor globular clusters to have a lower disruption rate. More generally, any variable (e.g., orbital motions) that correlates with metallicity and also couples to disruption rates would achieve a similar result (see Fall & Rees 1985 for a discussion of this point). This topic will be addressed more carefully in Paper II.

4.2. Choice of the Initial Cluster Mass Function

In this paper we have assumed a single power law for the initial mass function, since many works point to this form (see references in § 1). However, there have recently been suggestions that the initial mass function may be a broken power law (e.g., Mengel et al. 2005; Gieles et al. 2006) or even a Gaussian (e.g., Fritz-v. Alvensleben 2005; de Grijs et al. 2005). A full treatment of these forms is beyond the scope of this work, particularly since in our judgment, there are a number of reasons to remain suspicious of these suggestions.

Mengel et al. (2005) present a cluster mass function for the Antennae that is best fit by two power laws or a broken power law (i.e., the mass function is steeper at the high-mass end and flattens toward lower masses). They use a K-band limited cluster sample, which is much shallower than the HST data referred to here. As they make clear, their sample has very complicated and poorly understood selection effects, making it difficult to assess incompleteness and likely affecting the low-mass end of their mass function. Given these difficulties, Mengel et al. (2005) caution against “overinterpreting” the presentation of their mass function. A comparison of the Antennae mass function discussed in § 3.3 and that presented in Mengel et al. (2005) shows good agreement in the slope at the high-mass end (see S. M. Fall et al. 2007, in preparation). Their mass function begins to deviate (becomes shallower) from ours at lower masses, exactly as expected from a shallower data set without robust completeness corrections.

De Grijs et al. (2005) claim that the mass function for the 1 Gyr cluster population in M82 was initially a Gaussian. They find that the mean mass they derive for the 1 Gyr population is $10^5 M_\odot$, essentially the same as a normal population of old globular clusters. However, we note that this conclusion is not consistent with data from de Grijs et al. (2003b), since all nine of the clusters with ages less than 100 Myr in Figure 3 of that paper have masses less than $10^5 M_\odot$. The probability of having nine data points below the mean of a Gaussian distribution is ~0.002. Given the high and variable extinction in this galaxy, we suspect that completeness and extinction effects may be underestimated for their 1 Gyr population, resulting in an overestimate of the mean mass of the population.

We conclude by noting that in the Antennae, at least, clusters with ages less than 100 Myr clearly have a power-law mass distribution down to the limit at which clusters can be separated from stars based on luminosity (Zhang & Fall 1999; Fall 2004; S. M. Fall et al. 2007, in preparation), hence justifying our assumption of an initial power-law mass function in our model.

4.3. The Antennae as a Scaled-Up Version of the Milky Way

Lada & Lada (2003) make several of the same points about the need for rapid cluster disruption for very young clusters in the Milky Way as are made for the Antennae in this paper (§ 3.2). In fact, it is possible to view the cluster demographics in the Antennae as a scaled-up version of the Milky Way. According to Lada & Lada (2003), the Milky Way has roughly 100 known embedded young clusters within a distance of ~1 kpc. The most massive clusters in their sample are slightly more massive than $10^3 M_\odot$. If we were able to see the entire disk of the Milky Way the sample would be roughly 100 times larger. The current star formation rate per unit area in the Antennae is also roughly a factor of 10 higher than that of the Milky Way (Zhang et al. 2001); hence, the total enhancement for the number of young clusters in the sample might be roughly a factor of 1000. Assuming a universal power law with an index of ~$-2$ for the initial mass function, the most massive clusters with ages comparable to those of the Lada & Lada sample (i.e., ~10 Myr) would be predicted to be slightly more massive than $10^6 M_\odot$, just as they are in the Antennae.

In Figure 15 we plot cluster age distributions for our Antennae clusters more massive than $2 \times 10^7 M_\odot$, the Lada & Lada (2003) embedded cluster sample for the local Milky Way, and clusters in the SMC using data from Rafelski & Zaritsky (2005) but reanalyzed by Chandar et al. (2006a). We note that a $dN/d\tau \propto \tau^{-1}$ dependence fits the data in all three galaxies. Hence, we find that the cluster age distributions in these three very different environments (and covering different mass ranges) all exhibit the same characteristic behavior, with an ~$\tau^{-1}$ decline in the number of clusters with age over at least the first ~$10^8$ yr.
5. SUMMARY

Motivated by our previous results using age and mass distributions for the Antennae cluster system, and similar results for a few other galaxies, we have developed a framework for understanding the demographics of star clusters, along with a toy model. This model incorporates a universal initial power-law mass function, selected formation histories, selected disruption mechanisms, and a convolution with artifacts and selection effects. The three key parameters in the models are the power-law index for the initial mass function ($\beta \approx -2$; primarily affects the cluster mass distribution at the high-mass end and total galaxy mass and luminosity), the percentage of clusters disrupted in the infant mortality disruption law ($\approx 90\%$ per decade of $\tau$, i.e., $\tau^{-1}$ dependencies; primarily affects the total galaxy mass), and the mass-loss rate $\mu_\text{ev}$ (primarily affects the shape of the observed cluster mass function at very low masses and the total mass).

A wide variety of observations can be explained by this simple model. In this particular contribution we concentrate on the $M_*/(\text{brightest})$ versus log $N$ relationship and on extending the range of comparison for observations of the Antennae galaxies. In Paper II we will consider several other nearby galaxies.

1. The correlation between the brightest young cluster in a galaxy and the total number of young clusters [i.e., $M_*/(\text{brightest})$ versus log $N$] can be understood as a statistical size-of-sample effect rather than a difference in the physical process responsible for the formation of the clusters. One possibility for the much larger number of clusters in mergers is that conditions for making YMCs are globally present, while in spiral galaxies they are only locally present (i.e., in the spiral arms). The diagram is quite sensitive to the value of $\alpha$ for the initial power-law luminosity function, with relatively good agreement in the range $-2.0 < \alpha < -2.4$.

2. A detailed comparison is made between different cluster disruption laws, and it is demonstrated that the apparent break in the $dn/d\tau$ versus log $\tau$ diagram used to determine the parameters in the BL03 model may be produced by incompleteness at the break point.

3. A four-component model of the Antennae galaxies was first developed using only preexisting information (i.e., 90% infant mortality rate, $\beta = -2.0$, and cluster formation histories from Whitmore et al. 1999). The model showed reasonably good agreement with the data, although there were three minor areas of disagreement. In addition, the total magnitude predicted for the galaxy by extrapolating $\beta$ to the low stellar mass regime was $\approx 1$ mag brighter than the Antennae.

These shortcomings were addressed in a five-component model of the Antennae galaxies, which agrees with the observations quite well, based on matches to the magnitude-age, number-magnitude (i.e., luminosity function), color-magnitude, color-age, color-color, mass-age, and number-age (i.e., age distribution) diagrams. The model employs a two-stage cluster disruption law with 80% infant mortality in the first 100 Myr (i.e., $a \gamma \approx -0.7$ dependence, similar to the observed value of $\gamma \approx -1$; Fall et al. 2005) and a value of $\beta = -2.1$ for the index of the initial mass function (within the uncertainties of previous studies; Zhang & Fall 1999; Fall 2004; S. M. Fall et al. 2007, in preparation). This results in a predicted total galaxy mass and luminosity, and fraction of UV light in clusters, that are in reasonably good agreement with the observations (see Table 2).

4. We find no evidence for a truncation of the cluster mass function at the high-mass end. In fact, $\beta \approx -2.1$ gives a good match to the observed cluster mass function above the stellar contamination limit (see S. M. Fall et al. 2007, in preparation, for a more detailed treatment).

5. Four potential anomalies are examined and found not to be serious problems for the framework. The brightest cluster in NGC 1569 falls well within the statistical uncertainty expected by the model. Intermediate-age globular clusters have recently been found in nearby spiral galaxies such as M33 (e.g., Chandar et al. 2002, 2006b) and M31 (e.g., Barmby et al. 2000; Li & Burstein 2003; Puzia et al. 2005). The difference in the [Fe/H] distribution functions of stars and globular clusters in various galaxies may be due to less effective disruption of the metal-poor clusters (e.g., Fall & Rees 1985; Harris et al. 2006). Claims that the initial cluster mass function in various galaxies deviates significantly from a single power law may be due to completeness problems. The mass function for Antennae clusters younger than 10 Myr is well represented by a single power law down to $\approx 10^4 M_\odot$.

6. It is demonstrated that the basic demographics of the star clusters in the Antennae galaxies are consistent with a scaled-up version of the local neighborhood of the Milky Way. The required scaling factor ($\approx 1000$) is due to the larger volume and higher star formation rate in the Antennae. We find that the Antennae, the Milky Way, and the SMC all follow the same $dn/d\tau \propto \tau^{-1}$ relation over the first $\approx 10^8$ yr. These results provide further support for the universal nature of the framework outlined in this paper.

We thank Francois Schweizer and the referee for comments that improved the paper. This work was partially supported by NASA grants GO-09720-A and GO-10188-A. HST data were obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

REFERENCES

Applegate, J. H. 1986, ApJ, 301, 132
Ashman, K., & Zepf, S. 1998, Globular Cluster Systems (Cambridge: Cambridge Univ. Press)
Barmby, P., Huchra, J. P., Brodie, J. P., Forbes, D. A., Schroder, L. L., & Grillmair, C. J. 2000, AJ, 119, 727
Barnes, J. 1988, ApJ, 331, 699
Barth, A. J., Ho, L. C., Filippenko, A. V., & Sargent, W. L. 1995, AJ, 110, 1009
Bastian, N., Gieles, M., Lamers, H. J. G. L. M., Scheepmaker, R. A., & de Grijs, R. 2005, A&A, 431, 905
Baumgardt, H., & Makino, J. 2003, MNRAS, 340, 227
Beasley, M. A., Brodie, J. P., Strader, J., Forbes, D. A., Proctor, R. N., Barmby, P., & Huchra, J. P. 2004, AJ, 128, 1623
Beasley, M. A., Harris, W. E., Harris, G. L., & Forbes, D. A. 2003, MNRAS, 340, 341
Bica, E., Santiago, B. X., Dutra, C. M., Dotti, H., de Oliveira, M. R., & Pavani, D. 2001, A&A, 366, 827
Billett, O. H., Hunter, D. A., & Elmegreen, B. G. 2002, AJ, 123, 1454
Boily, C. M., & Kroupa, P. 2003a, MNRAS, 338, 665
———. 2003b, MNRAS, 338, 673
Boultoukos, S. G., & Lammers, H. J. G. L. M. 2003, MNRAS, 338, 717 (BL03)
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burstein, D., et al. 2004, ApJ, 614, 159
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Cervino, M., Valls-Gabaud, D., Luridiana, V., & Mas-Hesse, J. M. 2002, A&A, 381, 51
Chandar, R., Bianchi, L., & Ford, H. C. 1999a, ApJS, 122, 431
———. 1999b, ApJ, 517, 668
Chandar, R., Bianchi, L., Ford, H. C., & Sarajedini, A. 2002, ApJ, 564, 712
Chandar, L., & Whitmore, B. C. 2003, ApJS, 148, 115
Chandar, R., Bianchi, L., & Ford, H. C. 1999a, ApJS, 122, 431
———. 1999b, ApJ, 517, 668
Chandar, R., Bianchi, L., Ford, H. C., & Sarajedini, A. 2002, ApJ, 564, 712
Cervino, M., Valls-Gabaud, D., Luridiana, V., & Mas-Hesse, J. M. 2002, A&A, 381, 51
Davidge, T. 1981, ApJ, 242, 310
Davidge, T. 1984, ApJ, 285, 88
Davidge, T. 1985, ApJ, 293, 409
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burstein, D., et al. 2004, ApJ, 614, 159
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Cervino, M., Valls-Gabaud, D., Luridiana, V., & Mas-Hesse, J. M. 2002, A&A, 381, 51
Chandar, R., Bianchi, L., & Ford, H. C. 1999a, ApJS, 122, 431
———. 1999b, ApJ, 517, 668
Chandar, R., Bianchi, L., Ford, H. C., & Sarajedini, A. 2002, ApJ, 564, 712
