Star Cluster Formation Efficiency in the Andromeda Galaxy

Lent Clifton Johnson IV

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2015

Reading Committee:

Julianne J. Dalcanton, Chair

Anil C. Seth

Thomas R. Quinn

Program Authorized to Offer Degree:
Astronomy

University of Washington

Abstract

Star Cluster Formation Efficiency in the Andromeda Galaxy

Lent Clifton Johnson IV

Chair of the Supervisory Committee:
Professor Julianne J. Dalcanton
Astronomy

This work revolutionizes the study of star clusters in the Local Group galaxy Andromeda (Messier 31) using high spatial resolution, multi-wavelength imaging from the Hubble Space Telescope obtained as part of the Panchromatic Hubble Andromeda Treasury (PHAT) survey. I construct a cluster catalog using visual identification methods, including the use of image classifications collected from citizen scientist volunteers as part of the Andromeda Project. I perform analysis that combines large numbers of volunteer classifications with expert-derived identifications to yield a well-characterized census of star clusters. The resulting cluster catalog surpasses existing Galactic and extragalactic samples in terms of completeness and uniformity, serves as the basis for a wide range of current and future investigations of star formation and stellar evolution in M31, and is an important legacy data product of the PHAT survey.

I use the PHAT star cluster catalog and associated cluster age and mass characterizations to study two aspects of star cluster formation. First, I investigate cluster formation efficiency and its dependence on star formation intensity. This study combines detailed measurements of cluster ages and masses with star formation histories of underlying total stellar populations, where both sets of constraints are derived from fitting color-magnitude diagrams of individually resolved stars. I find that $\sim$4% of young stars (10-100 Myr old) in M31 are born in long-lived star clusters, and demonstrate that this fraction varies systematically as a function of star formation rate surface density ($\Sigma_{\rm SFR}$) and gas depletion

time ($\tau_{\rm dep}$). The results derived here agree with trends established by previous observations of nearby galaxies in which cluster formation efficiency correlates with star formation rate intensity. The spatially-resolved measurements of cluster formation efficiency in M31 are also consistent with theoretical predictions, providing observational support for a model of cluster formation in which star clusters are born in regions of high gas density and star formation efficiency within a hierarchically-structured interstellar medium.

In addition to star cluster formation efficiency, I also study the mass distribution of the young cluster populations (10-300 Myr old) in M31. I find that the mass function shape is well described by a Schechter function, with a power law index of $\alpha = -1.99 \pm 0.12$ and a characteristic mass of $M_c = 8.5^{+2.8}_{-1.8}\ M_{\odot}$. This exponential high-mass truncation of the cluster mass function occurs at a significantly lower mass in M31 than found for other nearby star forming galaxies. I show for the first time that the exponential truncation of the cluster mass function varies systematically with star formation rate intensity, such that the characteristic Schechter mass increases with star formation rate surface density as $M_c \propto {\Sigma_{\rm SFR}}^{\sim 1.3}$. Additionally, I explore the possibility that the M_c–$\Sigma_{\rm SFR}$ relation derived here might also apply to old globular cluster systems, and thus be useful in constraining properties of star formation environments in the early universe.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

This research would not have been possible without the efforts of many supportive mentors, collaborators, and friends. To my co-advisors, Julianne Dalcanton and Anil Seth, for their sage advice and guidance. To PHAT's "Team Cluster", including Morgan Fouesneau, Dan Weisz, Lori Beerman, and Nelson Caldwell. To my fellow UW astronomy grad students, especially Jim Davenport and Alexia Lewis. To my dissertation committee member Tom Quinn. And to Paul Hodge for our valuable conversations.

I am fortunate to have worked as part of two fantastic collaborations: members of the Zooniverse team and the PHAT collaboration were essential to the development and execution of my dissertation work. I wish to individually recognize Chris Lintott, Rob Simpson, Amit Kapadia, and Stuart Lynn from the Zooniverse, as well as Ben Williams and Phil Rosenfield from PHAT.

I owe a great debt of gratitude to the ∼30,000 Andromeda Project volunteers who made this work possible through their contributions as citizen scientists, but also for providing a constant reminder of why I dedicate myself to the work I do. Their names are listed at `http:// www.andromedaproject.org/#!/authors`.

I thank Pauline Barmby, Robert Braun, Karl Gordon, Adam Leroy, and Andreas Schruba for contributing data that was used during the course of my research. I also thank Robert Gendler for his participation in creating the beautiful PHAT survey mosaic, and for providing beautiful visuals of astronomical objects that I use in my work.

Support for this work was provided by NASA through grant number HST-GO-12055 from the Space Telescope Science Institute, which is operated by AURA, Inc., under NASA contract NAS5-26555.

Finally, I thank the team at Milstead & Co. in the Fremont neighborhood of Seattle for the delicious coffee and kind words that powered the writing of this document.

DEDICATION

To my wife, Nell, for her unending support and encouragement to live a fulfilling life.

Chapter 1

INTRODUCTION

Star clusters are groupings of stars that are born coevally out of a common molecular gas distribution. Long-lived star clusters that remain gravitationally bound for hundreds of millions of years or more are valuable astrophysical laboratories, improving our understanding in many fields of astronomical research, including star formation. In this dissertation, I describe my research studying cluster formation and its link to star formation behavior in the Andromeda galaxy (M31). Here I introduce the scientific questions that drive this observational investigation, and discuss how my work fits in the larger context of M31 studies and extragalactic cluster studies.

1.1 Clusters as Tracers of Star Formation

The strong link between star formation and star clusters is seen from the fact that a majority of star formation occurs in spatially-correlated stellar groupings. The seminal review from Lada & Lada (2003) presents the often-quoted result that 70–90% of stars form in embedded clusters, based on infrared observations of star forming regions in the Solar neighborhood. However, this same work goes on to state that only a small fraction (less than 4–7%) of these dust-embedded groupings remain bound and emerge as long-lived, exposed star clusters. While star formation occurs in spatially-correlated distributions, following the distribution of the natal, hierarchically-structured interstellar medium (ISM), the formation of long-lived clusters (e.g., the Pleiades, h and χ Persei) is a relatively rare event.

The relative rarity of long-lived star clusters comes as no surprise from a theoretical perspective and relates closely to the concept of "infant mortality" that was introduced by Lada & Lada (2003). This term describes the effect that gas expulsion has on concentrations of young embedded stars. In the simplified case of an isolated, spherical molecular cloud, theoretical N-body simulations have shown that the rapid removal of gas as a result of stellar

feedback from young stars (driven by supernova, radiative pressure, winds, and/or jets) leads to the dissolution of natal structure that was originally virialized and gravitationally bound (see e.g., Hills 1980; Goodwin & Bastian 2006; Baumgardt & Kroupa 2007). The rapid loss of gas mass from the gravitational potential, which represents a substantial fraction of the total mass in the region (under the assumption of low overall star formation efficiency), leads to an unbinding of the stars. These studies report that local star formation efficiencies of 30–50% are required for a young star forming region to remain bound and yield a long-live star cluster, under the assumption of isolated spherical initial conditions. We note that real star forming regions seldom resemble simple spherical distributions. In reality, star forming regions are shaped by continued gas and stellar infall from neighboring regions, forming out of a hierarchical, filamentary ISM.

While these idealized simulations lack the full complexity observed in star forming regions, they highlight an important principle of cluster formation: star clusters tend to form in regions with high local star formation efficiency (SFE). This result follows from the same logic used to explain infant mortality, where regions that convert larger fractions of the available gas supply will therefore loose less mass as a result of gas expulsion, making it easier for stellar distributions to remain gravitational bound. This reasoning represents the theoretical underpinning for the cluster formation framework presented by Kruijssen (2012). Briefly, Kruijssen combines prescriptions for gas density distributions, star formation efficiency, and stellar feedback to predict the cluster formation rate as a function of local average gas density. In this model, regions with high gas density also have shorter free fall times, leading to an increased total SFE over the duration of star formation (set by the timescale for supernova feedback) while still assuming a fixed SFE per free-fall time (following Elmegreen 2002).

These theoretical descriptions of star cluster formation lead to the idea that star clusters represent tracers of high intensity star formation, and observational studies support this hypothesis. A growing body of observational work (e.g., Goddard et al. 2010; Adamo et al. 2011; Cook et al. 2012) shows that formation rates of long-lived star clusters increase with rising star formation intensity, measured according to the clustered fraction of stars (Γ) and star formation rate surface density ($\Sigma_{\rm SFR}$), respectively. While the existence of a trend

is well-established based on galaxy-integrated measurements, the detailed dependence of cluster formation efficiency on galactic environment would benefit from additional, spatially-resolved studies that can examine trends on a region-by-region basis.

In addition to the increase in cluster formation efficiency with rising star formation intensity, there is also observational evidence for differences in the cluster mass spectrum as a function of star formation intensity. Studies by Larsen (2009) and Gieles et al. (2006b); Gieles (2009) show that the power law cluster mass function in nearby star forming galaxies has a high mass truncation that is well characterized by a Schechter function and its exponential truncation using a characteristic mass (M_c) of $2{\times}10^5$ $M_\odot$. While these studies derive a consistent characteristic mass value across a sample of normal spiral galaxies, analysis of the cluster mass function for the starbursting Antennae galaxies (Zhang & Fall 1999; Jordán et al. 2007) shows a larger M_c value of $\sim 10^6$ $M_\odot$. This result provides an initial indication that the shape of the cluster mass function may depend on star formation intensity, and suggests that extending the range of star forming environments with well characterized cluster mass functions might lead to new insights about cluster formation.

In this dissertation, I explore the link between cluster formation and star formation intensity, both in terms of the amount of mass formed in clusters in Chapter 4, as well as the mass function of emerging cluster populations in Chapter 5. My study of the Andromeda galaxy cluster population provides a new, detailed view of a star cluster population, allowing us to study cluster formation in great detail, and in a new environmental regime at low star formation intensity. Ultimately, we aim to improve our understanding of star cluster formation in order to infer properties of associated, coeval star formation using cluster populations that are relatively easy to observe (as compared to individual field stars), and potentially last for billions of years (in the case of old globular cluster populations).

1.2 The Andromeda Galaxy and its Clusters

The Andromeda galaxy is an excellent target for a detailed study of a star cluster population. The availability of rich, high quality data and the ability to expand cluster studies into a new regime of low intensity star formation make M31 a valuable observational objective. This Local Group neighbor of the Milky Way benefits from the distinct advantage of proximity,

lying at a distance of 785 kpc (McConnachie et al. 2005). Andromeda is the only massive L^* spiral galaxy for which we can resolve individual star cluster members, allowing us to construct color magnitude diagrams (CMDs). Fitting CMDs allows precise and accurate cluster age and mass determinations, and is consistent with methods used to analyze nearby star clusters in the Milky Way. In contrast, observations of clusters in more distant galaxies (>1 Mpc) only allow analysis of their integrated light, typically in the form of broadband spectral energy distributions (SEDs), and low resolution spectroscopy for the brightest clusters. Furthermore, M31 provides a complementary $\sim Z_{\odot}$ foil to the less massive, lower metallicity galaxies M33, the Large Magellanic Cloud (LMC) and Small Magellanic Cloud (SMC).

Andromeda serves an important role with respect to understanding how cluster formation varies with star formation rate (SFR) intensity, providing sampling at low $\Sigma_{\rm SFR}$, significantly extending the range of well-measured environments studied in the literature. Unlike its Local Group companions, Andromeda has a specific SFR and $\Sigma_{\rm SFR}$ that are less than half the values found for the Milky Way, M33, or the Magellanic Clouds. Along with this lower SFR intensity, the accompanying gas surface density in M31 is sufficiently low such that the atomic hydrogen component of the ISM dominates over the molecular hydrogen component. These properties lead to qualitatively different star formation behavior (see Bigiel et al. 2008; Leroy et al. 2008), characterized by lower SFE and longer gas depletion times ($\tau_{\rm dep}$). The ability to perform a detailed study of star cluster formation in a new regime of star formation intensity will greatly improve our understanding of the environmental dependence of these processes.

Given Andromeda's close proximity and visibility to observatories in the northern hemisphere, it is no surprise that my research builds upon decades of previous observational work on star clusters in M31. Edwin Hubble published the first work on M31's globular cluster population (Hubble 1932), only ∼8 years after first recognizing that the Andromeda "nebula" was an external galaxy. Most previous research has focused on these ancient globular clusters, as they are the brightest and therefore most easily observed of M31's star clusters. These studies include the construction of the Revised Bologna Catalog (Galleti et al. 2004), extensive spectroscopic follow up by Caldwell et al. (2009, 2011), and intriguing work that

studies the globular cluster population throughout the M31 halo (e.g., Huxor et al. 2005, 2008, 2014) and their connection to halo substructures (e.g., Mackey et al. 2010, 2014).

Relative to these studies of globular clusters, the young cluster population in M31 has received relatively less attention. From ground-based observations (e.g., Hodge 1979; Hodge et al. 1987; Fusi Pecci et al. 2005), it became clear that Andromeda hosts a robust young cluster population, but one which is difficult to identify using low-resolution ground-based images. Beginning with HST-enabled studies by (Williams & Hodge 2001a,b), the benefits of high spatial resolution imaging to M31 cluster studies became clear. These initial studies were followed by a systematic survey of the brightest young star clusters using individual targeted Hubble Space Telescope (HST) observations (Perina et al. 2009b, 2010). Alternately, a series of cluster searches published in Krienke & Hodge (2007, 2008, 2013) and Hodge et al. (2009, 2010) made use of assorted HST observations scattered across the disk of M31 to construct a substantial initial sample of clusters.

Although these early studies made great strides in identifying and analyzing Andromeda's young cluster population, the low efficiency of individual pointed observations and the heterogeneous nature of archival observations led to the conclusion that a wide-area HST imaging survey was the necessary next step towards answering substantive scientific questions about the M31 star cluster population.

1.3 The Panchromatic Hubble Andromeda Treasury Survey

The Panchromatic Hubble Andromeda Treasury survey (PHAT; Dalcanton et al. 2012) provides a transformative imaging dataset, allowing substantial forward progress in the study of M31's star clusters. The PHAT survey is a multi-cycle observing program executed using HST to observe one-third of M31's stellar disk. The survey utilized 828 orbits of observing time over the course of $\sim$3 years (2010–2013) to map a contiguous swath of Andromeda's stellar population. PHAT made use of the Advanced Camera for Surveys (ACS) and Wide Field Camera 3 (WFC3) instruments observing in parallel as part of an efficient observing protocol to observe 414 individual pointings. Imaging was obtained in six broadband filters spanning the electromagnetic spectrum from the near-ultraviolet (near-UV) to the near-infrared (near-IR): F275W (near-UV), F336W (U), F475W (g), F814W

Figure 1.1 PHAT survey footprint in the context of the full Andromeda galaxy. Image Credit: NASA, ESA, Z. Levay (STScI/AURA), and Robert Gendler

Figure 1.2 PHAT survey-wide mosaic created from optical (F475W+F814W) images. Image Credit: NASA, ESA, J.J. Dalcanton, B.F. Williams, and L.C. Johnson (University of Washington), the PHAT collaboration, and Robert Gendler

(I), F110W (J), F160W (H).

The PHAT survey footprint is located in the northeast portion of the M31 disk, and is shown in context with the full galaxy in Figure 1.1. As detailed in Dalcanton et al. (2012), the survey aimed to observe a "generous quadrant" of Andromeda in an effort to obtain significant sampling of star formation over a wide range of galactocentric radii. In addition, preference was given to the eastern side of the disk due to larger dust attenuation found on the western (near) side, and the northeast quadrant was preferred due to contamination from stars in the dwarf elliptical M32 that lies projected on the M31 disk in the southeast quadrant. A portion of the final data is shown in Figure 1.2. The high spatial resolution obtained using HST, combined with the contiguous wide-area coverage of the disk ($\sim$0.5 deg^2; $\sim$400 kpc^2), provide an incredible dataset for the study of resolved stellar populations.

The primary scientific product from the survey is the catalog of stellar photometry. As presented in Williams et al. (2014), the survey measured six-band photometry for 117 million individually resolved stars using crowded field, point-spread function (PSF) fitting techniques. This database of equidistant stars reaches a completeness limit of $\sim$28 mag in F475W for the outer disk of M31. The depth and spatial resolution (through its influence on crowding limits for the photometry) allow for the detection of individual main sequence stars down to $\sim$3 $M_\odot$. As discussed in the previous section, the ability to analyze individual cluster members permits a detailed level of analysis that closely resemble star cluster studies in the Milky Way and Magellanic Clouds.

Finally, the PHAT survey also benefits from multiple ancillary datasets that have observed M31 across a broad array of wavelengths. Specific to my dissertation research, gas phase observations are incredibly valuable for assessing the properties of the ISM out of which star clusters form. Measurements of atomic and molecular hydrogen surface densities, obtained from HI and CO observations (Nieten et al. 2006; Braun et al. 2009, A. Schruba, in preparation) were particularly valuable. In addition, Spitzer Space Telescope infrared observations (Barmby et al. 2006; Gordon et al. 2006), GALEX ultraviolet observations (Thilker et al. 2005), and ground-based spectroscopy of star clusters (Caldwell et al. 2009, 2011) and individual stars (e.g., Massey et al. 2009; Dorman et al. 2015) all contribute in the effort to understanding the galactic environment and stellar populations contained

within the M31 disk.

1.4 Outline

Using data from the PHAT survey, I perform a census of star clusters in the Andromeda galaxy and use this high-quality sample to address important scientific questions concerning cluster formation. In Chapters 2 and 3, I describe my efforts to construct a well-characterized cluster catalog using visual identification. I present an initial expert-led analysis conduct in a traditional manner, then present the final PHAT cluster catalog created using image classifications collected from the Andromeda Project citizen science website. Then, I explore the cluster formation efficiency (Γ) and its dependence on local star formation rate intensity in Chapter 4, and present analysis of the cluster mass function in Chapter 5. I close with a summary of my work and a discussion of future directions in Chapter 6.

Throughout this document, the pronoun "we" is used to refer to myself and my co-investigators who assisted in this research. While I was the lead investigator for all of the work presented here, I use the plural pronoun to properly credit the critical role played by my coauthors and PHAT collaboration team members whose contributions were essential to the successful completion of this work.

Chapter 2

YEAR 1 CLUSTER CATALOG

This chapter was published as L.C. Johnson et al. 2012, ApJ, 752, 95, and is reproduced by permission of the AAS.

We present the first results from the PHAT stellar cluster survey, the Year 1 cluster catalog. This work presents results derived from the first ∼25% of the PHAT survey data, where we identify 601 clusters that make up the Year 1 sample. This number represents more than a factor of four increase over previous catalogs within the Year 1 survey region (390 arcmin^2), and demonstrates our ability to construct a high-quality cluster catalog from Hubble Space Telescope imaging of M31 using visual cluster identification techniques. We present a catalog of clusters with positions, radii, and six-band integrated photometry. Along with a general characterization of the cluster luminosities and colors, we discuss the cluster luminosity function, the cluster size distributions, and highlight a number of individually interesting clusters found in the Year 1 search.

2.1 Introduction

Large, high-quality samples of stellar clusters provide key data for studies of a wide variety of astrophysical topics, including cluster evolution, stellar evolution, star formation, and galaxy evolution. However, current cluster samples suffer from serious limitations. For example, Milky Way clusters suffer from severe dust attenuation within the Galactic plane, resulting in disk cluster samples that are complete only within a small region (radius of ∼1 kpc) around the Sun (e.g., Dias et al. 2002; Piskunov et al. 2008). This limits the variety of objects and galactic environments explored, as exemplified by the dearth of massive, intermediate-age clusters known within the Milky Way (e.g., Davies et al. 2011). Although infrared surveys of the Galactic plane are improving this situation (e.g., Dutra et al. 2003; Mercer et al. 2005; Borissova et al. 2011), current samples of Galactic clusters do not probe the full

stellar cluster parameter space, limiting our ability to study mass, age, and environmental dependencies of evolutionary processes.

One solution to the incompleteness of Galactic samples has been to extend cluster studies to other galaxies. Extragalactic cluster samples have grown immensely over the past decade, thanks in part to the power of Hubble Space Telescope (HST) imaging. These analyses probe a variety of galactic environments, ranging from starbursting galaxy mergers (e.g., Whitmore et al. 1999) to quiescent spirals (e.g., Larsen 2002), producing provocative results concerning the environmental dependence of cluster formation and evolution (e.g., Larsen & Richtler 2000; Goddard et al. 2010). However, even with HST, clusters in most distant galaxies appear as marginally resolved single objects. This increases the difficulty and uncertainty associated with basic steps in cluster analysis, such as object identification, photometry, and the derivation of ages and masses. As a result, the interpretation of underlying cluster evolutionary processes has considerable associated uncertainty, even leading to cases of conflicting interpretations derived from the same dataset (e.g., in M83; Chandar et al. 2010c; Bastian et al. 2012b; Fouesneau et al. 2012).

Closer to the Milky Way, large cluster catalogs exist for the Large and Small Magellanic Clouds (LMC & SMC; Bica et al. 2008, 1999; Bica & Dutra 2000; Hunter et al. 2003). In the Clouds, the ability to resolve clusters into individual stars has resulted in a number of important results in stellar evolution (e.g., Chiosi et al. 1989; Frogel et al. 1990; Girardi et al. 2009), cluster evolution (e.g., Gieles & Bastian 2008; Chandar et al. 2010a), and cluster formation (e.g., Mackey et al. 2008). However, there are limitations associated even with this excellent sample of objects. On a practical level, while these cluster samples do not suffer from the same incompleteness issues as those from the Milky Way, their piecemeal assembly by many different groups has imprinted a complex and little-understood selection function. On a more fundamental level, there are additional limits to the applicability of stellar and cluster evolution results derived from these interacting, relatively low-mass galaxies. The LMC and SMC do not probe the range in galactic environments that are characteristic of a majority of the baryonic universe. More than 75% of all stars in the Universe today have metallicities within a factor of two of the solar value (Gallazzi et al. 2008), higher than those probed by the Clouds. No cluster sample comparable to those in

the Clouds is currently available in a large spiral galaxy.

The neighboring galaxy M31 (the Andromeda galaxy) is a prime target for stellar cluster studies. The galaxy's proximity allows resolution of individual bright stars in clusters and the robust detection of clusters down to faint ($< 10^4\ L_\odot$) luminosities. M31 also provides access to a range of widely-varying environments across the extent of the star-forming disk. Andromeda's role as a valuable laboratory was realized long ago, and decades of work have gone into exploring its cluster population. We defer the detailed review of existing cluster catalogs until Section 2.5, but beginning with Hubble (1932), previous work has mainly focused on M31's globular cluster population through ground-based imaging (e.g., Galleti et al. 2004; Peacock et al. 2010). The proximity of M31 also enables high-resolution spectroscopic follow-up of bright clusters (e.g., Colucci et al. 2009; Strader et al. 2011), as well as low-resolution spectroscopy of intrinsically faint clusters that are inaccessible in distant galaxies (e.g., Caldwell et al. 2009). Numerous studies have utilized HST's excellent spatial resolution to study massive clusters (e.g., Barmby & Huchra 2001; Perina et al. 2009b) and their individual resolved stars (e.g., Rich et al. 2005; Mackey et al. 2006; Perina et al. 2009a, 2011), as well as to identify and study less massive disk clusters (e.g., the Hodge-Krienke Catalogs; Krienke & Hodge 2007, 2008; Hodge et al. 2009, 2010, hereafter the HKC). While extremely valuable, these previous space-based observing programs have been limited to focused studies of a small number of targeted regions, as opposed to a wide-ranging survey to obtain a broad sampling of the galaxy and an overall assessment of the M31 cluster population.

The Panchromatic Hubble Andromeda Treasury (PHAT; Dalcanton et al. 2012) is poised to revolutionize the study of stellar clusters in M31. This on-going HST multi-cycle program will image one-third of the M31 disk at high spatial resolution, with wavelength coverage from the ultraviolet through the near-infrared. In terms of cluster studies, this survey provides a number of distinct advantages over existing work. High spatial resolution imaging allows us to resolve clusters into individual stars, permitting detailed characterization of their stellar populations through analysis of their color-magnitude diagrams (CMDs). Figure 2.1 demonstrates the data quality provided by PHAT for a previously unidentified cluster, showing the considerable gain over existing ground-based surveys. The wide wave-

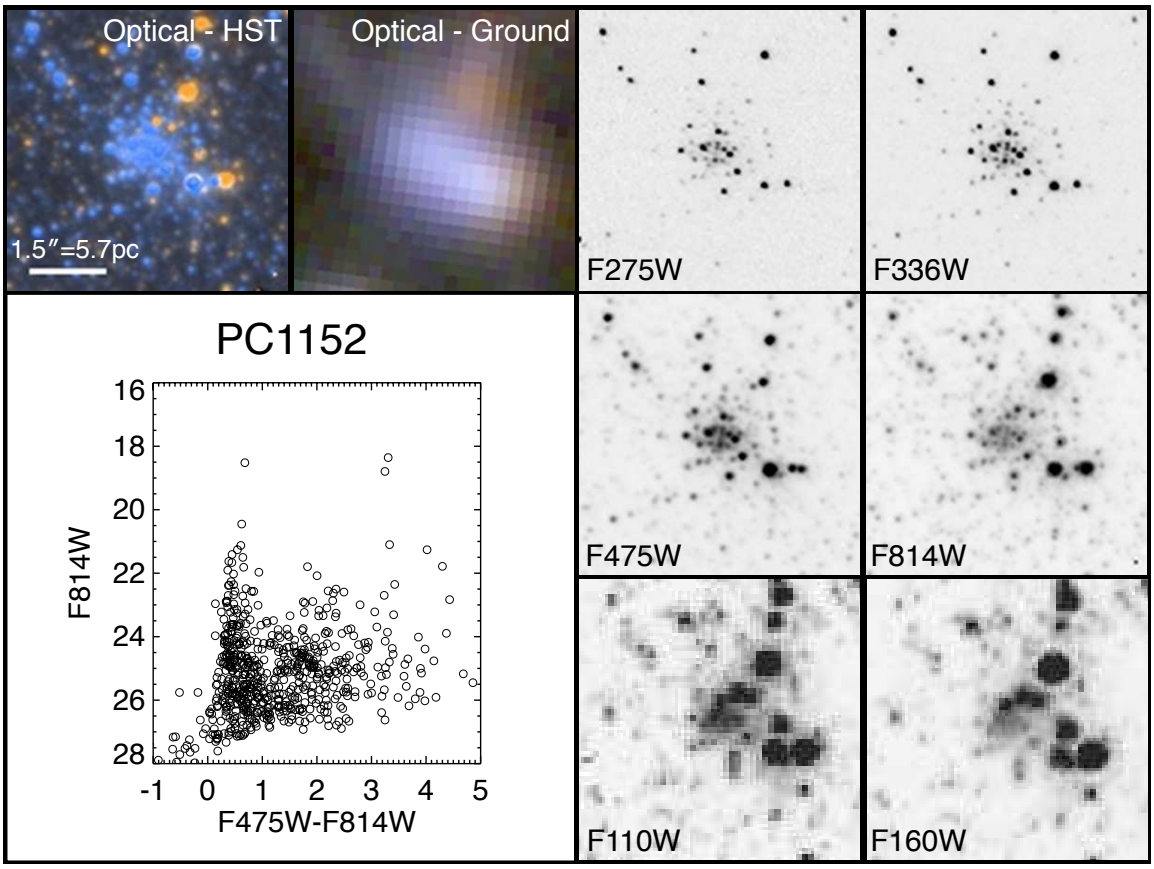

Figure 2.1 PHAT survey data quality example, showing newly-identified cluster PC1152. Six grayscale single-band images, as well as a color optical (F475W+F814W) mosaic, show the superior image quality provided by HST when compared to ground-based observations (from the Local Group Galaxy Survey; Massey et al. 2006). We also present a color-magnitude diagram of resolved photometry for objects that lie within the cluster cutout image. The cluster's main sequence forms the vertical sequence at F475W-F814W$\sim$0.5, along with three evolved giants at F814W$<$20. Other stars shown with F814W$>$20 and F475W-F814W$>$1.0 are likely coincident stars belonging to the background field.

length coverage enables accurate age-dating, as well as the ability to probe a broad range of stellar effective temperatures, from massive main sequence stars to evolved supergiants and AGB stars. As a result, these objects provide a wealth of valuable knowledge in terms of calibrating and refining stellar evolution models at high metallicity. Finally, these images allow the detection of faint clusters, reliably probing more than $\sim$2 magnitudes further down the luminosity function than was previously possible using ground-based datasets. The PHAT survey will increase the HST spatial coverage of the M31 searched for clusters by a factor of $\sim$3 (1800 arcmin^2 in total compared to 650 arcmin^2 surveyed in the HKC; 390 arcmin^2 in the current Year 1 dataset). However, this metric underestimates the scientific gain provided by high-quality, uniform PHAT data products, as opposed to the heterogeneous archival data used in the HKC work. Fundamentally, the PHAT survey represents the shift from a discrete, targeted mode of cluster study to a broad survey mode, allowing for comprehensive analysis of cluster evolutionary processes its environmental dependencies.

The stellar clusters identified as part of the PHAT survey will constitute the most comprehensive sample of clusters available for any large spiral galaxy. The large range of galactocentric radius (0-20 kpc) included in the survey spans a wide range of star formation intensities and gas densities. The diversity of galactic environments will be important for testing models of cluster formation and evolution. In addition, the simultaneous accessibility of objects over a >3 order of magnitude range in cluster luminosity provides a top-to-bottom view of the cluster population, given that we sample a continuous range of objects that extend from those equivalent to Galactic open clusters up to massive globular clusters.

This paper is the first in a series utilizing the PHAT dataset for studies of stellar clusters. Here, we present the first installment of a HST-based cluster catalog, which will serve as the basis for extensive study of Andromeda's cluster population. Catalog updates and improvements will be published over the course of this four year observing program. In this edition, we publish positions, sizes, and integrated photometry for the Year 1 cluster sample. Age and mass determinations derived from the integrated photometry will be presented in Fouesneau et al. (2014). Additional studies, including analysis of structural parameters, resolved star content, and integrated spectroscopy of the cluster sample will follow in subsequent work.

We summarize the PHAT observations in Section 2.2, while in Section 2.3 we describe our cluster identification procedures, present results from completeness testing, and introduce the Year 1 cluster catalog. Next, we describe and test our photometry methodology in Section 2.4, followed by a comparison between the PHAT cluster catalog and existing catalogs in Section 2.5. We present a basic characterization of the cluster catalog contents in Section 2.6, followed by discussions of luminosity functions, the cluster size distribution, and objects of interest in Section 2.7. We conclude with a summary and description of future work in Section 2.8. Throughout this work, we assume a distance modulus for M31 of 24.47 (785 kpc; McConnachie et al. 2005), for which $1''$ corresponds to a physical size of 3.81 pc.

2.2 Observational Data

This paper includes clusters identified in Year 1 PHAT imaging data, taken before June 2011. A full description of the PHAT observational design is available in Dalcanton et al. (2012), but we briefly summarize relevant details below. PHAT observations are grouped into 23 area units known as "bricks", each made up of 18 mosaiced HST fields of view in a rectangular 6×3 arrangement that covers a $\sim 12' \times 6.5'$ region of sky. Data are obtained simultaneously at different field centers with the ACS (using the F475W and F814W filters) and WFC3 (using the F275W, F336W, F110W, and F160W filters) instruments in two epochs, separated by ∼6 months. During each epoch, imaging is obtained by the cameras in two side-by-side, half-brick (3×3) arrays. Between the epochs, the orientation of the cameras change by 180 degrees due to the annual roll angle variation of HST. As a result, the half-brick mosaic obtained by the ACS camera during the first epoch is now imaged by the WFC3 camera in the second epoch, and vice-versa, completing six-filter imaging across the brick. In all, we obtain ∼130 minutes of exposure time at each of the brick's 18 field centers.

The Year 1 imaging used in this work includes four full bricks (designated B01, B09, B15, and B21) and the western halves of two additional bricks (B17W and B23W). These data sample locations along the major axis of M31 from the center out to a projected radius of ∼20 kpc. The full PHAT survey footprint, along with the locations of the Year 1 bricks

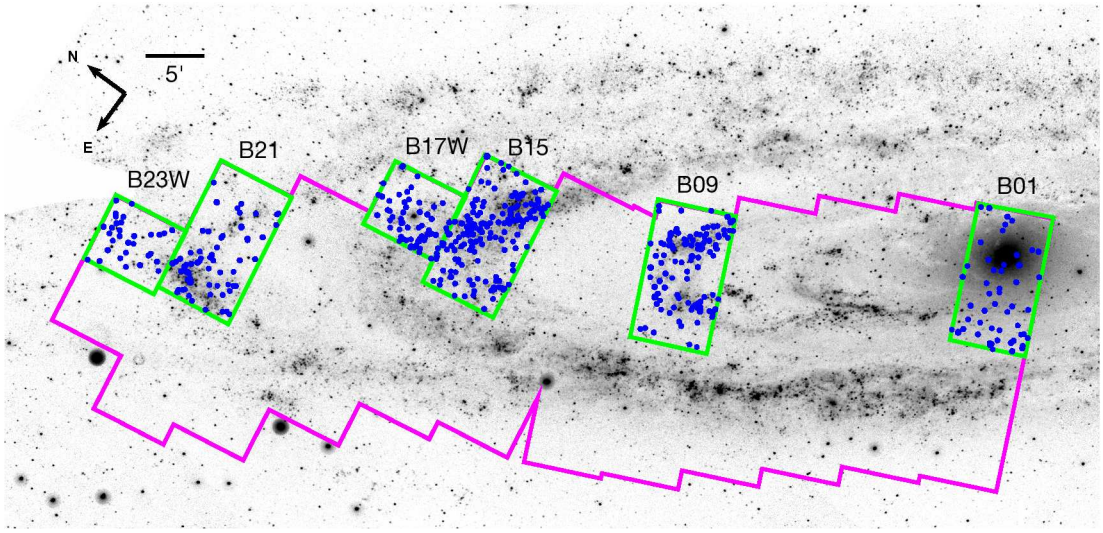

Figure 2.2 Footprint of the PHAT survey region (magenta) displayed on a GALEX NUV image of the northeast half of M31. Green rectangles represent the "bricks" that make up the Year 1 imaging data. Blue circles show the spatial distribution of clusters identified in the Year 1 cluster search.

are presented in Figure 2.2.

Combining imaging data across multiple pointings from three separate cameras (ACS/WFC, WFC3/UVIS, and WFC3/IR) requires astrometry with higher precision than that obtained using the standard telescope telemetry and data processing pipeline. Images of neighboring fields have sufficient overlap to enable us to derive an astrometric solution across a full (or half) brick. These astrometric solutions are obtained separately for each camera using photometric catalogs derived using DOLPHOT[1], a modified version of HSTPhot (Dolphin 2000) that has been updated to include specialized ACS and WFC3 modules. Affine distortion corrections (in addition to those already known for each camera from the IDCTABs) are required to obtain consistent brick-wide astrometric solutions. An additional correction brings the brick-wide astrometric solutions onto a global astrometric frame, defined using CFHT observations tied to 2MASS catalogs (Skrutskie et al. 2006). The global astrometric alignment agrees with that of the 2MASS reference system within an absolute level of ∼60 mas.

[1] `http://purcell.as.arizona.edu/dolphot`

Once aligned, we use the `multidrizzle` task within PyRAF (Koekemoer et al. 2002) for cosmic ray rejection and image creation along with `lacosmic` (van Dokkum 2001) for supplemental cosmic ray flagging. We note that cosmic ray correction is particularly difficult in the case of the WFC3/UVIS data, due to a large number of cosmic ray artifacts (particularly in F275W images) and the availability of only two frames for artifact detection in regions of non-overlapping field coverage. Pixel scales of the resulting images are (0.04, 0.05, and 0.065 arcsec/pixel) for the (WFC3/UVIS, ACS/WFC, WFC3/IR) cameras, where the WFC3/IR images are up-sampled from their native plate scale (0.128 arcsec/pixel) to take advantage of the higher effective resolution afforded by the survey's sub-pixel dither strategy.

2.3 Cluster Identification

The goal of the PHAT cluster survey is to identify and analyze a sample of gravitationally bound stars clusters in M31. In this paper, we undertake the first step toward achieving this goal: the visual inspection of Year 1 PHAT imaging to identify candidate bound clusters. As shown in Fig. 2.1, clusters appear in PHAT imaging as composite objects composed of a centrally concentrated overdensity of individual resolved stars and a broad unresolved light component. The ability to resolve these objects into individual stars allows for the clean separation between genuine stellar clusters and contaminants such as background galaxies or single stars.

Although PHAT imaging facilitates robust identifications of stellar overdensities, determining whether or not these clusterings are gravitationally bound is challenging. Quantitative assessment of an object's boundedness requires age, mass, and spatial profile information (e.g., Gieles & Portegies Zwart 2011). Further, determining the boundedness of young objects ($\lesssim$10 Myr) is made even more difficult, because dynamical evolution has had little time to evolve stellar structures from an initial hierarchical, scale-free spatial distribution (for further discussion, see e.g., Bastian et al. 2011, and references therein), blurring the distinction between bound and unbound stellar groupings. In future work, we will utilize age, mass, and structural characteristics to assess boundedness for each object (e.g., using the Π statistic; Gieles & Portegies Zwart 2011), but that analysis is beyond the scope of

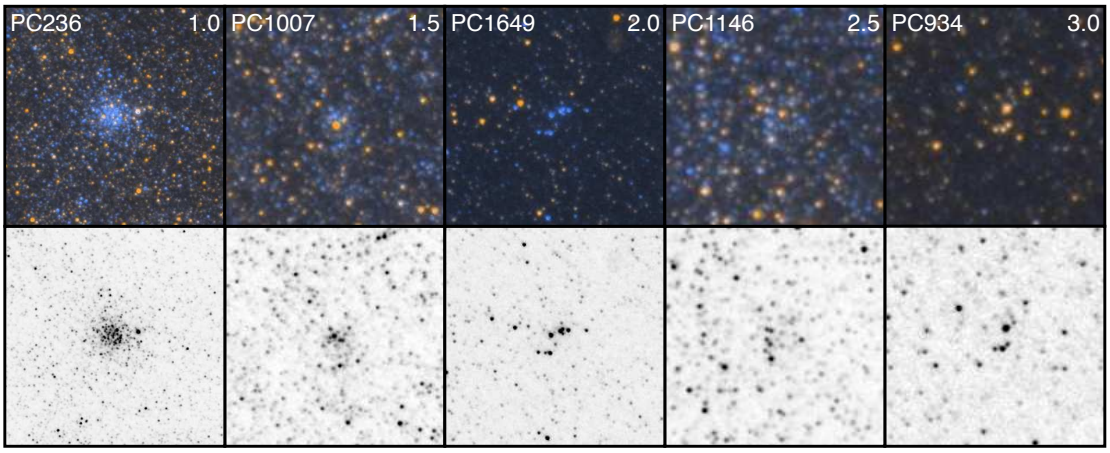

Figure 2.3 Cluster cutout images showing objects across the range of possible S_{by-eye}. The top row shows color optical (F475W+F814W) images, and the bottom row shows grayscale F475W images. Labels present the object name along with the average S_{by-eye}, which runs on a scale of 1 (definite cluster) to 3 (unlikely cluster).

the initial classification work presented here.

Given the current limitations in assessing an object's boundedness, we adopt a liberal approach for cluster candidate identification. We prioritize sample completeness over purity and therefore include all cluster-like stellar overdensities as part of the object catalog. We acknowledge that the cluster catalog presented here will likely include both bound and unbound groupings of stars, particularly among the youngest objects. As a result, throughout this work all objects are formally considered cluster candidates, although for brevity we will refer to them simply as clusters.

With the intent of the cluster search well defined, we proceed to a description of our search methodology and a presentation of the cluster search results in Section 2.3.1. We conclude our discussion of cluster identification by characterizing the completeness characteristics of the sample in Section 2.3.2.

2.3.1 By-Eye Search

We undertake a systematic by-eye search of the Year 1 PHAT images following a precedent set by previous M31 cluster studies (e.g., Barmby & Huchra 2001; Krienke & Hodge 2007). Drawing on this rich history of visual cluster identification, we make a concerted effort to

improve upon previous work through the use of uniform analysis techniques, redundancy, cross-validation, and improved characterization of the resulting selection function.

When using any method of cluster identification, manual or automated, it is important to understand the biases and limitations inherent to the technique. We use artificial cluster tests in order to assess the completeness characteristics of our search methodology, which we discuss in detail in Section 2.3.2. We are also developing automated methods of cluster identification to further reduce subjectivity of the identifications and enable more robust completeness testing in future work with this dataset. The cluster sample presented here will act as an important comparison sample to help refine these automated search techniques for use with the PHAT dataset.

Our by-eye search consists of two stages: an initial search of the available imaging for all viable cluster candidates, followed by a re-evaluation of each preliminary candidate in a systematic manner. The completeness and accuracy of this process are enhanced by the redundancy of eight experienced astronomers conducting each stage of the search.

To perform the initial image search, the survey area is subdivided using the footprint of the WFC3-IR camera ($2.3' \times 2.1'$), providing 18 contiguous, minimally-overlapping search fields within each brick. Each field is searched by three or four team members in a "blind" manner, meaning that these individuals are not provided with the locations of clusters previously identified by other PHAT team members or previous surveys. To identify cluster candidates, searchers use a suite of images that include a two-band optical color image, all six available single-band images spanning from the UV to the NIR, and two star-subtracted optical images used to identify diffuse emission that makes up a cluster's unresolved light. These images are accessed using a custom image viewer that allows users to switch between the available spatially-aligned, full resolution images, as well as the ability to alter brightness and contrast levels for optimal visualization.

Once candidate objects are selected in the initial visual search, we cross-match and combine the identifications of all team members. Based on this preliminary catalog, we perform initial photometry and create cutout images for each cluster candidate. The preliminary catalog is reviewed independently by all eight team members on an interactive web site, and each individual assigns scores for each candidate, providing an assessment of the likelihood

that an object is a cluster. Scores (S_{by-eye}) are based on a scale of 1-to-3, where $S_{by-eye} = 1$ represents a definite cluster, $S_{by-eye} = 2$ signifies a likely cluster, and $S_{by-eye} = 3$ represents an unlikely or non-cluster object. We present examples of the cluster scoring system in Fig. 2.3.

We average the scores from all team members for each object and use these average rankings to divide the candidates into three subsamples: clusters, possible clusters, and unlikely objects. We choose thresholds of $S_{by-eye} < 2.0$ for clusters, $2.0 \leq S_{by-eye} < 2.5$ for possible clusters, and $S_{by-eye} \geq 2.5$ for unlikely objects. We discard unlikely objects from the catalog, while retaining clusters and possible clusters in two separate catalogs. The average scores for each cluster are provided as part of the cluster catalogs as an assessment of candidate quality. To determine the reliability of these average scores, we built a ranking experiment into our classification procedure. During the course of the ranking work, 24 objects appeared twice within the preliminary catalog. As a result, these clusters were each ranked two separate times by each team member. When the resulting average scores of the duplicate entries are compared, we find the standard deviation of the 1-to-3 ranking differences to be 0.27, showing good consistency and repeatability for the scores provided by our search team.

The Year 1 cluster search yielded a catalog of 601 high-scoring clusters. Table 2.1 presents positions for each object, as well as other descriptive information (radii, photometric measurements) that will be described in Section 2.4. Cutout images for each object are presented in Fig. 2.4. In addition, the spatial distribution of the clusters are shown in Fig. 2.2. Tabulated information and image cutouts for 237 possible clusters are presented in Table 2.2 and Fig. 2.5, respectively. Information about the possible clusters is provided for completeness, but due to the uncertain nature of their classifications, we exclude these objects from further analysis.

During the cluster search, we also identified 370 putative background galaxies. We did not explicitly search for these objects, therefore this catalog does not constitute a complete sample of objects. However, the potential usefulness of these identifications (e.g., as astrometric references, multi-wavelength source catalog cross-correlation) warrants its release. We present this catalog of objects in Section 2.9.1.

Table 2.1. PHAT Year 1 Cluster Catalog

PC ID S_{by-eye}	RA (J2000) R_{ap} ($''$)	Dec (J2000) R_{eff} ($''$)	F275W F814W	σ σ	F336W F110W	σ σ	F475W F160W	σ σ	ApCor[a] Alternate Name
1	11.638274	42.193887	17.54	0.02	17.68	0.03	18.82	0.02	-0.12
1.00	1.57	0.68	18.34	0.04	17.95	0.07	17.73	0.12	Hodge10-85
2	11.637139	42.209936	15.60	0.02	15.91	0.02	17.33	0.01	-0.00
1.43	2.51	0.64	17.23	0.04	17.39	0.19	17.30	0.35	Hodge10-84
20	11.630550	42.200631	...	...	22.55	0.25	21.97	0.63	-0.13
1.00	1.10	0.48	19.83	0.07	19.16	0.23	18.63	0.39	...
21	11.631591	42.199991	19.60	0.02	19.82	0.01	20.89	0.10	-0.21
1.57	1.00	0.52	20.98	0.28	22.88	1.80	...	...	...
22	11.630849	42.201656	23.38	0.36	22.52	0.11	22.22	0.06	-0.31
1.71	0.75	0.46	21.39	0.67	22.47	1.72	...	...	...

Note. — Table 2.1 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content. The sample presented here consists of objects classified as clusters, with $S_{by-eye} < 2.0$.

[a]Aperture Corrections are provided such that $m_{Total} = m_{Aperture} + ApCor$.

Table 2.2. PHAT Year 1 Possible Cluster Catalog

PC ID S_{by-eye}	RA (J2000) R_{ap} ($''$)	Dec (J2000) R_{eff} ($''$)	F275W F814W	σ σ	F336W F110W	σ σ	F475W F160W	σ σ	ApCor[a] Alternate Name
4	11.664001	42.192391	19.37	0.06	19.43	0.05	20.38	0.02	-0.09
2.43	2.10	0.83	19.80	0.14	19.42	0.22	18.63	0.10	...
6	11.668817	42.197452	21.50	1.01	20.93	0.13	21.86	0.29	-0.01
2.29	1.13	0.32	21.97	1.72	21.57	0.69	21.71	3.33	...
7	11.675393	42.193555	...	...	25.59	2.19	22.63	0.14	-0.07
2.29	0.80	0.29	20.96	0.13	20.38	0.11	19.73	0.08	...
11	11.648244	42.227958	25.42	1.35	23.98	0.24	22.56	0.15	-0.08
2.43	1.00	0.38	20.12	0.12	18.77	0.03	17.82	0.07	...
12	11.661360	42.190571	21.39	0.16	21.02	0.05	21.38	0.04	-0.02
2.00	1.05	0.31	20.74	0.06	20.82	0.13	21.39	1.90	...

Note. — Table 2.2 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content. The sample presented here consists of objects classified as possible clusters, with $2.0 \leq S_{by-eye} < 2.5$.

[a]Aperture Corrections are provided such that $m_{Total} = m_{Aperture} + ApCor$.

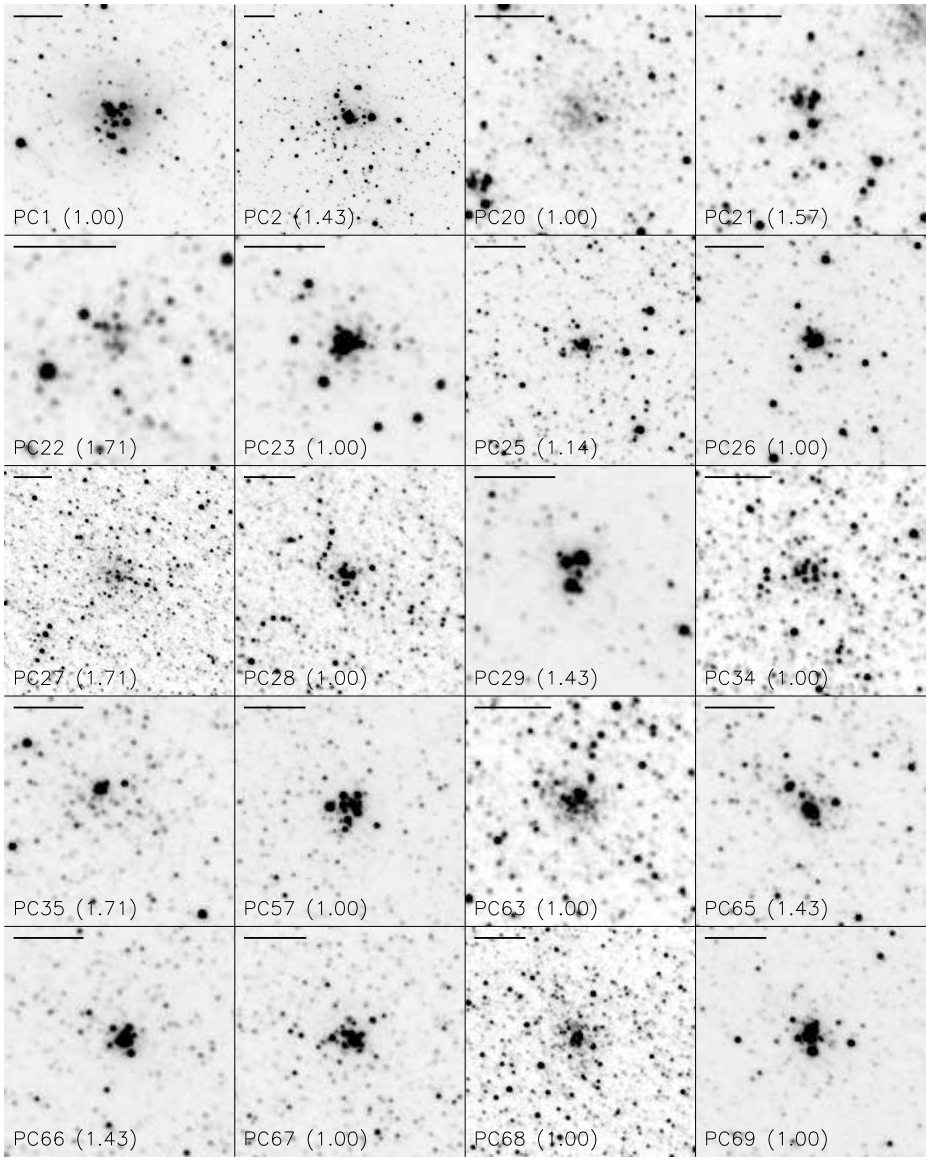

Figure 2.4 Cutout images for clusters in Table 2.1 with $S_{by-eye} < 2.0$. Images are grayscale F475W images, scaled to three times the cluster radius, and aligned such that North is up and East is left. Along with the PHAT cluster identifier, the average S_{by-eye} is provided for each object in parenthesis. The scale bar in each image represents $2''$. Figures 2.4.1–2.4.31 are available in the online version of the Astrophysical Journal.

2.3.2 Catalog Completeness

To characterize the completeness of our cluster sample, we conduct artificial cluster tests that mimic the selection procedure of the clusters. We use artificial clusters that span the

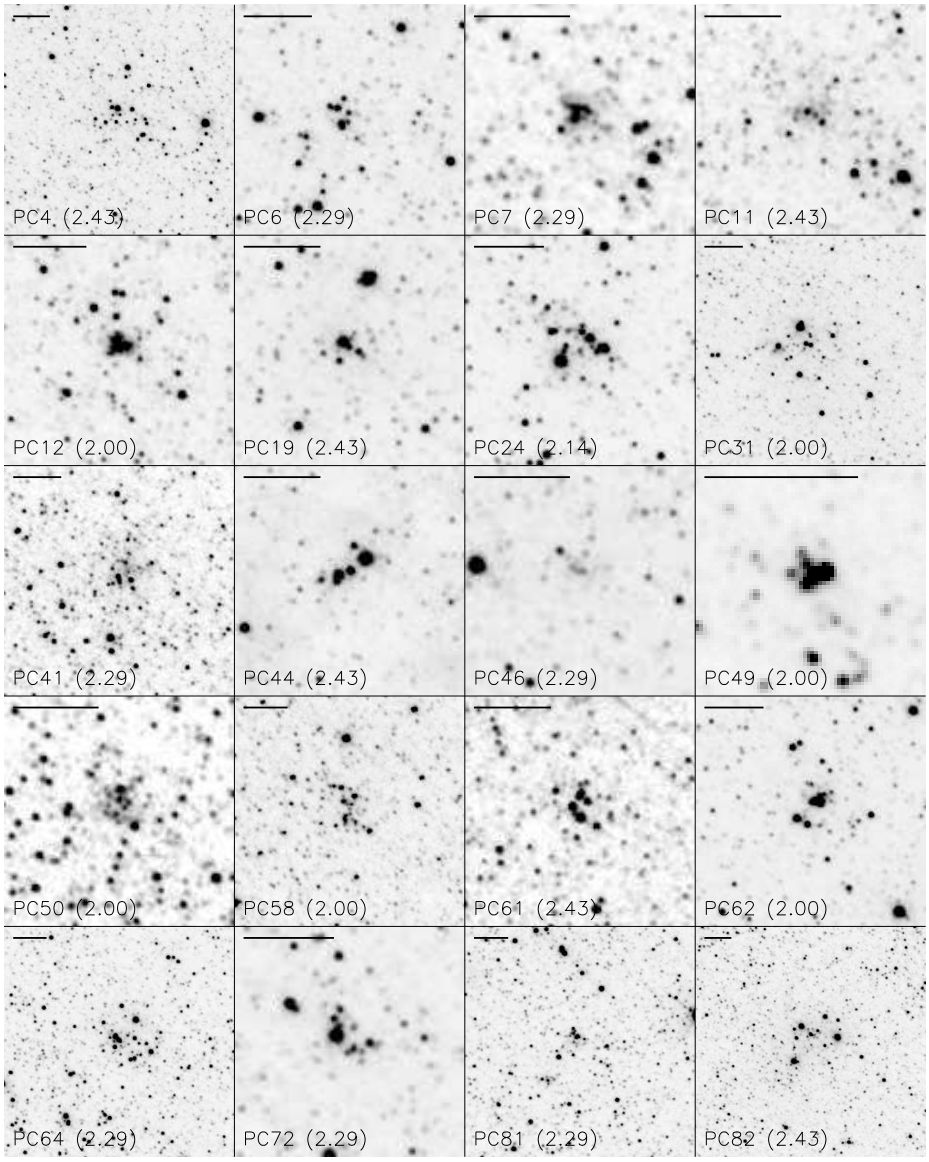

Figure 2.5 Same as Figure 2.4, but for possible clusters listed in Table 2.2 with $2.0 \leq S_{by-eye} < 2.5$. Figures 2.5.1–2.5.12 are available in the online version of the Journal.

range of ages and masses we expect to find in the cluster sample, while the size distribution is chosen to sample the minimum, average, and maximum sizes of the true sample (see Section 2.7.2). Ages and masses are chosen from a logarithmic grid of values, while sizes are drawn from the set of three characteristic values. Specifically, we select ages ranging from 4 million to 10 billion years, masses ranging from 10^2 to 10^5 $M_\odot$, and profiles that have

effective radii (R_{eff}; equivalently, half-light radii) of 1, 3 or 7 pc (0.26, 0.79, or 1.84 arcsec).

We create artificial clusters by populating a Padova isochrone (Girardi et al. 2010) of the appropriate age using a Kroupa (2001) stellar initial mass function. Next, the stars are spatially distributed according to a King (1962) profile, assuming no mass segregation. Finally, the size of the cluster and magnitudes of the individual stars are scaled appropriately to account for the distance of M31, assuming zero Galactic foreground or other internal M31 extinction. For each test field, we insert 46 artificial clusters into individual raw (FLT) images using new functionality developed for the DOLPHOT photometry package. The clusters are randomly positioned within each image to prevent search bias (as would result from a regular grid pattern), though we ensure that clusters are well-separated within the image so that they do not overlap or interfere with the photometry of other artificial clusters. We drizzle the resulting images together to create searchable images in the optical F475W and F814W passbands. In all, we created three fields of artificial clusters for each of the four full bricks in the Year 1 dataset, resulting in a total of ∼550 simulated objects. In addition to the completeness tests that follow, these artificial clusters are also used for quality assurance of our photometric results, as discussed in Section 2.4.2.

After creating the artificial cluster images, we identify clusters in the same way we searched the original images, with all eight team members searching each field and rating the reliability of the clusters. By comparing the resulting cluster identifications to the full list of inserted clusters, we can estimate the completeness limits of the sample based on various cluster input parameters; the results of this analysis are shown in Fig. 2.6. The left panel shows the best estimate of the sample's characteristic cluster recovery fraction, computed using a subsample of 110 "typical" simulated clusters with R_{eff} of 3 pc distributed within the outer three bricks (B09, B15, & B21). We estimate a 50% completeness limit of m_{F475W} ∼21.2. Considering distance and Galactic foreground dust reddening ($E(B-V) =$ 0.062; Schlegel et al. 1998), this translates to an absolute magnitude limit of M_{F475W} ∼ -3.5.

The true completeness of the sample as a whole, however, is a complicated function of cluster luminosity, size, and location within M31. Using the full complement of simulated clusters, we explore the dependence of completeness on the latter two cluster properties.

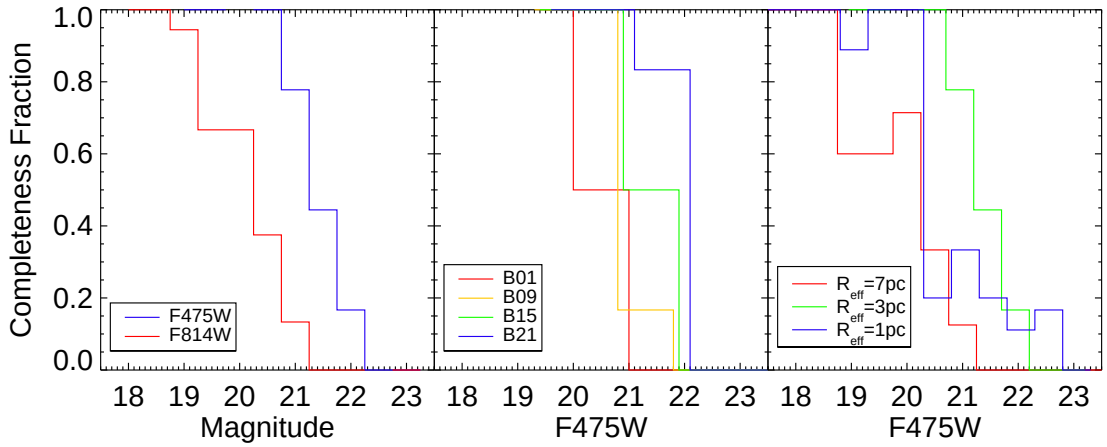

Figure 2.6 Cluster completeness derived from artificial cluster tests. Left: Completeness in the two optical filters, derived from 3 pc artificial clusters inserted in B09, B15, and B21. Center: Variation in completeness as a function of brick membership, derived from $R_{eff} = 3$ pc artificial clusters. Right: Variation in completeness as a function of R_{eff}, derived from artificial clusters inserted in B09, B15, and B21.

The middle panel of Fig. 2.6 shows that variations in galactocentric position result in a $\sim$1 magnitude difference in completeness limits. Detection limits are brighter for clusters located in the bulge-dominated inner galaxy due to luminous, crowded background fields. Cluster size also plays a role, in which increasing size at constant luminosity results in reduced surface brightness and lower detection efficiency. The right panel of Fig. 2.6 shows that there is a $\sim$1 magnitude difference in completeness when comparing clusters with R_{eff} of 1 versus 7 pc. To account for these multiple dependencies, we will undertake a larger and more rigorous set of completeness tests in future work to better characterize the subtleties of the completeness function.

2.4 Integrated Photometry

We use aperture photometry to measure the six band integrated fluxes of each cluster. Aperture photometry consists of two main analysis tasks: defining an aperture (center and size) and determining the local background flux level. We build upon photometry procedures used in M31 by Barmby & Huchra (2001) and Krienke & Hodge (2007), with

Table 2.3. Photometric Zeropoints

Passband	*Vegamag* Zeropoint
F275W	22.65
F336W	23.46
F475W	26.16
F814W	25.52
F110W	26.07
F160W	24.70

several refinements related to the assessment of local background levels and accounting for light that lies outside the photometric aperture. A detailed description of photometric analysis procedures and results is provided in Section 2.4.1, followed by artificial cluster validation analysis in Section 2.4.2. Photometry zeropoints for the ACS and WFC3 cameras were obtained from the STScI webpage[2] and are listed in Table 2.3. All photometry is presented in the Vega magnitude system using the native HST passbands; we do not perform passband conversions.

2.4.1 Aperture Photometry Procedure

The first step in aperture photometry is to define an aperture center. Centers are estimated by centroiding on a F475W image that is smoothed using a $0.3''$ (6 pixel) FWHM Gaussian kernel. For certain clusters, particularly low luminosity objects whose light is dominated by a small number of bright sources, the flux-weighted positions determined by the automated procedure do not always accurately reflect the cluster center. For this reason, we visually inspect central positions and manually adjust incorrect determinations.

We adopt a photometric aperture size that provides the largest signal-to-noise ratio for

[2] `http://www.stsci.edu/hst/acs/analysis/zeropoints` and `http://www.stsci.edu/hst/wfc3/phot_zp_lbn`

flux measurement by enclosing a maximum amount of cluster light while including as little background light as possible. The clusters considered in this study vary by a factor of ∼10 in radius, and consequently, the chosen photometric aperture radii vary by the same amount. We define circular apertures using growth curve analysis to determine an appropriate radius. The aperture limit is defined at the radius where the cluster profile drops below the level of the noise in the background, equivalent to the point at which the curve of growth turns over and the increase in cumulative flux as a function of radius stops. An illustrative example of a cluster image and growth curve is provided in Fig. 2.7. The aperture radii (R_{ap}) are determined by visual inspection of the growth curves for each cluster and are reported with the photometric results. While it would be preferable to adopt an algorithmic approach for defining R_{ap}, the relatively noisy character of the local background significantly complicates automated determinations. We perform aperture definition and growth curve analysis on the F475W image, which provides the best combination of signal-to-noise and contrast between cluster and field populations for a wide range of cluster ages. Apertures of the same angular size are used for the five other images. Photometric aperture radii range between $0.5''$ and $6''$ for the Year 1 sample.

Aperture photometry depends greatly on the determination of the background flux level. Following traditional photometric terminology, we also refer to non-cluster background light as the "sky" or the "sky background". For the PHAT cluster sample, the sky is made up of two components: individual resolved stars and unresolved light. Traditionally, the background is determined using the mode of sky region pixel values. However, the fact that resolved stars are a true component of the background light calls for an alternate statistical treatment.

We define ten annular sky regions that encircle the photometric aperture, extending radially from $1.2 \times R_{ap}$ to $3.4 \times R_{ap}$. Each of the annuli are equal in area to the photometric aperture, to accurately measure the dispersion of the sky background based on an equal number of pixel samples. Figure 2.7 shows an example of the aperture layout, where we denote the inner and outer extent of the sky measurement annuli. Next, we calculate the total integrated flux within each of the sky regions. We perform iterative 2σ rejection on these ten fluxes, thereby excluding regions that contain bright stars or other objects not

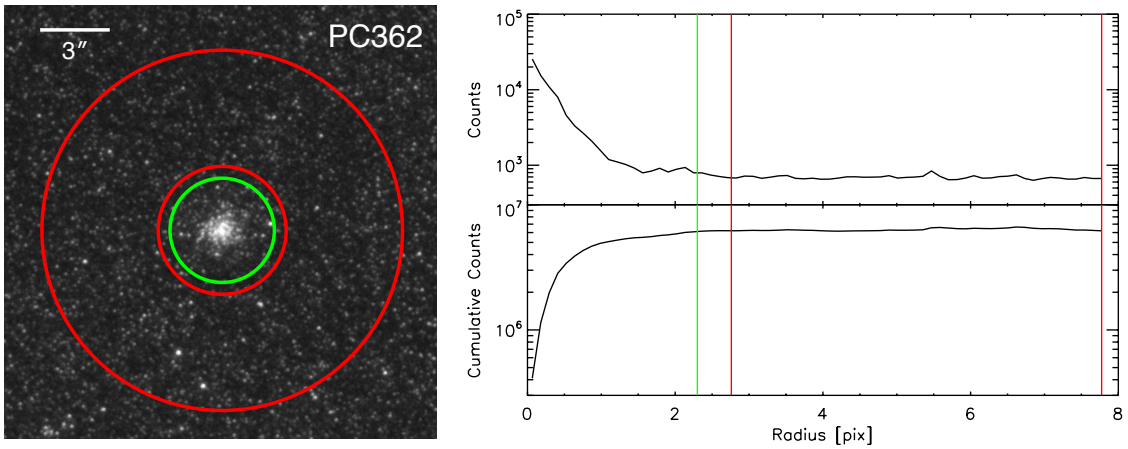

Figure 2.7 Cluster aperture layout and growth curve for PC362. Left: F475W image of cluster with aperture locations overplotted. The green circle denotes the photometric aperture. The red circles denote the inner and outer edges of the annulus used for background determination; this region is subdivided into ten equal-area annular subregions to estimate the variance in the background. Right: The flux profile (top) and cumulative growth curve (bottom) for the cluster; vertical red and green lines correspond to the radii plotted in the left panel.

representative of the typical sky background. We adopt the mean of the non-rejected sky fluxes as the sky background value, and the standard deviation of these fluxes as a measure of the uncertainty in the sky background determination. We propagate the uncertainty from the sky level determination into the overall cluster photometry by adding it in quadrature with the cluster's flux measurement uncertainty.

In agreement with the HKC cluster studies, the uncertainty in the sky background determination dominates the overall uncertainty in cluster magnitudes. When compared to previous ground-based cluster photometry in M31 or other HST-based extragalactic cluster photometry, our magnitude uncertainties appear larger. Our inclusion of sky level uncertainty into the reported values account for these larger overall magnitude uncertainties. Uncertainty in the determination of cluster fluxes, comparable to the errors reported in other cluster catalogs, never rise above 0.01 mag for any object in the Year 1 catalog.

The resulting magnitudes measured within R_{ap}, hereafter referred to as *aperture magnitudes*, are presented in Table 2.1 for the cluster sample and Table 2.2 for the possible cluster

sample. These aperture magnitudes represent high signal-to-noise, spatially matched measurements of cluster light, and are optimal for calculating cluster colors. We plot the photometric uncertainties as a function of magnitude for the cluster sample in Fig. 2.8. To assess relative reliability in the six passbands, we count the number of well-determined ($\sigma < 0.5$ mag) photometric measurements in each band, and record the results in Table 2.4. The F475W imaging provides the highest-quality measurements, as shown by that filter's small photometric uncertainties, followed by the F336W image which probes a similar wavelength regime. The F814W image has increased levels of photometric error due to the reduced contrast between cluster and field populations. The quality of the measurements is lower in the three remaining filters due to intrinsic wavelength-dependent limitations; the F275W measurements suffer from low signal-to-noise for all but the youngest, bluest clusters, while the F110W and F160W measurements suffer from high sky background levels and poor cluster-field contrast. In addition to the problem of faint signal, we note that $\sim$10% of F275W magnitudes (and F336W magnitudes, but at a lower level) are affected by cosmic ray artifacts. While a vast majority of these defects are adequately corrected for by our image processing, we caution that a few percent of the UV magnitudes might still be adversely affected.

In addition to the aperture magnitudes, we also provide estimates of the effective radius (R_{eff}; equivalently, the half-light radius) for each cluster in Tables 2.1 and 2.2. These estimates are obtained by measuring the cluster flux profile and interpolating this curve to find the radius that contains half of the light within R_{ap}. We use the F475W cluster light profiles to make these assessments, given their good data quality. We discuss the resulting sizes in Section 2.7.2 and use them to calculate aperture corrections in the next section. We recognize that the R_{eff} estimates will systematically underestimate the true cluster sizes due to the fact that they are calculated using aperture magnitudes that fall short of measuring the full luminosity of the cluster. However, we expect the impact of this underestimation to be small due to the steepness of the cluster light profile.

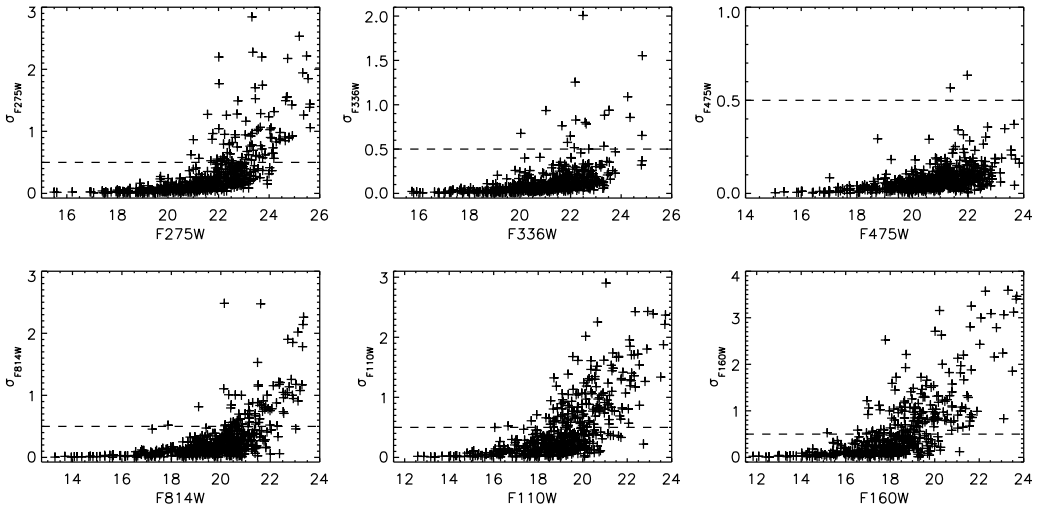

Figure 2.8 Photometric errors for integrated cluster photometry in each of the six PHAT passbands. The scale of the y-axis varies between panels. A standard reference at $\sigma = 0.5$ mag is provided for comparison.

Table 2.4. Passband Photometric Quality Comparison for Cluster Sample

Passband	N(Valid Measurements)	N(Well-determined Measurements)
F275W	552 (91.8%)	447 (74.4%)
F336W	590 (98.2%)	566 (94.2%)
F475W	600 (99.8%)	597 (99.3%)
F814W	593 (98.7%)	514 (85.5%)
F110W	518 (86.2%)	358 (59.6%)
F160W	472 (78.5%)	313 (52.1%)

Note. — Valid measurements denote magnitudes that result from positive fluxes (signal measured above sky level) and suffer no other failures (e.g., image artifacts). Well-determined measurements denote magnitudes where $\sigma < 0.5$ mag.

2.4.1.1 Aperture Corrections

The aperture magnitudes presented above measure a majority of the cluster light. However, these values do not account for light that lies below the noise level of the sky background, beyond the limits of our photometric aperture in the faint outer wings of the cluster profile. We correct for this missing flux by using R_{eff} determinations to approximate the cluster's luminosity profile shape, then extrapolate this profile past the limits of our photometric aperture to make an estimate of the cluster's total light. The magnitude difference calculated between the original aperture magnitudes and this total light estimate are equivalent to an aperture correction.

To make this profile-based extrapolation, we require a cluster profile shape and a normalization for that profile. We adopt a King (1962) profile, assuming a concentration ($c = R_{tidal}/R_{core}$) that matches the characteristic profile of PHAT clusters, as obtained during preliminary cluster profile fitting ($c = 7$). Next, we use a cluster's R_{eff} to set the radial scaling of that characteristic King profile. Finally, we normalize the scaled profile such that the integrated flux within R_{ap} matches the cluster's aperture magnitude measurement. Once normalized, we calculate the fraction of flux that lies outside R_{ap} and transform this value into an aperture correction in magnitudes, which is presented in Tables 2.1 and 2.2 for all clusters. These corrections may be applied to the aperture magnitudes to obtain *total magnitudes*[3]. These total magnitudes are optimal for the estimation of total cluster magnitudes and luminosities. While the aperture corrections were derived in the F475W passband, they may be used for all filters under the simplifying assumption of flat radial color profiles in the outer parts of the cluster.

The amplitude of the aperture corrections are presented in Fig. 2.9. Over the sample of clusters, the corrections vary from 0.0 to -0.6 mag, with a median correction of -0.1 mag. These corrections are negligible for the brightest clusters, where a majority of the light is detectable above the noise level of the sky background. The corrections become larger for fainter clusters, due to their low cluster-to-field flux contrast. The simplifying assumption of a universal cluster profile shape that varies only as a function of R_{eff} provides suitable

[3]Total Magnitude = Aperture Magnitude + Aperture Correction

accuracy for this correction, as shown by artificial cluster tests that follow in Section 2.4.2. Full cluster profile and structural parameter fitting is currently underway (M. Fouesneau et al., in preparation), and these results could be used to further refine these profile extrapolations and improve upon our estimates of total cluster light in future work. However, we expect little overall improvement in photometry as a result of increased aperture correction precision because these corrections are comparable in size to the amplitude of the photometric uncertainties for most clusters.

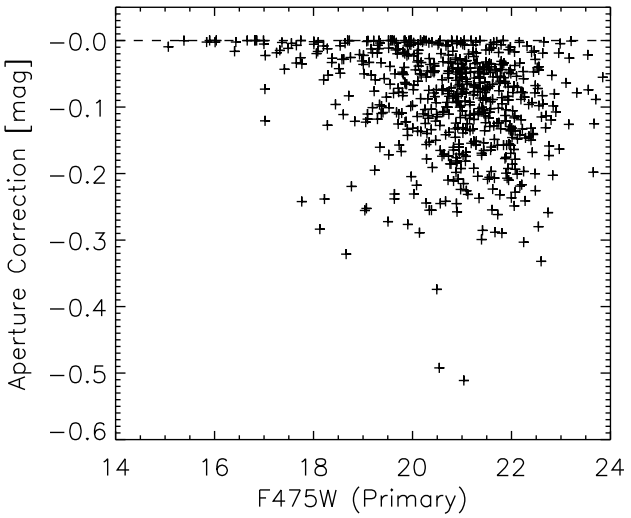

Figure 2.9 Aperture corrections, used for converting aperture magnitudes to total magnitudes. The corrections are derived from cluster profile extrapolation.

2.4.2 Artificial Cluster Photometry Experiments

We use artificial clusters to assess the uncertainties and biases associated with our photometry procedures. For these tests, we only consider clusters that lie in the outer three bricks, have R_{eff} of 3 pc, and were detected as part of our artificial cluster search in Section 2.3.2. Analysis of this sample of 76 artificial clusters provides an evaluation of photometric accuracy for "typical" clusters in the Year 1 sample. We process the simulated objects using photometric procedures identical to those described in the previous section.

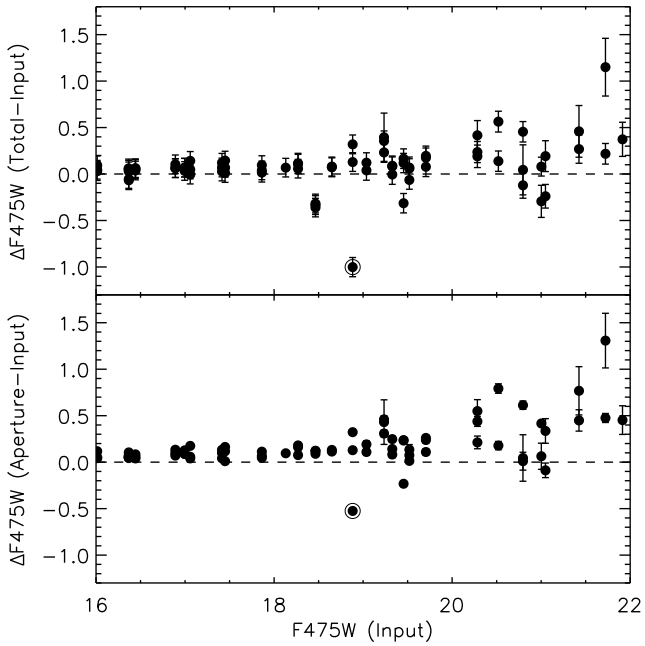

Figure 2.10 F475W magnitude differences between input and recovered magnitudes for a subset of simulated clusters inserted into B09, B15, and B21, adopting a R_{eff} of 3 pc. Top: Differences between the measured total magnitudes and the input magnitudes. Bottom: Differences between the measured aperture magnitudes and the input magnitudes. The lowest outlier in both panels (circled) differs due to a bright (F475W$\sim$19) nearby field star that overlaps the cluster.

The results of these tests are shown in Fig. 2.10. The artificial cluster tests show good agreement (ΔF475W $\lesssim$0.1 mag) between the input and output magnitudes for bright clusters (F475W < 19). For less luminous clusters, aperture magnitudes are fainter than input magnitudes by up to $\sim$0.4 mag at F475W $\sim$ 21.5. However, total magnitudes are more successful in recovering input magnitudes accurately, showing a smaller faintward bias of $\sim$0.2 mag at F475W $\sim$ 21.5. This bias represents a $\sim 0.5\sigma$ deviation when compared to the $\sim$0.4 mag scatter in the photometry. Finally, we acknowledge that these photometric experiments are slightly idealized, for instance due to our consideration of only a single family of cluster profile shapes. However, this testing confirms that our photometry techniques provide accurate assessments of cluster properties and their associated uncertainties.

2.5 Comparison to Existing M31 Cluster Studies

As discussed in the introduction, there is a long history of stellar cluster studies in the Andromeda galaxy. Decades of effort have produced a wealth of knowledge on this topic. To place our findings in context, in this section we cross-reference our cluster identifications with existing catalogs, allowing us to reference previous work on the same objects and compare the results of our cluster analysis to existing catalogs.

To compile a list of known clusters located within the Year 1 footprint, we began by searching the Revised Bologna Catalog[4] (RBC; Galleti et al. 2004, last updated 2009 December to v4.0). This excellent resource has aggregated all known cluster identifications from early catalogs (e.g., Vetešnik 1962; Sargent et al. 1977; Crampton et al. 1985; Battistini et al. 1987, 1993; Barmby et al. 2000, among many others) as well as more recent works (e.g., Kim et al. 2007; Huxor et al. 2008; Caldwell et al. 2009). Other than the RBC, we searched the HKC and other works published since the most recent RBC revision (Vansevičius et al. 2009; Peacock et al. 2010; Fan et al. 2010). Our search of the HKC produced additional objects for cross-matching, however no new objects were recovered from the three other catalogs. The Peacock et al. (2010) catalog is composed solely of edits and reclassifications from an earlier version of the RBC (v3.5). Similarly, the objects studied in Fan

[4] `http://www.bo.astro.it/M31/`

et al. (2010) are derived directly from v4.0 of the RBC. Finally, we find no overlap between any part of the PHAT footprint (existing or planned coverage) and the southwest region of M31 studied by Vansevičius et al. (2009). In the discussion that follows, we consider cluster classifications with respect to those provided in the RBC v4.0.

The RBC and HKC contain a total of 146 published clusters, 32 cluster candidates, and 84 other non-cluster classifications (foreground stars, background galaxies, and HII regions) that lie within the Year 1 PHAT survey footprint. We cross-match these previously known objects with all identifications made as part of the Year 1 search (cluster, possible clusters, and unlikely objects). Further, we re-examined the PHAT data at positions of previously cataloged objects that were not matched to PHAT candidates to ensure the complete re-analysis of all existing catalog entries that lie within the Year 1 footprint. In total, we classify 132 of the previously known objects as clusters, 12 as possible clusters, and reject the 118 remaining objects. We note that all confirmed Year 1 PHAT clusters were identified independently as part of our blind cluster search. Tables 2.1 and 2.2 provide cross-matched names of clusters using the naming convention of Barmby et al. (2000), consisting of the identifier from the Revised Bologna Catalog, followed by the identifier of the next most significant cluster catalog, where such exists. Additional details concerning the comparison and reclassification of existing cluster identifications is provided in Section 2.9.2. There we provide catalog-specific commentary on the validity of the previously published classifications, as well as object-by-object classification revisions for the RBC and the HKC.

A comparison between ground-based and space-based M31 cluster catalogs reveals the importance of high spatial resolution imaging in cluster identification work. In Fig. 2.11, we present a histogram of the number of cluster as a function of apparent magnitude for the Year 1 cluster sample of 601 clusters and the 132 previously known objects confirmed as clusters in the PHAT data. It is immediately apparent from the plot that HST imaging has allowed the PHAT cluster survey to identify hundreds of low luminosity clusters that could not be identified from ground-based data. As shown in the inset of Fig. 2.11, the sample completeness associated with previously known clusters discovered from ground-based data (green dashed histogram) drops precipitously at $m_{F475W} > 18$ ($M_{F475W} > -6.5$). This fall-off in completeness reflects the difficulty in differentiating between single unresolved stars

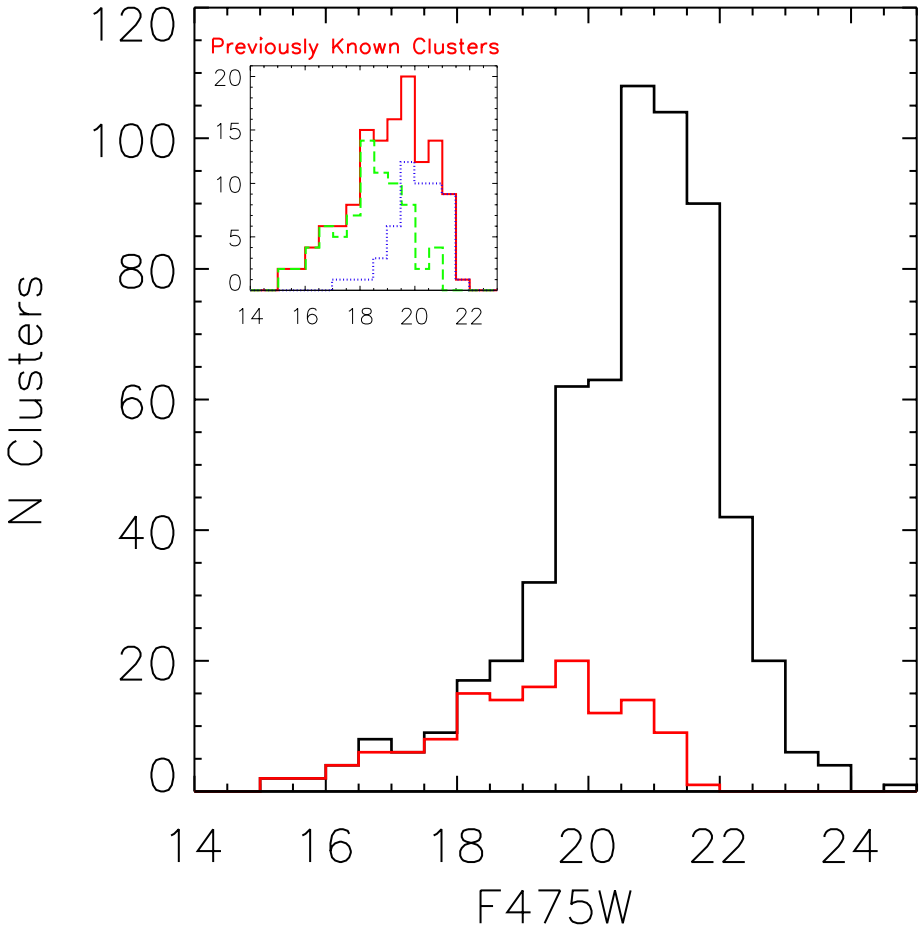

Figure 2.11 Magnitude histogram comparing the Year 1 PHAT clusters (black) to objects from existing catalogs that fall within the Year 1 survey footprint (red). The inset shows a breakdown of the previously known clusters into objects that were identified using ground-based imaging (green) and those that were identified using HST imaging (blue).

and compact clusters in low resolution images. In contrast, high spatial resolution imaging from HST enables the identification of clusters 2-3 magnitudes fainter than previous ground-based surveys. Within this fainter luminosity range, Fig. 2.11 shows an order of magnitude increase in the number of objects identified in the PHAT cluster catalog when compared to previous HST-based cluster survey work (HKC; blue dotted histogram in inset). This improvement results from the order of magnitude increase in spatial coverage provided by PHAT when compared to the limited number of previous, targeted HST observations that fall within the Year 1 survey footprint (35 arcmin^2 versus 390 arcmin^2 in PHAT Year 1).

In addition to the catalog comparison presented here, in Section 2.9.3 we compare photometry results presented in this work to those of existing catalogs. This analysis acts as quality assurance for the photometry presented here, and provides the reader with an assessment of the inherent differences between the sets of photometric results.

2.6 Year 1 Clusters: Photometric Properties

The Year 1 cluster sample, derived from $\sim$1/4 of the total expected PHAT survey data, provides the first glimpse of what can be expected from the full balance of the PHAT stellar cluster survey. In Section 2.5, we showed that our catalog represents a considerable increase in the number and diversity of clusters known in M31. Our excellent HST-based imaging should lead to factor of $>$4 increases in the number of known clusters within the PHAT survey footprint.

To obtain a better sense for the type of clusters we have identified in the Year 1 sample, we present color and magnitude distributions from the catalog photometry. For this analysis, we select a subset of objects with well-determined photometry, where the uncertainties in the F336W, F475W, and F814W magnitudes are each less than 0.5 mag. This quality cut results in a subsample of 482 well-characterized objects that we use to explore the properties of the catalog.

We plot a color-color diagram for clusters with well-determined photometry in Fig. 2.12. This diagram aids our ability to assess the cluster age distribution. To provide reference points to guide the eye, we overlay the stellar evolution model predictions from the Padova group (Girardi et al. 2010). These model tracks assume solar metallicity and are reddened

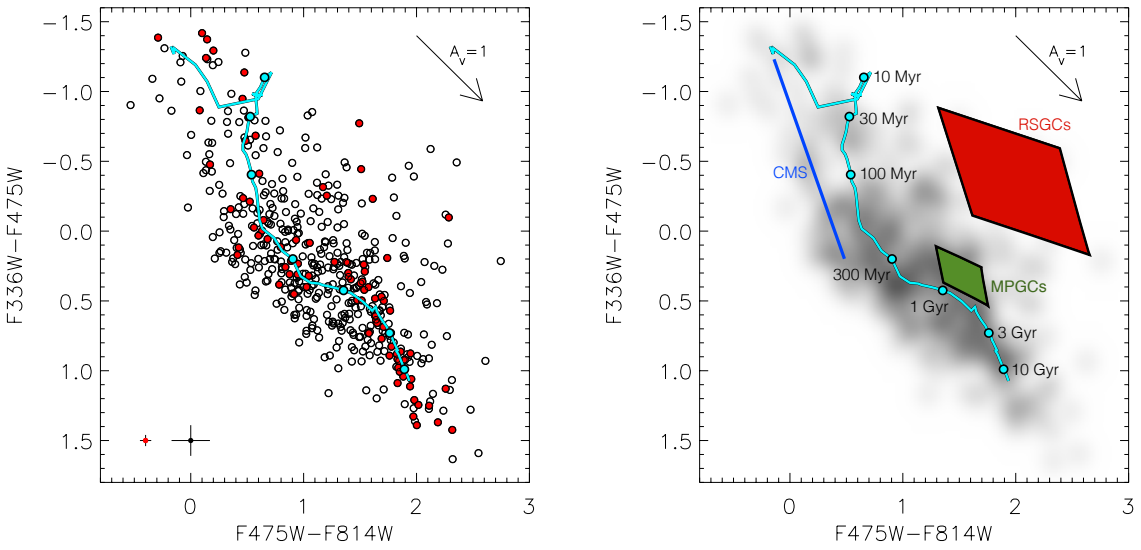

Figure 2.12 Left: Color-color diagram for the 482 Year 1 clusters with well-constrained ($\sigma < 0.5$ mag) photometry in the F336W, F475W, and F814W passbands. A subset of the 92 most luminous clusters (F475W < 19.5) are highlighted in red. Padova SSP models for Solar metallicity, reddened to account for Galactic foreground extinction, are plotted as a cyan line for reference. Cyan points act as age indicators, spaced at 0.5 dex increments beginning at 10 Myr. Median error bars for the bright (red) and faint (black) cluster samples are displayed in the lower left corner. Right: A smoothed, grayscale version of the color-color diagram, with SSP models shown again as reference. The red box denotes the location of anomalous red supergiant clusters (RSGCs), which contain luminous evolved stars that strongly affect the cluster's integrated colors. The green box denotes the parameter space populated by metal poor globular clusters (MPGCs). The blue line denotes the modeled location of the cluster main sequence (CMS), reddened by A_V=0.4 mag to match the color distribution of the Year 1 cluster sample. The CMS represents the color-color sequence populated by low mass clusters that host no evolved stars due to stochastic sampling of the cluster's stellar mass function.

to account for foreground Galactic extinction ($E(B-V) = 0.062$; Schlegel et al. 1998). On this plot, ages increase as we follow the evolutionary track from the upper left to the bottom right. In addition to the initial cluster selection based on photometric uncertainties, we define a second subset of clusters based on cluster luminosity. The red points denote the well-determined subsample's 92 most luminous clusters, with F475W < 19.5.

The first conclusion we draw from Fig. 2.12 is that the Year 1 cluster sample includes a wide range of ages. Clusters populate the full length of the model evolutionary track, with a large number of objects populating an intermediate age range (300 Myr to 3 Gyr). However, the mapping from position in the color-color diagram to age suffers from well-known degeneracies with extinction and metallicity. For example, old (12-14 Gyr) metal-poor ([Fe/H] $\lesssim -1.0$) globular clusters inhabit the same position on this diagram as $\sim$100 Myr old, solar metallicity clusters with A_V of $\sim$1.5 mag. The green region in the right panel of the figure denotes the shared color-color region where this particular age-metallicity degeneracy exists.

Second, we observe that the cluster color distribution is affected by the effects of stochastic stellar mass function sampling in low-mass ($< 10^4\ M_\odot$) clusters. The red region in the right panel of Fig. 2.12 highlights color outliers that have anomalous red F475W-F814W colors. While young clusters (<50 Myr) suffering from large amounts of dust attenuation (A_V >2 mag) can populate this region of the diagram, these red colors are more frequently caused by the presence of a small number of bright evolved supergiant stars that can bias integrated cluster colors. We discuss this behavior in greater detail in Section 2.7.3.1. In addition to the red outliers, stochastic effects can also cause integrated cluster colors to appear bluer than model predictions. The fluctuation in the small number of evolved stars sometimes results in the complete absence of supergiant cluster members, meaning that the cluster's integrated light is emitted exclusively by main sequence stars. Such clusters fall onto a linear sequence in color-color space we refer to as the cluster main sequence. We highlight this feature in blue in the right panel of Fig. 2.12.

Next, we plot a cluster color-magnitude diagram in Fig. 2.13 to assess the cluster mass range probed by the Year 1 sample. As in Fig. 2.12, we plot foreground reddened, solar metallicity Padova stellar models for reference. The PHAT clusters span $\sim$8 mag in F475W

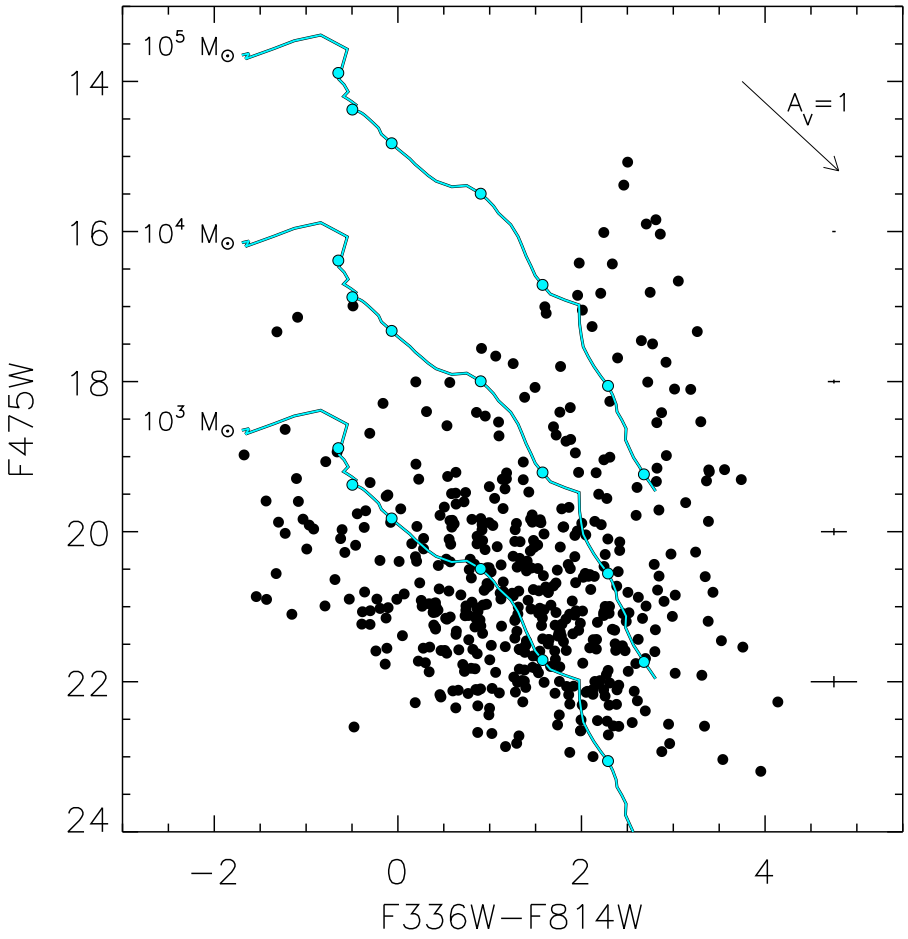

Figure 2.13 Color-magnitude plot for the 482 Year 1 clusters with well-constrained ($\sigma < 0.5$ mag) photometry in F336W, F475W, and F814W passbands. Padova SSP models for three cluster masses (10^5, 10^4, 10^3 $M_\odot$) at Solar metallicity are plotted for reference, with age indicators spaced at 0.5 dex increments beginning at 10 Myr. Characteristic median error bars are shown on the right side of the plot, where each point describes the uncertainties of points within a 2 mag bin in luminosity.

luminosity, translating to >3 orders of magnitude in cluster mass. This range indicates that the PHAT cluster sample hosts a wide variety of clusters, spanning systems that contain hundreds of solar masses up to those with a million solar masses. The most luminous clusters, however, all appear to have red colors (F336W-F814W $\sim$ 2.5), forming a vertical sequence of objects on the right side of the plot. This results from the fact that most massive ($> 10^5\ M_\odot$) clusters in the Year 1 sample are old globular clusters.

Figures 2.12 and 2.13 show that the sample of clusters assembled from the PHAT dataset provide a top-to-bottom assessment of the M31 cluster population. Few datasets have the ability to sample objects across a variety of stages in cluster evolution over such a large, uninterrupted mass range. This diversity makes the PHAT cluster sample a valuable tool to better understand the formation and dissolution of star clusters.

2.7 Discussion

2.7.1 Luminosity Functions

Luminosity functions are a basic, model independent measure of a stellar cluster population. Up to this point, observations of spiral galaxies have provided largely consistent results, where luminosity functions are well fit by a power-law ($N \propto L^{-\alpha_L}$) with indices of about -2 or steeper (Larsen 2002; Gieles 2010; Chandar et al. 2010c). In addition, there is growing evidence of steeper slopes at the brightest cluster luminosities ($M_V < -9$; Whitmore et al. 1999; Gieles et al. 2006a; Haas et al. 2008), suggesting the possibility of a Schechter-like cluster mass function with an exponential truncation at the high-mass end (Larsen 2009).

These previous studies have generally focused on the bright end of the luminosity function, fitting clusters with luminosities greater than $\sim 2 \times 10^4\ L_\odot$ (equivalent to $M_V \lesssim -6$). In contrast, the faint end of luminosity function has relatively poor constraints due to detection and catalog completeness difficulties encountered by previous extragalactic cluster studies. However, due to our ability to identify low luminosity clusters in PHAT survey data, we can use the Year 1 cluster sample to probe the shape of the luminosity function down to limits only previously accessible in the Magellanic Clouds . These new constraints are interesting because of the sensitivity of the faint end luminosity function to cluster

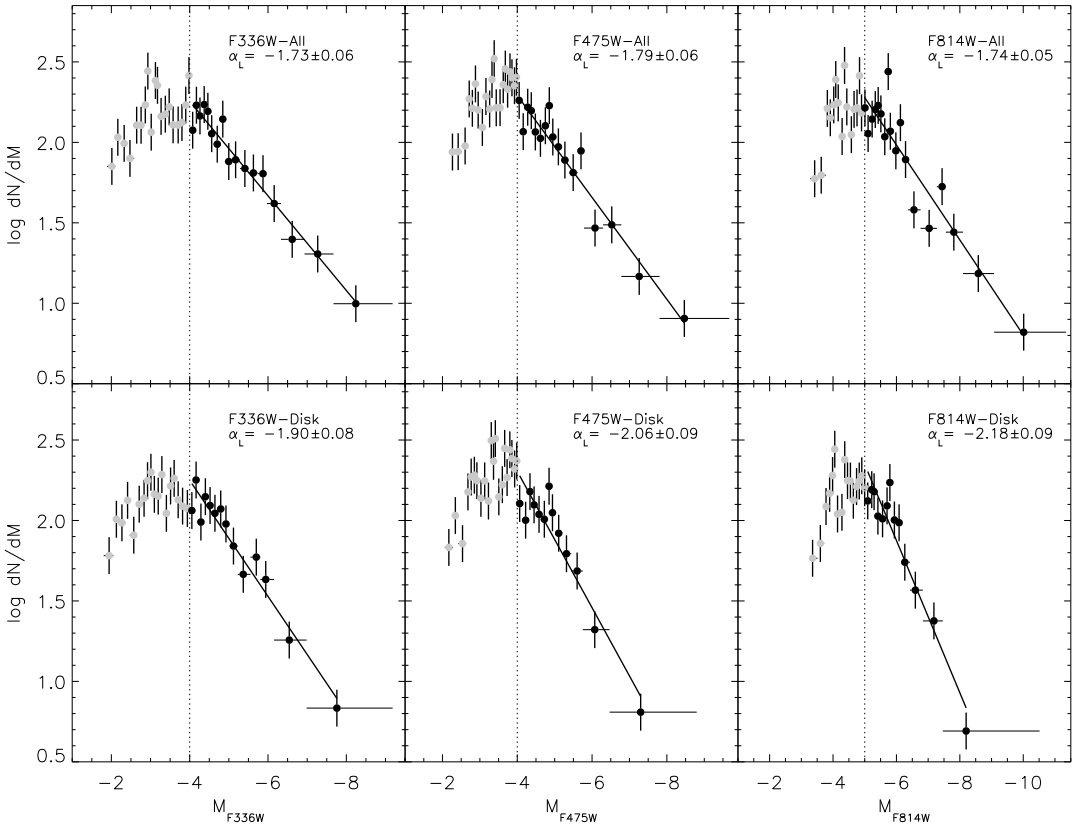

Figure 2.14 Luminosity functions for the F336W (left), F475W (center), and F814W (right) passbands. Top panels display the luminosity functions for clusters drawn from the complete Year 1 sample, and the bottom plots show the luminosity functions considering only disk clusters that lie outside the bulge (i.e., excluding B01 clusters). Only clusters with well-constrained ($\sigma < 0.5$ mag) photometry are included in the analysis. Clusters have been corrected for Galactic foreground dust attenuation, but not attenuation within M31. Data is plotted using variable binning, such that each point represent an equal number of clusters (N=15). Vertical error bars represent Poisson uncertainties due to the number of objects per bin, and horizontal error bars represent the magnitude range of each bin. The luminosity function slope is fit down to completeness limits denoted by vertical dotted lines at -4.0 for F336W and F475W, and -5.0 for F814W.

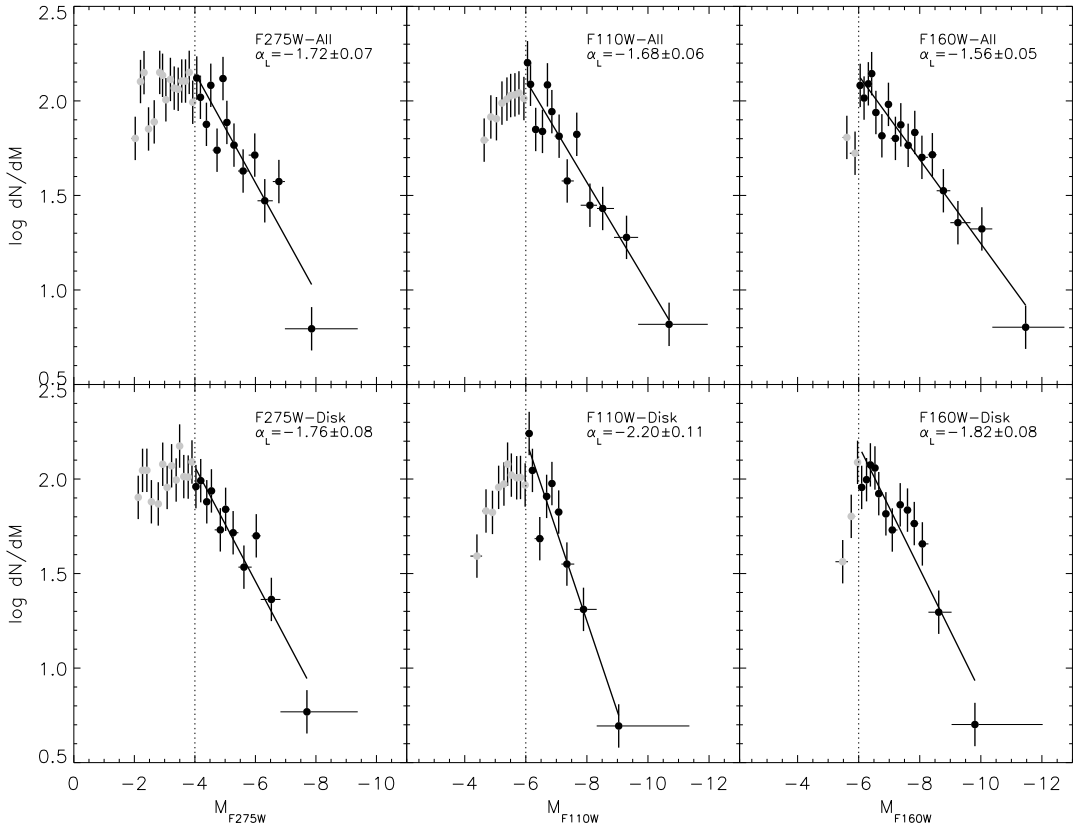

Figure 2.15 Same as Figure 2.14, but for the F275W (left), F110W (center), and F160W (right) passbands. Completeness limits are -4.0 for F275W and -6.0 for F110W and F160W.

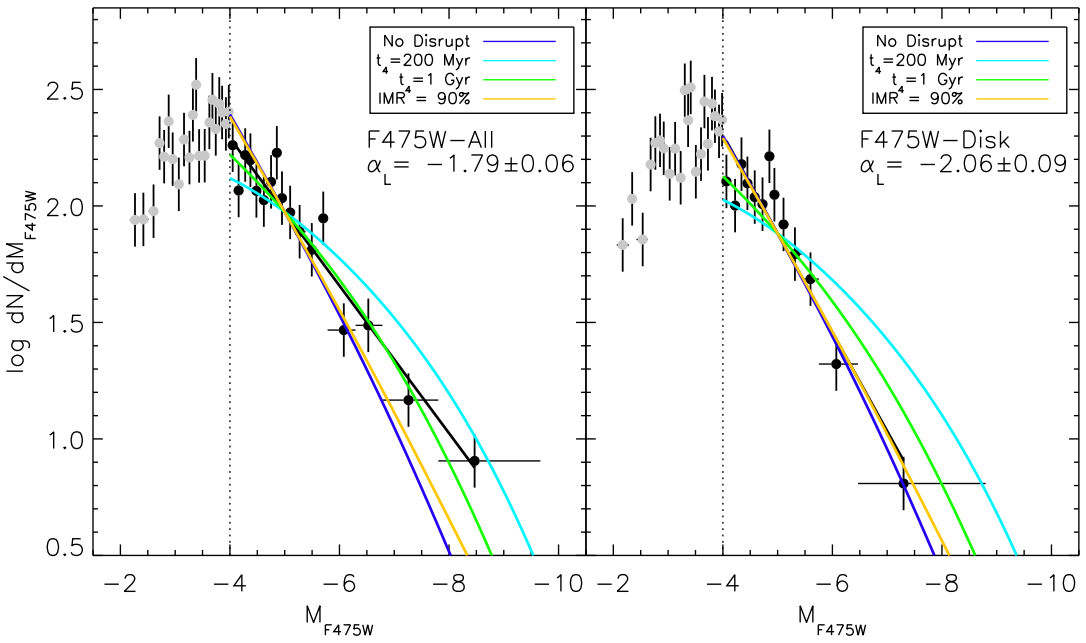

Figure 2.16 F475W luminosity functions from Fig. 2.14 with model luminosity functions from Larsen (2009) overplotted. The left panel shows the luminosity function constructed using all Year 1 clusters, and the right panel shows the luminosity function constructed using a subset of Year 1 disk clusters that lie outside the bulge-dominated B01. The outcomes of four different cluster dissolution models are compared: no disruption, two models of mass-dependent disruption as parameterized by Lamers et al. (2005) for two dissolution timescales (t_4; characteristic destruction timescale for a 10^4 $M_\odot$ cluster), and mass-independent disruption in which 90% of clusters are destroyed within every logarithmic age interval. Models are normalized to the data at $M_{F475W} = -5$. Although the samples appear to prefer particular dissolution models, the effects that complex underlying cluster age and mass distributions impart on the luminosity functions do not allow us to draw any conclusions about cluster disruption at this time.

disruption (Larsen 2009).

To assess the shape of the PHAT Year 1 luminosity functions in each of the six filter passbands, we consider two different samples of objects for fitting: all 601 clusters, and a subsample of 534 disk clusters that excludes objects from the bulge-dominated Brick 1. The cluster population within Brick 1 is dominated by old, massive globular clusters associated with the galaxy's bulge component, in addition to the significantly brighter completeness limits in this region. The disk sample represents a set of predominately younger objects that are likely more homogeneous in terms of formation and destruction history, simplifying the interpretation of its resulting luminosity functions. We narrow the samples further to exclude clusters with uncertain photometry ($\sigma > 0.5$ mag), performing the quality cut individually for each passband. The fraction of objects in each filter that meet this quality requirement is listed in Table 2.4. Finally, we convert aperture corrected, total magnitudes for the selected clusters to absolute magnitudes, correcting for the M31 distance modulus and Galactic foreground reddening. We note that no correction for dust attenuation within M31 has been applied; we will derive these correction factors on an object-by-object basis as part of age and mass fitting analysis (Fouesneau et al. 2014).

To characterize the power law slope of the luminosity function, we perform a linear fit to constrain the slope of $\log \mathrm{d}N/\mathrm{d}M$ (the logarithm of the number of clusters per magnitude) as a function of absolute magnitude. For each passband, we bin the clusters using variable-sized magnitude bins to ensure fair weighting of the data, following the suggestion of D'Agostino & Stephens (1986) (and more recently in Maíz Apellániz & Úbeda 2005; Haas et al. 2008) to group the data such that each bin represents an equal number of clusters, N. We choose $N = 15$, but find the results are insensitive to the particular number chosen. We fit to datapoints brighter than absolute magnitude completeness limits: -4.0 for F275W, F336W, and F475W; -5.0 for F814W; -6.0 for F110W and F160W. These limits are conservative; they are equivalent to >80% completeness for all bricks, as found in Section 2.3.2. We convert the resulting magnitude-based slopes to equivalent luminosity function slopes (α_L) and report these as our primary results. The luminosity functions and their associated power-law fits for the complete and disk-only cluster samples are presented in Figs. 2.14 and 2.15; the values of the fitted slopes are listed in Table 2.5. We separate the results

Table 2.5. Luminosity Function Fits

Passband	α_L(All Clusters)	α_L(Disk Clusters)
F275W	-1.72 ± 0.07	-1.76 ± 0.08
F336W	-1.73 ± 0.06	-1.90 ± 0.08
F475W	-1.79 ± 0.06	-2.06 ± 0.09
F814W	-1.74 ± 0.05	-2.18 ± 0.09
F110W	-1.68 ± 0.06	-2.20 ± 0.11
F160W	-1.56 ± 0.05	-1.82 ± 0.08

into two groups to isolate the F275W, F110W, and F160W passbands, which carry the potential for larger systematic uncertainties due to the smaller fraction (∼50-75%) of clusters with well-determined photometry available at these wavelengths. As such, we will focus our subsequent discussion on results from the F336W, F475W, and F814W passbands in Fig. 2.14.

Considering the complete sample of Year 1 clusters, we find luminosity function power law slopes in the F336W, F475W, and F814W passbands that are all flatter than −2. While these measurements agree with the general trend of flatter slopes measured at fainter luminosities (e.g., in NGC45 and M51; Mora et al. 2007; Haas et al. 2008), we suggest these flat slopes result from the inclusion of a relatively large number of luminous, evolved globular clusters associated with Brick 1 and included in the complete cluster sample. We find that although clusters in Brick 1 make up ∼15% of the complete sample by number, ∼50% of sample members with $m_{F475W} < 19$ ($M_{F475W} \lesssim -5.5$) are Brick 1 clusters with spectroscopically derived ages of >10 Gyr (Caldwell et al. 2009, 2011). Globular clusters are known to follow a Gaussian-like luminosity function shape with a peak magnitude around M_V of −7 and −8, depending on factors such as metallicity and age (e.g., Harris 1991; Barmby et al. 2001). The dissimilar luminosity function shape of the globular cluster subpopulation

significantly influences the slope of the overall complete Year 1 luminosity function.

To determine the luminosity function shape that is characteristic of the dominant, young cluster population found in M31, we examine the fitting results for the disk-only sample of clusters, presented in the bottom row of Fig. 2.14. We find uniformly steeper slopes ($\alpha_L < -1.9$) for the F336W, F475W, and F814W disk-only luminosity functions when compared to the complete Year 1 sample results. In addition to the closer agreement between these slopes and the canonical -2 power law, we also observe a significant trend where the luminosity functions at bluer wavelengths are flatter than at redder wavelengths. This behavior can be explained by wavelength and age-dependent cluster mass-to-light ratios (Gieles 2010).

To visualize the sensitivity of the faint end luminosity function slope to differences in cluster dissolution, we compare our F475W luminosity functions to shapes predicted by four canonical dissolution models in Fig. 2.16. The analytical luminosity functions were calculated by Larsen (2009) under the assumption of a constant cluster formation history, an underlying Schechter mass function (with cutoff of $2{\times}10^5$ $M_\odot$), and continuous, non-stochastic sampling of the stellar initial mass function[5]. In the case of mass-dependent cluster dissolution, Larsen (2009) finds that the slope of the luminosity function should flatten to > -2 at $M_V > -8$ or -9, while scenarios with mass-independent dissolution (and the case of zero dissolution) result in slopes that always remain steeper than -2.

The complete Year 1 cluster sample appears to be better described by a mass-dependent disruption model, while the steeper disk cluster luminosity function appears to agree with the shape of the mass-independent or the no disruption case. However, due to the degenerate effects of underlying cluster age and mass distribution differences and variations in cluster disruption, such simple interpretation is not possible. We opt to defer further analysis and interpretation until we obtain robust age, mass, and extinction determinations for individual clusters. This crucial information will allow for the separation of the degenerate effects that influence the shape of the luminosity function and better constrain the characteristics of

[5]We note that in terms of luminosity function modeling, we find little evidence for any noticeable differences when stochastic sampling of a cluster's stellar initial mass function is incorporated, validating the present model comparison.

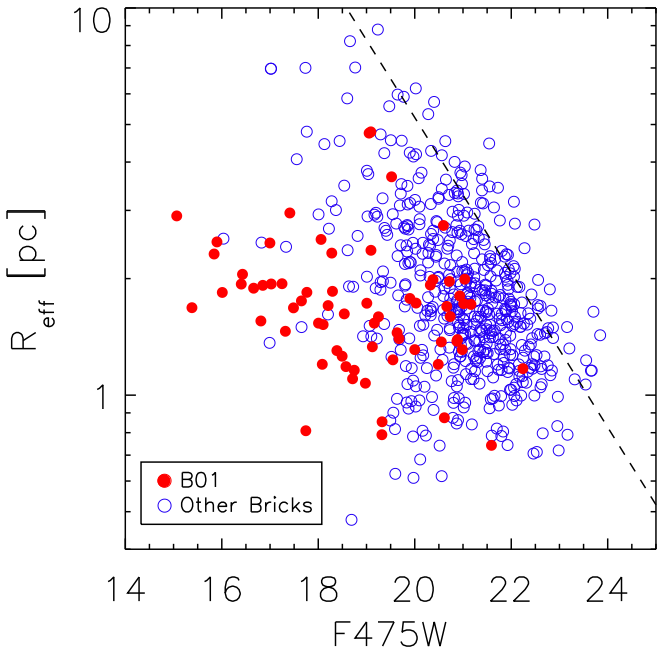

Figure 2.17 R_{eff} as a function of F475W magnitude. The old, globular cluster dominated population in B01 (red solid circles) occupy a different region of parameter space than the younger disk clusters that dominate the remaining outer bricks (blue open circles). The black dashed line denotes an approximate 50% completeness limit, corresponding to a uniform surface brightness limit consistent with a F475W cutoff of 21.2 at an R_{eff} of 3 pc.

the underlying cluster population.

2.7.2 Cluster Sizes

Cluster size is a fundamental parameter that provides constraints on the dynamical state of stellar systems. While many interesting correlations involving cluster sizes require the determination of other cluster parameters (e.g., mass-radius and age-radius relations), from sizes alone we can compare the overall properties of the cluster sample to those in other galaxies and look for interesting sub-populations of clusters within the sample.

We plot R_{eff} as a function of F475W magnitude in Fig. 2.17. In this parameter space, the sample separates well into two groups: Brick 1 clusters that are predominately old (∼10

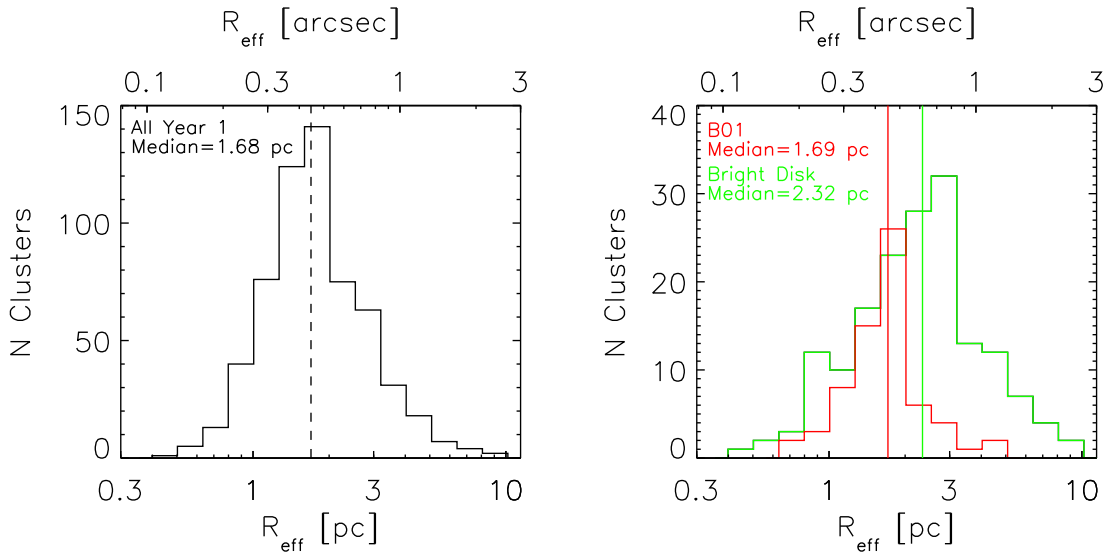

Figure 2.18 Histograms of R_{eff} derived from F475W data. Left: Distribution of R_{eff} for the full Year 1 cluster sample. Median R_{eff} is 1.68 pc (0.44″). Right: Distribution of R_{eff} for the B01 (red) and bright disk (green) subsamples. Median R_{eff} for the two subsamples are 1.69 pc (0.44″) and 2.32 pc (0.61″), respectively.

Gyr), massive globular clusters associated with the galaxy bulge as discussed in Section 2.7.1, and clusters from the remaining bricks that are associated with Andromeda's star forming disk. Overall, the Brick 1 clusters are more compact and luminous than the disk clusters, and show a trend in which R_{eff} decreases with decreasing F475W luminosity. Along with the cluster data, we plot a line of constant surface brightness anchored at F475W of 21.2 and a R_{eff} of 3 pc, representing the 50% catalog completeness limit as determined in Section 2.3.2. The distribution of clusters with respect to this line suggests that, to first order, catalog completeness roughly follows a constant surface brightness cutoff.

The disk cluster sample includes objects with a broad range of ages and suffers from significant completeness truncations, making its magnitude-radius distribution difficult to interpret. On the other hand, the Brick 1 clusters are similarly aged and are generally much brighter than the completeness limit. This allows us to conclude that the moderate trend in which fainter Brick 1 clusters are more compact is real. We do not currently know the physical origin of this trend, but plan to follow up using ages and metallicities from Caldwell

et al. (2011) and masses from Strader et al. (2011).

The left panel of Fig. 2.18 presents the principal result of our size analysis, which shows that the overall size distribution of the Year 1 PHAT cluster sample can be described approximately as a log-normal distribution, where the median R_{eff} value is 1.68 pc (0.44″). This median R_{eff} is similar to, but smaller than values found in other galaxies. In M83, Bastian et al. (2011) find a overall median value of ∼2.4 pc for a sample of mostly young clusters, while McLaughlin & van der Marel (2005) find a median value of ∼3.3 pc for a sample of Milky Way globular clusters. Additionally, Barmby et al. (2007) find a median value of ∼2.5 pc for a sample of previously known M31 globular clusters.

We can explain the differences in median cluster sizes by considering how the composition of the Year 1 PHAT cluster sample differs from the M83, Milky Way, and existing M31 samples. One possible reason for the median mismatch with the Milky Way and M31 globular cluster samples is that the Year 1 PHAT sample is a heterogeneous mix of mainly young clusters, while the globular samples are uniformly much older. However, an age-based explanation is ruled out because the old, globular-dominated Brick 1 subsample has a similar, but still smaller median R_{eff} of 1.69 pc. The R_{eff} distribution for the Brick 1 subsample is shown in the right panel of Fig. 2.18.

An explanation for the median differences becomes clear when we consider the galactocentric radii of the existing globular cluster samples. A correlation between R_{eff} and galactocentric radius has been show to exist in many galaxies such that inner clusters are more compact (Barmby et al. 2007). When we limit the Milky Way and previous M31 cluster samples to include only objects with galactocentric radii less than 3 kpc, which roughly matches the radial extend of the Brick 1 sample, the median R_{eff} for these two samples drops to ∼2.0 pc. These lower medians are in better agreement with the Brick 1 subsample, easing the tension between the values derived for the two globular cluster samples and for the PHAT clusters.

To explain the median R_{eff} difference between the young M83 cluster sample and the Year 1 PHAT sample, instead of age and galactic position, we consider how completeness limits influence the shape of the R_{eff} distribution. In terms of numbers, the Year 1 cluster sample is dominated by compact, low luminosity objects that lie just above our detection

threshold. At these faint magnitudes, surface brightness limits prevent us from detecting clusters across the full range of possible R_{eff} values. As a result, the overall cluster sample is biased towards compact objects, driving the median towards smaller values. However, if we only consider clusters with F475W < 20.5, the biasing effects of surface brightness limits are greatly reduced. The R_{eff} distribution for this luminous subsample is shown in the right panel of Fig. 2.18. Without the faint, compact portion of the sample, the median cluster size rises to 2.32 pc, which agrees well with the observed value obtained for clusters in M83. In conclusion, cluster sizes for the Year 1 sample are consistent with results derived by previous works in similar spiral galaxies once sample completeness and composition are taken into account.

2.7.3 Objects of Interest

2.7.3.1 Red Supergiant Clusters

We identified a subsample of objects with unusually red F475W-F814W colors in Fig. 2.12. Within the selection box, we identify ∼15 clusters that show extreme integrated colors. Upon inspection, we find that many of these clusters appear to host luminous red supergiant (RSG) stars, which explain their anomalous integrated colors. Two examples of this class of cluster, PC57 and PC1127, are shown in Fig. 2.19.

Red supergiant stars are massive (8 to 25 $M_{\odot}$) stars that have left the main sequence and are traversing the top of the CMD during the helium burning phase of stellar evolution. Individual luminous RSG stars have the ability to bias the integrated colors of low mass ($< 10^4$ $M_{\odot}$) clusters because stochastic sampling of the stellar mass function in low mass systems can cause the small number of evolved stars to vary by factors of a few (low mass clusters typically host between 0 and 4 evolved stars). This variation creates wild fluctuations in the total luminosity and integrated color of the cluster, which depend on the particular number of evolved stars progressing through this relatively short evolutionary phase.

The existence of these color outliers confirms that stellar population modeling of low mass clusters must account for stochastic sampling of the stellar initial mass function to

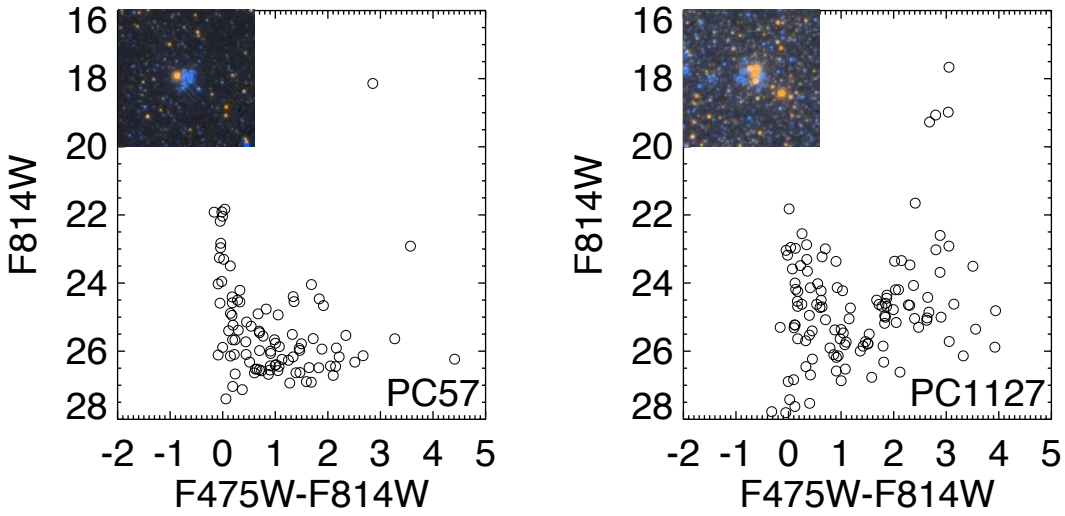

Figure 2.19 Resolved star color-magnitude diagrams and color optical (F475W+F814W) image cutouts for two candidate red supergiant clusters. These clusters have unusually red F475W-F814W colors due to the presence of luminous evolved stars that strongly bias integrated light measurements.

obtain accurate age and mass determinations (see e.g., Fouesneau & Lançon 2010). Models that assume a fully sampled stellar initial mass function cannot reproduce objects with integrated colors as red as those shown by RSG clusters. Our future cluster analysis will benefit from the use of cluster models that account for stochastic fluctuations in integrated light caused by the RSGs.

More importantly, our ability to resolve individual stars in clusters provides us a number of important benefits with respect to the RSG clusters. Resolved star photometry enables us to use CMD fitting techniques to obtain cluster parameter determinations, allowing us to avoid the use of biased integrated measurements all together. In addition, once we obtain age, mass, and attenuation information for the clusters, we can use this information to tag individual RSGs and use these constraints to improve calibration of late stage stellar evolution for massive stars at high metallicity.

2.7.3.2 Massive Clusters

Globular clusters, with old characteristic ages (∼10 Gyr) and large masses ($> 10^5\ M_\odot$), have long been a target of study in M31. Although the Year 1 cluster sample is numerically dominated by low mass clusters, it contains many massive clusters as well. While we identified few new globular clusters as part of the Year 1 search, our high spatial resolution imaging has enabled us to confirm or eliminate a large number of globular cluster candidates from existing catalogs. In addition to the 63 out of 71 highly-ranked clusters we confirm from the RBC, we affirm 12 and reject 19 possible candidates. The PHAT survey also enables detailed analysis of these objects by means of resolved star photometry for cluster members. We plan to place better constraints on red giant branch (RGB) and horizontal branch (HB) morphology as a function of metallicity through these detailed studies.

In addition to the old globular clusters, we have also newly identified a number of intermediate mass ($\sim 10^4\ M_\odot$), intermediate age (∼1 to 3 Gyr) clusters. These objects are interesting targets for study because of the relative rarity in the Milky Way (Friel 1995), where most similarly aged Galactic clusters are less massive. While a small number of these intermediate mass, intermediate age clusters are already known to exist in M31 (Caldwell et al. 2009, 2011), increasing the sample size for this class of object should help to better understand their origin and evolution. Investigation of these objects will complement the study of younger (<1 Gyr) massive clusters by Fusi Pecci et al. (2005), Caldwell et al. (2009), and Perina et al. (2010).

The most remarkable massive cluster in the Year 1 catalog is the cluster SK142C (PC1156). This object is the sample's most discrepant color outlier, appearing at the redward extreme of all cluster photometric measurements, with a F475W-F814W color of ∼3.5. This cluster appears to be an intermediate mass, intermediate age, highly-reddened cluster. Along with cutout images showing the cluster's heavily reddened appearance, we plot the cluster's resolved star CMD in Fig. 2.20. The CMD shows an extended, highly-reddened RGB, indicative of an intermediate cluster age. We perform a visual comparison to solar metallicity Padova isochrones (Girardi et al. 2010) and find that the CMD is consistent with an age of ∼1–3 Gyr and A_V of 2.25–2.75 mag, assuming large uncertainties associated

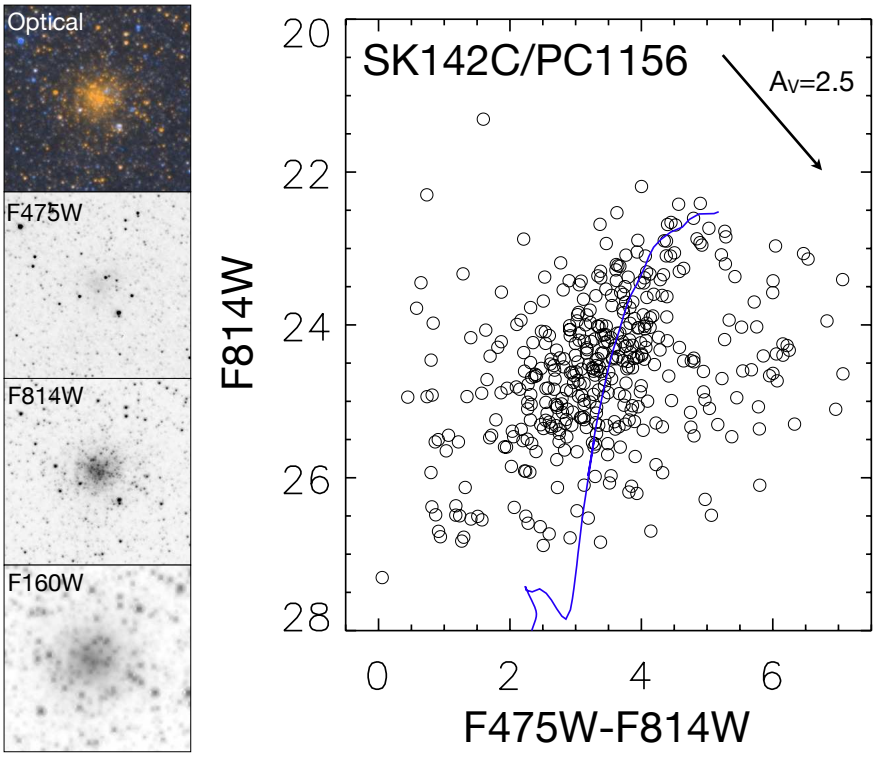

Figure 2.20 Left: Image cutouts showing highly-reddened, intermediate age cluster SK142C (PC1156). Right: The resolved star color-magnitude diagram for SK142C shows a reddened RGB, indicative of an intermediate age population. We overplot a Padova stellar isochrone with a log(Age) of 9.35 (∼2 Gyr), A_V of 2.5 mag, and solar metallicity to show that the cluster is consistent with these parameter values.

with visual "chi-by-eye" isochrone fitting. Using this combination of age and reddening, we use the cluster's F160W luminosity to estimate a cluster mass of $\sim 1.5 \times 10^5\ M_\odot$. Due to its massive nature and its somewhat intermediate age, this cluster is a rare and interesting object akin to newly discovered objects in the Milky Way (Figer et al. 2006; Davies et al. 2011, 2012).

2.8 Summary & Future Work

In this paper, we presented the first installment of the PHAT stellar cluster catalog. We introduced the Year 1 cluster sample, consisting of 601 clusters identified through a visual search of high spatial resolution HST imaging. The PHAT cluster sample significantly increases the number of clusters known in M31; the catalog presented here represents more than a factor of four increase in the number of known clusters within the current survey area. We presented a basic assessment of the cluster sample, including positional, size, and photometric information.

We have shown that the PHAT cluster sample hosts a large range in ages and masses. This wide-ranging sample provides the opportunity for a top-to-bottom study of stellar cluster evolution processes. The PHAT sample includes low-luminosity objects missed in studies of distant external galaxies, while covering an uninterrupted range of cluster masses, unlike Milky Way cluster samples.

When the survey is complete, the PHAT cluster catalog will be among the largest and most comprehensive surveys of star clusters in any galaxy. This work presents results derived from the first 25% of the survey data; we estimate that the final sample will include ∼2500 clusters. Over the duration of the survey, we plan to periodically publish updates to the catalog to include new clusters and to revise object classifications as we gather additional information about the sample. Age and mass determination analysis for the Year 1 sample is currently underway (Fouesneau et al. 2014), which will provide the means to explore a host of different topics, including the cluster mass function and cluster dissolution. The Year 1 cluster sample will also be the basis for analysis of structural parameters, resolved star content, and integrated spectroscopy, but we also look forward to future studies enabled by the complete, four year PHAT cluster catalog.

Table 2.6. Background Galaxy Catalog

ID	Brick	RA (J2000)	Dec (J2000)	R (arcsec)
1	21	11.677181	42.192120	1.41
2	21	11.644169	42.197941	1.20
3	21	11.647937	42.202160	1.37
4	21	11.685157	42.218735	0.97
5	21	11.616282	42.196312	2.05

Note. — Table 2.6 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

2.9 Supplementary Information

2.9.1 Background Galaxy Catalog

As introduced in Section 2.3.1, we present a catalog of objects visually identified as background galaxies in Table 2.6. Along with positional information, we present a rough assessment of the galaxy size.

2.9.2 Commentary on Existing Cluster Catalogs

We performed a comparison of object classifications from existing cluster catalogs to those in the PHAT Year 1 cluster catalog in Section 2.5. Here we provide comments on individual catalog comparisons, including detailed classification statistics and object-by-object classification revisions. We discuss individual classification statistics for the RBC (Galleti et al.

Table 2.7. Summary of Existing Cluster Catalog Classifications and Revisions

Catalog	Clusters		Candidates		Galaxies		HII Regions		Stars		Other[a]		Total
RBC	78	(71)	4	(34)	3	(3)	1	(3)	98	(76)	3	(0)	187
Kim (Total)	12	(10)	2	(28)	2	(0)	0	(0)	22	(0)	0	(0)	38
Kim A	4	(10)	1	(0)	2	(0)	0	(0)	3	(0)	0	(0)	10
Kim B	4	(0)	0	(15)	0	(0)	0	(0)	11	(0)	0	(0)	15
Kim C	4	(0)	1	(13)	0	(0)	0	(0)	8	(0)	0	(0)	13
Caldwell	68	(73)	1	(0)	0	(0)	1	(0)	113	(110)	0	(0)	183
Peacock	77	(58)	4	(25)	2	(2)	0	(2)	41	(40)	3	(0)	127
HKC	54	(75)	8	(0)	0	(0)	0	(0)	0	(0)	13	(0)	75

Note. — In each column, the first number represents the number of objects per category after PHAT reclassification and the second number (in parenthesizes) represents the number of original classifications from the published catalog. The A classification from Kim et al. (2007) maps to a cluster classification in the RBC, while B and C map to candidate classifications.

[a]The "other" classification signifies clusters that were not recovered by the PHAT search or were duplicate catalog entries in the RBC and Peacock et al. (2010) catalogs. In the case of the HKC, the "other" classification signifies objects that were deemed non-cluster asterisms.

2004), Kim et al. (2007), Caldwell et al. (2009, 2011), Peacock et al. (2010), and three of the four HKC studies (Krienke & Hodge 2007; Hodge et al. 2009, 2010; no objects from Krienke & Hodge 2008 lie within the Year 1 PHAT footprint). Classification results broken down by object type are presented in Table 2.7.

The Revised Bologna Catalog — Our search for RBC objects that lie within the Year 1 PHAT imaging footprint produces a catalog of 187 objects. These objects include 71 clusters, 36 cluster candidates, and 80 non-cluster objects. After reanalyzing these classifications using the by-eye ranking results and following up on other objects that were not selected as PHAT candidates, our revised object classifications show excellent agreement with the original RBC results. Individual object revisions are provided in Table 2.8, but the bulk of the changes are due to our ability to determine the nature of the RBC's candidate

objects (RBC Flag = 2). Of the 36 candidates, we confirm 12 objects as clusters in the PHAT catalog, reject 23 candidates (including two originally controversial cases, one duplicate entry, and one object not found in PHAT imaging), and transfer the final candidate to the PHAT possible cluster sample. In terms of catalog accuracy, only five "confirmed" clusters (RBC Flag = 1) were confidently revised to non-cluster objects, four of which were Kim et al. (2007) objects (see additional discussion below). Overall, the high level of consistency between the two catalogs gives us great confidence in galaxy-wide accuracy of the RBC cluster sample.

Kim et al. (2007) — The catalog content of Kim et al. (2007) is also accounted for in the results of the RBC cross-comparison, but the resulting cluster confirmation and rejection statistics from this work is worth individual mention. As pointed out in Caldwell et al. (2009) and Peacock et al. (2010), the Kim et al. (2007) catalog has a high level of contamination, namely from misclassified stellar sources. Considering all three cluster quality categories from Kim et al. (2007) together, we find 38 objects from this catalog that fall within the Year 1 PHAT footprint. Of these, only 37% of those objects appear in the PHAT catalog as confirmed or possible clusters. Even when taking the quality rankings into consideration, we find that (50%, 27%, 38%) of the (A, B, C) objects appear in the catalogs from our present work.

Caldwell et al. (2009, 2011) — These two papers revised classifications from v3.4 of the RBC (many of which were adopted in v4.0) using low-resolution spectra, ground-based imaging from the Local Group Galaxy Survey (LGGS; Massey et al. 2006), and HST-based images. As this catalog has been incorporated into the latest version of the RBC, discussion of individual object classifications are included in Table 2.8. When comparing our current PHAT classification with those from Caldwell et al. (2009, 2011), we find good agreement between the two works. We revise classifications for 6 clusters out of 183 that fall within the Year 1 footprint.

Peacock et al. (2010) — The clusters considered by Peacock et al. (2010) are included in the RBC comparison discussed above, and classifications for each of these objects is already provided in Table 2.8. Many of the reclassifications performed in Peacock et al. (2010), which was based on v3.5 of the RBC, have proven to be accurate (including the stellar

classification of numerous Kim et al. 2007 candidates). Overall, we find good consistency between the Peacock et al. (2010) catalog and our PHAT classifications.

Hodge-Krienke Catalogs — Similar data and methodology (including common authors in the case of P.H.) leads to strong consistency between the HKC and the work presented here. We find 54 clusters from the HKC that match entries in the PHAT cluster catalog and 8 additional objects that match possible clusters, out of a total of 75 objects that lie within the boundaries of the PHAT Year 1 footprint. Unlike in the case of ground-based catalogs, the 13 HKC objects not recovered as part of the PHAT cluster search mostly fall into the category of asterisms. In other words, these objects were interpreted by the PHAT search team as uncertain or unlikely cluster candidates, and omitted from the final catalog presented in this work. This subset of objects shows that for the least luminous, least massive clusters, defining the difference between an object that is potentially a bound stellar system and simply a chance collection of stars seen in projection becomes a difficult and subjective task. Table 2.9 presents proposed revisions to the HKC. In closing, we note that we do not compare to cluster identifications presented in Williams & Hodge (2001a) because classifications in this work were superseded by the HKC.

2.9.3 Comparison to Existing Cluster Photometry

We provide a comparison between PHAT cluster photometry and the results of Peacock et al. (2010), Fan et al. (2010), Hodge et al. (2009), and Hodge et al. (2010). We compare photometry for these selected studies because they provide magnitudes derived in similar photometric passbands as part of a uniform analysis. We choose not to compare to photometry provided by the Revised Bologna Catalog due to the heterogenous nature of their unified photometric measurements (see Galleti et al. 2004, for details). We note that both Peacock et al. (2010) and Fan et al. (2010) find good agreement between photometric results and that of Barmby et al. (2000), upon which the photometry of the Revised Bologna Catalog is based.

The optical aperture photometry of Peacock et al. (2010) was derived from SDSS *ugriz* imaging. We compare photometric results for 77 clusters found in both datasets. We

Table 2.8. Revised Bologna Catalog Revisions

Cluster Name	PC ID	New Flag	Old Flag	Comments
BH16	1381	1	2	···
B523	1383	1	2	···
SK118C	641	1	2	···
SK134C	1349	1	2	···
M028	544	1	2	···

Note. — Table 2.8 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

Table 2.9. Hodge-Krienke Catalog Revisions

Cluster Name	PC ID	PHAT Classification
WH13	1119	Possible Cluster
WH18	1089	Possible Cluster
KHM31-195	1282	Possible Cluster
KHM31-241	964	Possible Cluster
Hodge09-57	977	Possible Cluster

Note. — Table 2.9 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

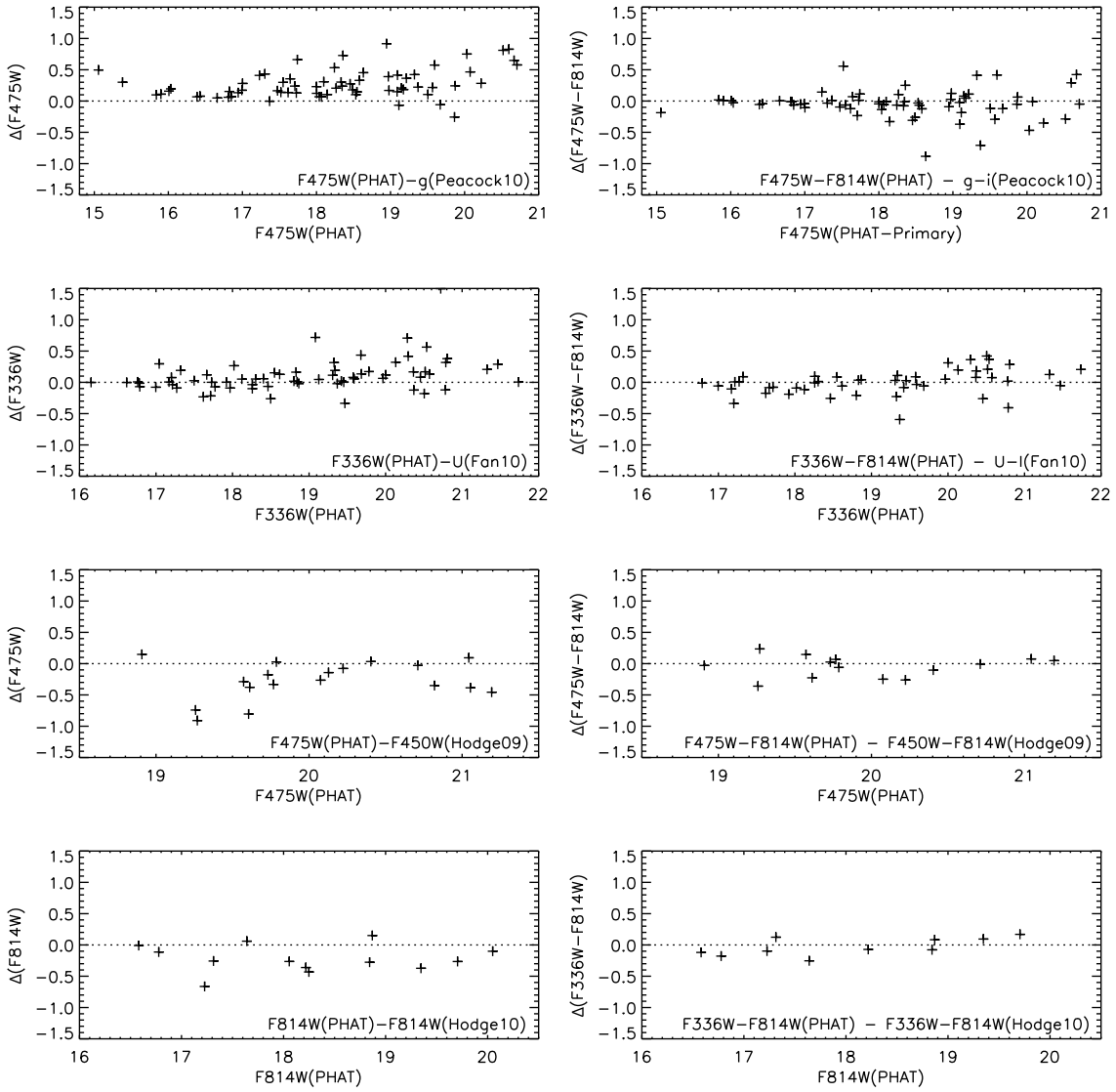

Figure 2.21 Comparison of PHAT photometric results with those of Peacock et al. (2010) (top row), Fan et al. (2010) (second row), Hodge et al. (2009) (third row), and Hodge et al. (2010) (bottom row).

calculate transformations to convert from AB-based g and i-band photometry to Vega-based F475W and F814W magnitudes using transformations from Girardi et al. (2008), assuming a single-age, 1 Gyr solar metallicity population that represents a median value for the age-dependent transformation:

$$F475W = g_{AB} + (0.06 \pm 0.02), \tag{2.1}$$

$$F814W = i_{AB} - (0.54 \pm 0.10). \tag{2.2}$$

The resulting photometric comparison is presented in the top row of Fig. 2.21. We find that the PHAT magnitudes are fainter than those of Peacock et al. (2010) by $\sim$0.25 mag. This offset can be explained by the difference in aperture sizes, as seen in Fig. 2.22. The Peacock et al. (2010) apertures have radii that are twice as large on average, and up to four times as large as those used in this work. In the case of the brightest comparison cluster, B127-G185 (PC1425), the size of the aperture ($10''$) was large enough to include the nearby cluster NB89 (PC1426), explaining the $\sim$0.5 mag difference for this object. In a comparison of F475W-F814W versus $(g-i)$ optical colors, the agreement is quite good for brighter clusters (F475W $<$17) with increasing scatter and a redward bias of Peacock et al. (2010) colors for fainter clusters. These differences are likely caused by field contamination within the large Peacock et al. (2010) apertures, and by blending in their low-resolution (1-2$''$ seeing) ground-based imaging. In summary, we find systematic differences between the photometry derived in this work and that of Peacock et al. (2010), but the overall agreement is adequate and remaining differences can be readily explained by the effects of image resolution and aperture size differences.

In another recent compilation of ground-based aperture photometry, Fan et al. (2010) derived Johnson-Kron-Cousins $UBVRI$ magnitudes from LGGS imaging (Massey et al. 2006). The Vega-based PHAT F336W and F814W magnitudes are nearly equivalent to the U and I magnitudes from Fan et al. (2010), allowing us to directly compare photometric results for 68 clusters common to both samples. We present the results in the second row of Fig. 2.21, showing good agreement between the two datasets. While Fan et al. (2010) magnitudes appear slightly brighter for faint clusters, the overall agreement is improved with respect to the Peacock et al. (2010) results. One explanation for the improvement between

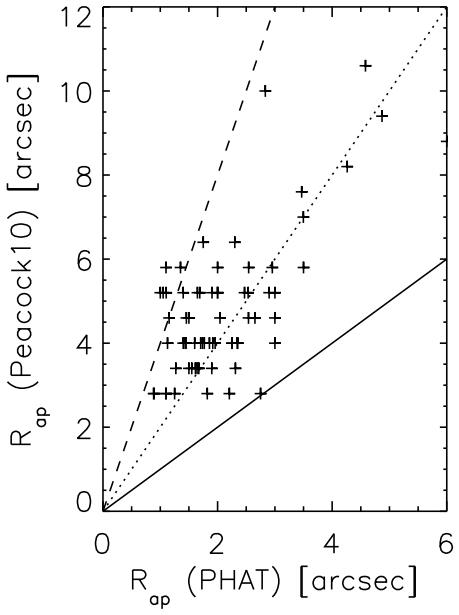

Figure 2.22 Comparison of photometric aperture radii used in Peacock et al. (2010) and in this work. The solid line represents a 1-to-1 relation, the dotted line a 2-to-1 relation, and the dashed line a 4-to-1 relation.

ground-based and HST-based results could trace back to the superior image quality of the LGGS data when compared to the SDSS imaging of Peacock et al. (2010). The average seeing for the LGGS data is $\sim 1''$, while varying between 1-2$''$ for the SDSS imaging. This results in the ability for Fan et al. (2010) to reduce their aperture sizes, such that they are only a factor of $\sim$1.5 bigger than those used by PHAT.

The photometry of Hodge et al. (2009) and Hodge et al. (2010) allows us to compare two sets of HST-based photometric results. Out of the four M31 cluster catalog papers that compose the HKC, we choose two where the tabulated photometry is provided in native, Vega-based HST filter magnitudes. Although the data used in these studies are similar to those used in this work, there are fundamental differences in the photometry techniques employed. As opposed to our growth curve method of aperture size determination, the HKC adopt an isophotal aperture determination, such that apertures extend out to a chosen surface brightness limit. Further, the methods of sky background determination differ. Our sky levels are determined using local sky background estimates measured in annular rings around the photometric aperture, while the HKC uses samplings of the sky background taken across the HST image. Due to these differences in methodology and the lack of any correction to account for total cluster light, we expect the comparison to show some scatter and a systematic offset such that the PHAT photometry is brighter.

For the Hodge et al. (2009) comparison, we use 18 objects common to both datasets and present the results in the third row of Fig. 2.21. We compare our F475W photometry to magnitudes measured in a similar F450W passband, using a transformation derived in the same way as Eqs. 2.1 and 2.2:

$$F475W = F450W - (0.11 \pm 0.10). \tag{2.3}$$

There are several significant (Δ >0.5 mag) outliers in the F475W magnitude comparison, in which the photometry from this work is brighter by up to $\sim$1 mag. To check the validity of our measurements, we performed photometry for the three most discrepant outliers on the WFPC2 images used by Hodge et al. (2009), employing photometry procedures used in this work. We found resulting magnitudes that agree within $\sim$0.1 mag of our original PHAT photometry values. This result indicates that the magnitude offsets are likely due to

differences in photometric technique, specifically related to sky background determination.

The bottom row of Fig. 2.21 presents the comparison of Hodge et al. (2010) photometry for 13 common objects. We find better overall agreement with this dataset when compared to the Hodge et al. (2009) results. We compare F814W magnitudes (no filter transformations required) and F336W-F814W colors, and find a scatter of $\sim$0.2 mag between the two datasets. The comparison of photometry between our current PHAT study and the work of Hodge et al. (2009, 2010) shows good overall consistency. Systematic biases and a small number of significant outliers likely stem from differences in measurement technique.

Chapter 3

ANDROMEDA PROJECT CLUSTER CATALOG

This chapter was published as L.C. Johnson et al. 2015, ApJ, 802, 127, and is reproduced by permission of the AAS.

We construct a stellar cluster catalog for the Panchromatic Hubble Andromeda Treasury (PHAT) survey using image classifications collected from the Andromeda Project citizen science website. We identify 2,753 clusters and 2,270 background galaxies within ∼0.5 deg^2 of PHAT imaging searched, or ∼400 kpc^2 in deprojected area at the distance of the Andromeda galaxy (M31). These identifications result from 1.82 million classifications of ∼20,000 individual images (totaling ∼7 gigapixels) by tens of thousands of volunteers. We show that our crowd-sourced approach, which collects >80 classifications per image, provides a robust, repeatable method of cluster identification. The high spatial resolution Hubble Space Telescope images resolve individual stars in each cluster and are instrumental in the factor of ∼6 increase in the number of clusters known within the survey footprint. We measure integrated photometry in six filter passbands, ranging from the near-UV to the near-IR. PHAT clusters span a range of ∼8 magnitudes in F475W (g-band) luminosity, equivalent to ∼4 decades in cluster mass. We perform catalog completeness analysis using >3000 synthetic cluster simulations to determine robust detection limits and demonstrate that the catalog is 50% complete down to ∼500 $M_\odot$ for ages <100 Myr. We include catalogs of clusters, background galaxies, remaining unselected candidates, and synthetic cluster simulations, making all information publicly available to the community. The catalog published here serves as the definitive base data product for PHAT cluster science, providing a census of star clusters in an $L^\star$ spiral galaxy with unmatched sensitivity and quality.

3.1 Introduction

Observations of our Local Group neighbor, M31, present the best opportunity for a detailed yet comprehensive study of a large spiral galaxy, providing a local analog to the disk-dominated systems that populate wide-field galaxy surveys. While the Milky Way allows analysis at the highest level of detail, studying our host galaxy on the whole proves difficult due to distance ambiguities and large amounts of dust attenuation within the Galactic plane. Conversely, studying galaxies beyond the local group necessitates a substantial decrease in data quality and content due to reduced spatial resolution and rising photometric completeness limits.

Similarly, Andromeda is an excellent target for obtaining a big picture view of a galaxy's stellar cluster population. While many extragalactic cluster samples exist, each offering galaxy-wide coverage unattainable in the Milky Way, M31's proximity provides a number of sensitivity-based advantages. Using the power of the Hubble Space Telescope (HST), we can obtain a census of Andromeda's star cluster population that extends deep into the low-mass regime while simultaneously resolving individual stars within each cluster. The ability to resolve individual stars also allows for thorough analysis of M31's field star populations, leading to detailed comparisons of field and cluster populations, enabling studies of cluster formation and dissolution in the context of the galaxy's overall star formation activity.

The Panchromatic Hubble Andromeda Treasury survey (PHAT; Dalcanton et al. 2012) provides contiguous, high spatial resolution imaging of approximately one-third of the M31 disk using the HST, observed in six broadband passbands that span from the near-UV to the near-IR. The Year 1 cluster catalog (Johnson et al. 2012, hereafter, Paper I) presented cluster results from the first 20% of the survey data. In this paper, we present a final, survey-wide cluster catalog created through a crowd-sourced, visual search of the data. The contribution of citizen scientists to astronomical research is not novel: projects such as Galaxy Zoo (Lintott et al. 2008; Willett et al. 2013), the Milky Way Project (Simpson et al. 2012), and Planet Hunters (Schwamb et al. 2012) have previously made use of crowd-sourcing. In this work we analyze image classifications collected from the Andromeda Project, a website established explicitly for the identification of star clusters in the PHAT dataset.

We utilize these data to assemble a cluster catalog that reaches cluster masses below 10^3 $M_\odot$. This level of catalog completeness represents a significant extension to previous ground-based studies of M31, which mainly focused on old massive globular clusters, presented in the compilations of Caldwell et al. (2009, and updates via its corresponding website[1]), the Revised Bologna Catalog[2] (RBC; Galleti et al. 2004, last updated 2012 August to v5), Huxor et al. (2014), and references therein. The new catalog also builds upon previous space-based efforts in M31 by Williams & Hodge (2001a) and the series of Hodge-Krienke catalogs (Krienke & Hodge 2007, 2008, 2013; Hodge et al. 2009, 2010, hereafter collectively referred to as the HKC). The HST's high spatial resolution imaging allows for the identification of less massive clusters through its ability to differentiate between single stars and compact clusters, but previous HST-based studies were limited to isolated targeted observations. In contrast, PHAT's contiguous wide-area coverage allows us to study cluster populations across the entire northeast quadrant of M31.

The catalog presented here serves as the basis for future work that will further characterize the sample: basic cluster parameter determinations (age, mass, A_V; Beerman et al. 2012; Fouesneau et al. 2014, Beerman et al., in prep), spatial profiles (Fouesneau et al., in prep), and comparison to spectroscopically-derived properties of the globular cluster population (Caldwell et al., in prep). Once characterized, the star clusters presented here will be used as input for a variety of explorations by the PHAT collaboration and others. As part of PHAT, we will place constraints on the high-mass stellar initial mass function (Weisz et al. 2015), and measure cluster formation efficiency throughout the galactic disk (L.C. Johnson et al., in prep) to test theoretical model predictions (Kruijssen 2012). Further, we will constrain cluster dissolution time scales (M. Fouesneau et al., in prep) in an effort to differentiate between competing models (mass-dependent versus mass-independent dissolution; Fall et al. 2009; Boutloukos & Lamers 2003; Bastian et al. 2012b; Chandar et al. 2010c).

We begin with a description of the citizen science website and data in Section 3.2. Section

[1] http://www.cfa.harvard.edu/oir/eg/m31clusters/M31_Hectospec.html

[2] http://www.bo.astro.it/M31/

3.3 discusses the process of converting contributions from citizen scientists into a catalog of objects, while Section 3.4 characterizes the make-up and completeness of the final catalogs. We present our cluster catalog and accompanying integrated photometry in Section 3.5. Section 3.6 includes a comparison of the current catalog with our previous Year 1 work and a discussion of how this cluster sample fits within the context of other well-known cluster catalogs. We conclude with a summary of our work in Section 3.7. Throughout this work, we assume a distance modulus for M31 of 24.47 (785 kpc; McConnachie et al. 2005), where 1 arcsec corresponds to a physical size of 3.81 pc.

3.1.1 Cluster Definition

A star cluster can be defined in the most general sense as a grouping of stars that are spatially and temporally correlated. Beyond this broad definition, the notion of a star cluster can vary significantly, depending mostly on whether the system is still embedded in its natal gas or exposed (Lada & Lada 2003). Older (>10–30 Myr) gas-free systems are relatively straightforward to classify using a criterion based on the gravitational boundedness of individual members to a larger group. In contrast, young groupings of stars that are still embedded within the ISM make classification a difficult, uncertain task. These embedded clusters are still forming through hierarchical merging of sub-clumps (Allison et al. 2010), and the application of various stellar density thresholds to identify distinct features of a continuous (scale-free) distribution leads to interpretative challenges (Bressert et al. 2010; Gieles et al. 2012). Embedded environments are dynamically evolving and membership within a particular gravitational grouping is neither well-defined nor unique.

For the PHAT cluster catalog, we work mostly in the exposed, gas-free regime because our identification is based on optical imaging. Once the gas has been expelled from a star cluster and its stars have evolved through multiple dynamical times, it becomes possible to infer whether a grouping of stars is either gravitationally bound or expanding and dissolving (Gieles & Portegies Zwart 2011). Therefore, uncertainties pertaining to boundedness are minimal for our sample because a majority of PHAT clusters are already many dynamical times old, as inferred from the age and mass distributions of the Year 1 catalog (Fouesneau

et al. 2014).

At young ages (<10 Myr), the use of boundedness as a selection criterion for clusters becomes difficult. Due to the similar appearance (i.e., radial spatial profile) of bound clusters and unbound stellar associations at young ages, determining boundedness becomes an uncertain and contentious enterprise (e.g., see Chandar et al. 2010a; Bastian et al. 2012b; Whitmore et al. 2014). In the work that follows, we include all objects identified as part of our search. As a result, our catalog may include a heterogeneous mix of bound and unbound objects at ages <10–30 Myr. We choose this approach in an effort to maximize the return for science cases that do not depend on the differentiation between bound clusters and unbound associations, while allowing open discussion of differing cluster definitions where they affect the resulting scientific interpretation. Overall, we seek a catalog of objects that are spatially and temporally correlated and can be reasonably approximated as simple stellar populations. While this goal is easily achieved for a majority of the sample, we will make a point to identify regions of parameter space that contain debatable objects, allowing the reader to make informed decisions with regards to boundedness. A full exploration of the question of boundedness requires detailed age and spatial structure information (Gieles & Portegies Zwart 2011), which is beyond the scope of this work.

The inclusive philosophy that we adopt in this work represents a shift from the approach we took in Paper I, where we discarded objects that were classified as likely associations. This paper's inclusive methodology leads to a modest ∼15% increase in clusters when compared to the Year 1 catalog within their shared imaging footprint. We discuss these catalog differences in detail in Section 3.6.2, but find good overall agreement between the two samples.

3.2 The Andromeda Project

In Paper I, we presented a sample of 601 clusters identified in a visual search carried out by eight professional astronomers, which examined the first 20% of imaging acquired by the PHAT survey. This task was time consuming; the initial identification of cluster candidates and subsequent quality ranking of the candidates required more than a month of effort from each scientist involved. This cost limited our cluster search in two significant ways. First,

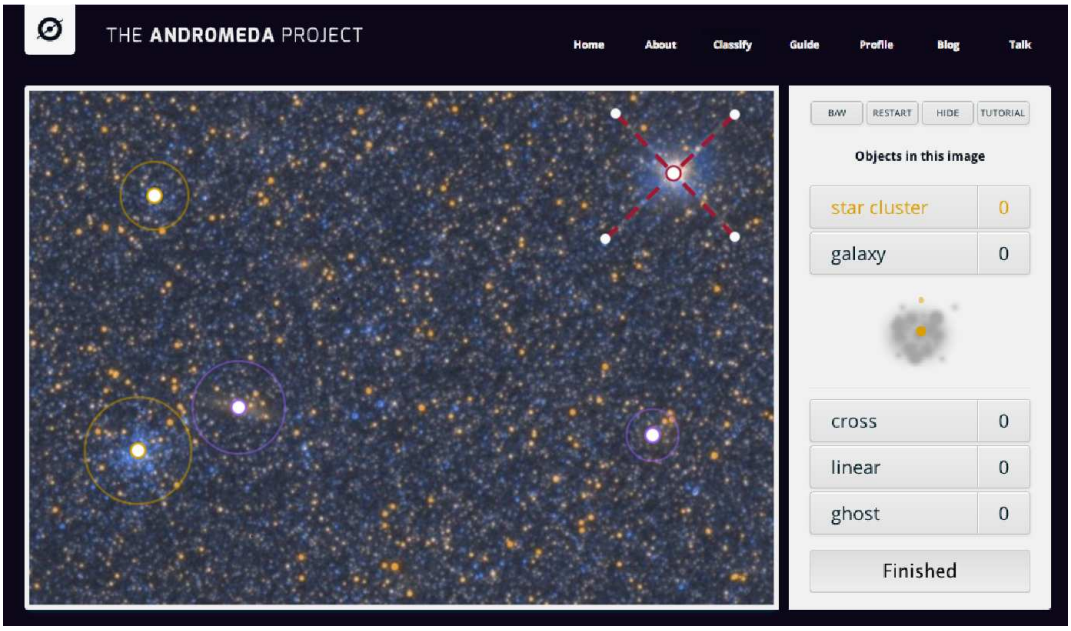

Figure 3.1 The web-based classification interface for the Andromeda Project. The tutorial image used to train participants is shown here, which includes all three object types: clusters (yellow), background galaxies (purple), and artifacts (i.e., a saturated star with diffraction spikes; red).

only 3–4 people looked at each image to make initial identifications of cluster candidates. Of the 601 clusters, 23 were originally identified by just a single person, suggesting that a small number of additional good cluster candidates were probably missed in our initial search. Second, characterization of the cluster completeness was done with a sample of 550 synthetic (artificial) clusters. This relatively small sample of synthetic clusters limited our ability to track the completeness as a function of age, mass, cluster size, and galactocentric radius.

Our original plan for extending the cluster search to the full PHAT footprint was to devise an automated algorithm to identify clusters using the Year 1 sample as a training set. This approach proved challenging because all of the automated techniques we tested produced samples with at least as many contaminants as true clusters. Expert by-eye verification would have been necessary to reduce the number of contaminants to an acceptable level. This verification would have been time consuming and the resulting catalog would

still suffer from subjectivity issues. In addition, the goal of robustly characterizing the catalog selection function becomes difficult, requiring an understanding of human and machine behavior and their joint interaction.

The failure to devise a fully automated cluster identification technique, combined with the difficulty of scaling our original by-eye techniques to the full dataset, led us to create the Andromeda Project. This crowd-sourced solution allows us to scale a by-eye search to the volume of data available from PHAT, improve the robustness and repeatability of cluster identifications, and accurately characterize the catalog completeness function.

3.2.1 Interface

The Andromeda Project[3] (AP) is a website built and hosted by the Zooniverse[4] citizen science platform. The AP interface is based on previous tools and code used for the Seafloor Explorer project, another Zooniverse project that aims to survey scallops, seastars, and other aquatic life using underwater imaging.

Upon entering the AP website, visitors are presented with the primary option to start classifying data, as well as links to find out more about the project. Individuals who start classifying for the first time are directed to a tutorial image, where the basic functionality of the classification screen is explained. The classification screen is shown in Figure 3.1. By default, the site displays a color image constructed from F475W and F814W imaging. By clicking on the "B/W" button, participants can change the image to an inverted F475W gray scale image in which it is often easier to distinguish individual stars and faint image features. The site's marking tool is set for cluster identification by default; modes for identifying background galaxies and three types of image artifacts are also available. Markers for clusters, galaxies, or ghost artifacts are circular, positioned by clicking the center of an image feature and dragging outward to select the desired radius. Only the cluster and galaxy markings are utilized in this paper.

After clicking on the "Finished" button, volunteers are shown the location of the field

[3] `http://www.andromedaproject.org`

[4] `http://www.zooniverse.org`

they were classifying within M31 and given the option to discuss the images in the AP Talk[5] forum. This feature enables new volunteers to get help identifying clusters, and allows participants to highlight interesting or confusing objects and discuss them with other volunteers and the science team. After choosing whether or not to enter the Talk forum, volunteers are presented a new search image; the AP image database ensures that no user sees the same image twice.

Volunteers are urged to log-in or create a Zooniverse account, but participants are allowed to classify an unlimited number of images as an unregistered user. Unregistered users do, however, receive periodic messages suggesting that they log-in or create an account. Registration allows analysis of volunteers' classification behavior using consistent (anonymous) identifiers. Input from unregistered users can still be aggregated from within a single classification session, however the (anonymous) identifiers tend not to carry over from session to session and could be shared by multiple unregistered users, limiting the depth of analysis we can perform.

3.2.2 Input Data & Synthetic Clusters

Each search image shown on the AP site was extracted from high-resolution (0.05 arcsec pixel^{-1}) HST/ACS images of M31. A vast majority of these images came from the PHAT dataset; we show the survey's imaging footprint in Figure 3.2. The prominent rectangular regions in Figure 3.2 that divide the survey into 23 parts are referred to as "bricks"; their numbering increases from SW to NE along the major axis, starting with the brick enclosing the galaxy nucleus, B01 (see Fig. 1 in Dalcanton et al. 2012). In addition to the optical (F475W, F814W; equivalent to g, I) ACS images, PHAT also obtained near-UV (F275W, F336W; the latter is equivalent to U) and near-IR (F110W, F160W; similar to J, H) imaging using the HST/WFC3 instrument. Additional information about PHAT imaging data and survey design is available in Dalcanton et al. (2012) and Paper I.

In addition to the PHAT data, we also processed and prepared ACS data from the HST archive (PID: 10273; PI: Crotts) that covered portions of M31 not imaged by PHAT. The

[5] `http://talk.andromedaproject.org`

imaging footprint for these data are also shown in Figure 3.2. This program obtained two-filter optical imaging using filters (F555W, F814W) similar to those used by PHAT, allowing easy incorporation into the AP search. Due to significant differences in data richness for objects identified in the archival dataset compared to the PHAT imaging, we choose not to include these objects in any further analysis, but we present object catalogs in Section 3.8.5.

We created AP search images using 725×500 pixel (36.25×25 arcsec; ∼6.9×4.8 pc in projected size) extractions from single-field ACS images. This subimage size efficiently divides the image area, includes 100 pixels of overlap between neighboring subimages to reduce incompleteness and biases caused by image edges, and allows participants to search images at full resolution. The parent ACS images have missing data due to the camera's chip gap that we filled using overlapping data from neighboring ACS images. Gaps, edges, and other artifacts are still present in some images, but our efforts mitigated most issues concerning missing data. We created a total of 13,017 subimages (4.7 gigapixels) from imaging that spans the entire PHAT survey region, as well as 1,728 addition subimages from archival imaging.

In addition to the normal imaging, we also produced additional search images that included synthetic clusters. The primary reason for inserting these synthetic test objects is to measure the cluster catalog completeness as a function of age, mass, size, and environment. In addition, the synthetics provided feedback to our volunteers: when a participant identified a synthetic cluster, they were notified that the object was synthetic and congratulated on their find. Participants on the site's Talk forum confirm that these notifications acted as positive reinforcement that they were performing the task they set out to accomplish.

We used the Year 1 cluster catalog results and its small number of accompanying synthetic cluster tests to create a realistic variety of clusters for insertion into AP search images. To begin, we choose age, mass, metallicity, attenuation, and effective radius values for the synthetic clusters drawn from distributions in each parameter:

- Ages were drawn from a flat distribution of discrete log(Age/yr) values between 6.6 and 10.1, spaced at an increment of 0.05 dex to match the grid of stellar isochrones

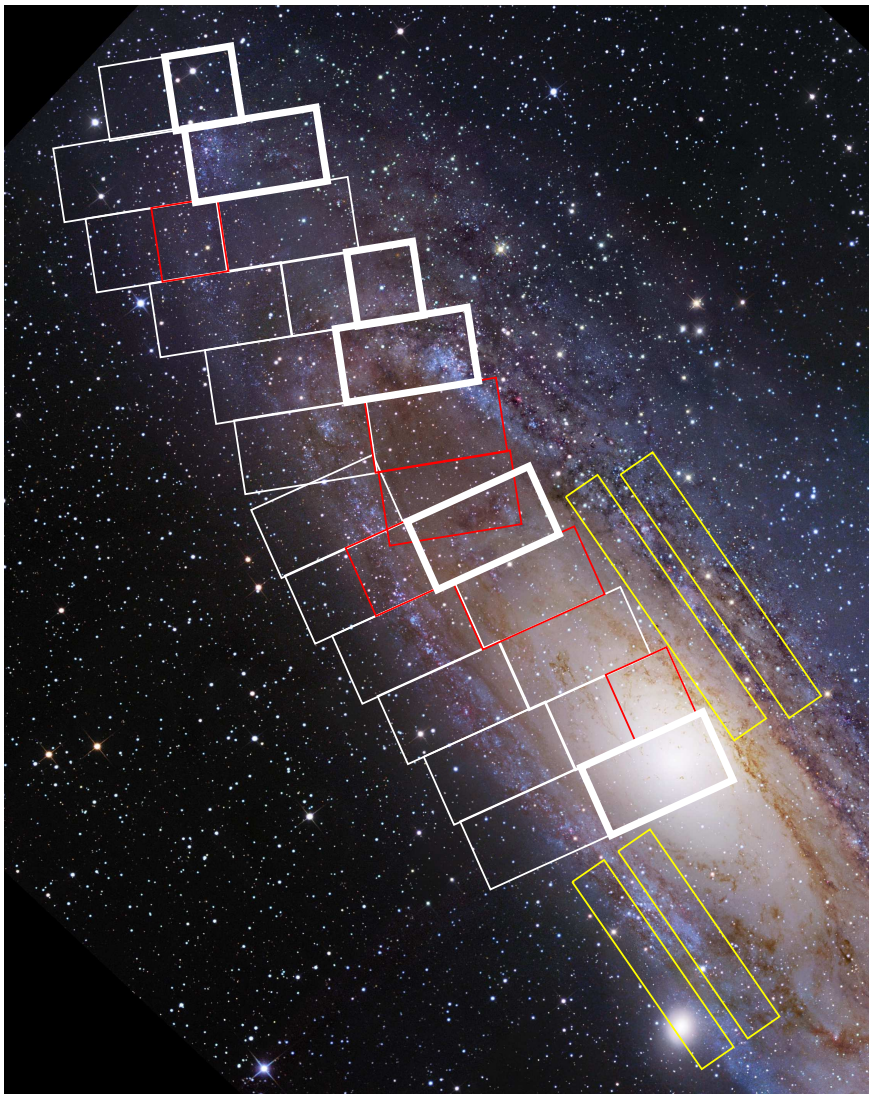

Figure 3.2 Map of M31 showing HST imaging footprints, oriented such that north is up and east is left. Color-coded regions denote various subsets of data: PHAT images searched during 2012 AP campaign (white); PHAT images searched during 2013 AP campaign (red); HST archival images searched during 2013 AP campaign (yellow). Bold white regions show areas searched during Year 1 effort (Paper I). Image Credit: Robert Gendler.

from the Padova stellar evolution models (Marigo et al. 2008; Girardi et al. 2010).

- Masses were drawn randomly from a continuous flat distribution of $\log(\mathrm{Mass}/M_\odot)$ between 2.0 and 5.0, yielding usable sample sizes across the full range of masses.

- Solar metallicity ($Z = 0.019$) was assumed for ages less than 5 Gyr. For ages greater than 5 Gyr, the metallicity was selected from a grid of Z to simulate the presence of metal-poor globular clusters: 0.0001 (0.005 $Z_\odot$), 0.001 (0.05 $Z_\odot$), 0.004 (0.2 $Z_\odot$), 0.008 (0.4 $Z_\odot$), 0.019 ($Z_\odot$).

- Extinctions were drawn from an exponential A_V distribution ranging from 0.17 mag (foreground Galactic extinction; Schlafly & Finkbeiner 2011) to 3.0 mag following the expression

$$P(A_V) \propto e^{-A_V/1.34}. \tag{3.1}$$

  This distribution was chosen to match the extinction distribution derived by Fouesneau et al. (2014) from their integrated light fitting of the Year 1 cluster catalog.

- Spatial profiles are defined using a King (1962) profile with a fixed concentration ($R_{\mathrm{tidal}}/R_{\mathrm{core}} = 30$) and an effective radius (R_{eff}; equivalent to half-light radius) drawn from a distribution of measured half-light radii presented in Paper I, but with a linear bias towards larger radii. We include this bias to boost the number of extended objects and ensure our ability to characterize the completeness of diffuse clusters. The resulting R_{eff} distribution peaks at 1.5 pc (0.39 arcsec) and extends from 0.5–9.0 pc (0.13–2.4 arcsec).

After drawing cluster parameter combinations, we populated individual cluster star lists using the Padova models, assuming a Kroupa (2001) IMF. We computed total luminosities for each cluster and selected a subset for insertion into search images that straddle the detection limit, as computed for the Year 1 catalog. In Paper I, we found that the sample was 100% complete for clusters brighter than $m_{F475W} = 18.5$ and 0% complete for clusters fainter than $m_{F475W} = 23.5$. Furthermore, when we take cluster age into account we can narrow

the range of acceptable m_{F475W} values even more: for $6.6 < \log(\mathrm{Age/yr}) < 8.0$ we adopted $18.5 < m_{F475W} < 22.0$; for $8.0 < \log(\mathrm{Age/yr}) < 9.0$ we adopted $19.5 < m_{F475W} < 22.5$;1 and for $9.0 < \log(\mathrm{Age/yr}) < 10.0$ we adopted $20.0 < m_{F475W} < 23.0$. These ranges allow us to efficiently map the functional form of completeness as a function of F475W magnitude at all ages.

Once we were satisfied with the sample, we inserted synthetic clusters into F475W and F814W images using the DOLPHOT photometry package, an updated version of HSTphot (Dolphin 2000) that is used by the PHAT collaboration for point-spread function photometry. These synthetic clusters were added into search images, one cluster per subimage, positioned pseudo-randomly within the image but always >120 pixels from the image edge. We spatially distributed the synthetic clusters across the PHAT survey footprint, covering a wide range of galactic environments to ensure our ability to evaluate completeness throughout M31. We selected fields that sample the survey's image variety, as defined by per-image red giant branch (RBG) star counts[6]. We inserted synthetic clusters into fields with $10^2 < N(\mathrm{RGB}) < 10^3$, and inserted proportionally less synthetics into fields with $N(\mathrm{RGB}) < 400$ to achieve a uniform number of synthetic clusters per $N(\mathrm{RGB})$ bin within this range. This selection results in the placement of synthetic clusters into regions where a majority of cluster identifications are made.

3.2.3 Data Collection & Classification Statistics

We obtained AP data during two rounds of collection; the first ran from 5–21 December 2012 and included 72% of the PHAT images. The remaining PHAT images and archival images were searched between 22–30 October 2013. Defining a classification as a volunteer's submitted response to a single image (containing zero to many individual markings), AP volunteers performed a total of 1.82 million image classifications. This corresponds to an average rate of about 70,000 classifications per day; our peak classification rate was over 80,000 classifications per hour.

A total of 29,262 unique users participated in the AP; 9,663 of these participants logged

[6]RGB stars are defined as sources with F110W−F160W > 0.5 and F160W < 21.0; see Section 3.4.

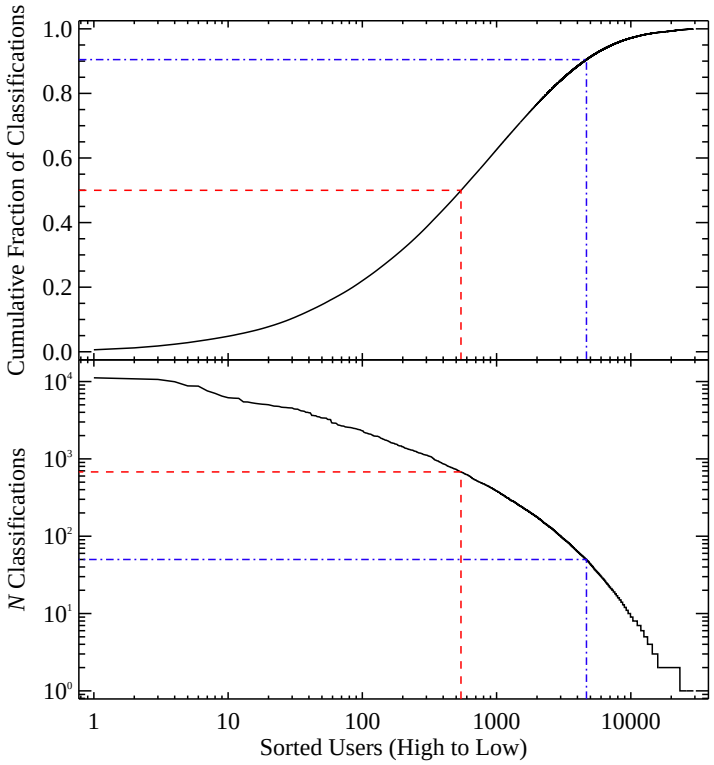

Figure 3.3 Classification statistics for AP. Participants are sorted as a function of decreasing contribution to the project and plotted logarithmically on the x-axis. Top: Cumulative fraction of 1.82 million classifications submitted by the top N volunteers. Bottom: The number of classifications submitted individually by the Nth volunteer. The red dashed lines highlight that half of the classification work (cumulative fraction = 0.5) was performed by the top 543 participants, who each classified $\geq$678 images. The blue dash-dotted lines highlight that 90.5% of the total classifications were submitted by the 4,671 participants who each classified $\geq$50 images.

in using a Zooniverse account. While the median number of classified images among all users was only 3 images (27 when only considering registered participants), 90.5% of our image classifications were performed by volunteers who examined at least 50 images. The distribution of work completed by the AP volunteer community is shown in Figure 3.3. The combined effort of Andromeda Project volunteers totals approximately 24 months of constant human attention.

Each image was classified a minimum of 80 times, but the distribution of classifications per image extends up to 108 with a median of 88. The classification counts vary slightly

between the two rounds of data collection: the median for the 2012 campaign is 86, while the median for the 2013 campaign is 101. In all, participants made >2 million individual cluster and galaxy identifications.

3.3 Catalog Construction

The primary goal of this work is to construct a catalog of clusters from the identifications provided by the project's participants. In this section we describe the process of converting clicks to scientifically-valuable data products. We evaluate the reliability of the crowd-sourced results and choose appropriate catalog thresholds by comparing to the PHAT Year 1 catalog (Paper I), an expert-derived "gold standard" reference.

The first step of catalog construction is to synthesize a merged list of identifications. We describe the details of our catalog creation algorithm and show examples of its application in Section 3.8.1. To briefly summarize: we spatially merge object identifications on an image-by-image basis, then merge these intermediate results into a survey-wide catalog. The resulting raw catalog includes ∼54,000 candidate clusters and galaxies, although a vast majority of these are low significance detections as we discuss below. Synthetic cluster tests are analyzed using outputs from the per-image catalogs. Also, artifact identifications are processed separately from the cluster and galaxy identifications and will not be discussed as part of this work.

After assembling a set of candidate objects, we use three metrics to identify cluster candidates and separate them from galaxies:

- $f_{\rm cluster}$ – the fraction of volunteers who viewed the search image and identified the object as a cluster.
- $f_{\rm galaxy}$ – the fraction of classifications for an object that identified it as a galaxy.
- $f_{\rm clst+gal}$ – the fraction of volunteers who viewed the search image and identified the object as either a cluster or a galaxy.

These quantities are related by:

$$f_{\rm cluster} = f_{\rm clst+gal} \times (1 - f_{\rm galaxy}) \tag{3.2}$$

The $f_{\rm cluster}$ scores provide relative rankings for AP cluster candidates. The top panel of Figure 3.4 shows the overall distribution of $f_{\rm cluster}$ scores for all AP identifications. This plot shows a large number of low significance detections with respect to higher significance detections. The distribution of $f_{\rm clst+gal}$ values is nearly identical to the $f_{\rm cluster}$ distribution.

We begin our comparison between AP and Year 1 results by cross-matching the two catalogs. The bottom panel of Figure 3.4 compares the distribution of AP $f_{\rm cluster}$ scores for three categories of Year 1 cluster cluster classifications. We confirm the expectation that increasing AP $f_{\rm cluster}$ scores correlate with a greater likelihood that candidates are clusters.

The distribution of $f_{\rm galaxy}$ values is presented in Figure 3.5. The top panel shows a clear bimodality in $f_{\rm galaxy}$ values, signaling a clear cluster-versus-galaxy classification preference for a majority of candidate objects. The bottom panel confirms the accuracy of these classification preferences; the expert-derived cluster and galaxy classifications from the Year 1 catalog map to low and high $f_{\rm galaxy}$ scores, respectively. We also observe that $f_{\rm galaxy} = 0.3$ defines a division between clusters and galaxies that leads to a minimal number of misclassifications.

It is interesting to note that there is an apparent bias at intermediate $f_{\rm galaxy}$ values ($0.3< f_{\rm galaxy} < 0.5$), such that a majority vote of AP participants would not classify these objects accurately, according to expert-derived labels. We hypothesize that this bias may be caused by the default cluster setting for the site's marking tool, leading to the tendency to mark candidates, particularly questionable ones, as clusters. Whatever the cause may be, only a small number of objects in this range of $f_{\rm galaxy}$ could plausibly be considered for inclusion in the AP catalog as a cluster instead of as a galaxy: there are 125 (13) objects with $f_{\rm cluster} > 0.2$ (0.5) in the full catalog of AP identifications that fall within $0.3< f_{\rm galaxy} < 0.5$. Nevertheless, we adopt an $f_{\rm galaxy}$-based selection criterion to account for this bias and incorporate as much information as possible during classification. We use the observed $f_{\rm galaxy} = 0.3$ threshold throughout the remainder of the paper to differentiate cluster and galaxy candidates.

To select a catalog of likely clusters from the set of AP identifications, we use selection criteria based on the cluster candidate's $f_{\rm cluster}$ and $f_{\rm galaxy}$ values. While we've clearly defined an $f_{\rm galaxy}$-based selection criterion, we now need to define an $f_{\rm cluster}$ threshold that

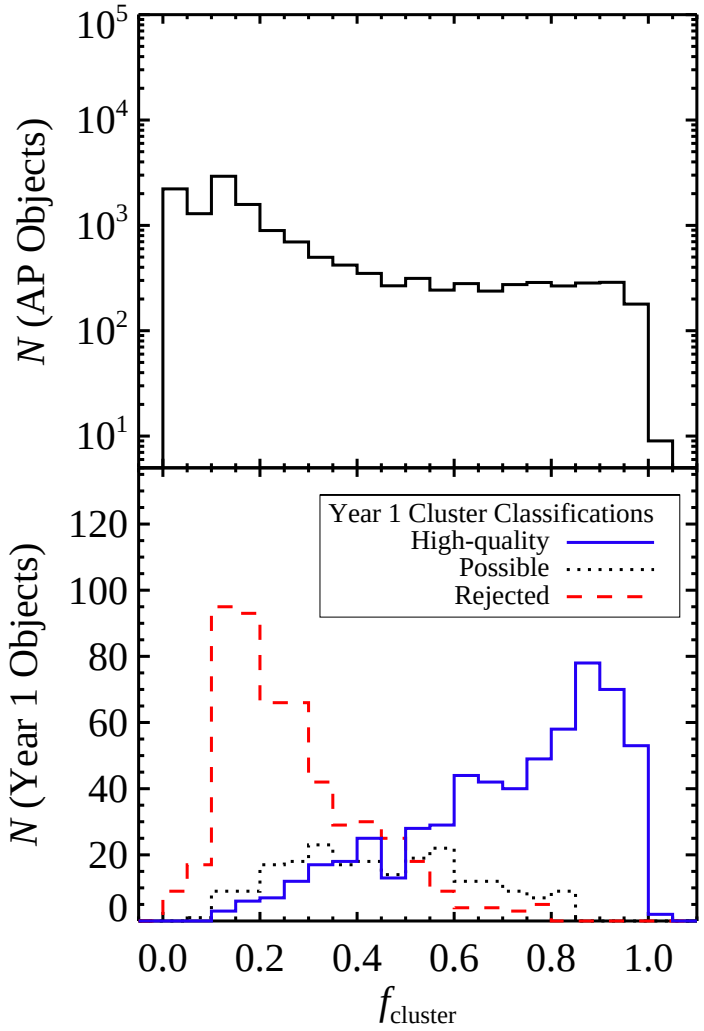

Figure 3.4 Top: Histogram of f_{cluster} values for the full catalog of AP identifications. Bottom: Histograms of f_{cluster} values for cross-matched high-quality (blue solid), possible (black dotted), and rejected (red dashed) Year 1 cluster candidates.

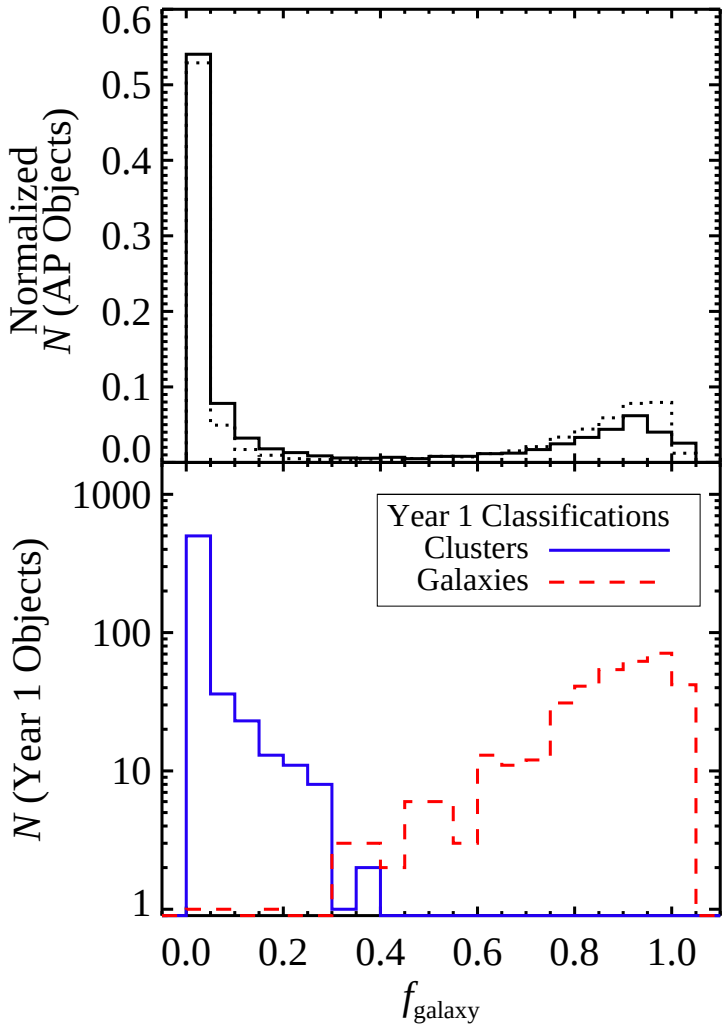

Figure 3.5 Top: Histogram of $f_{\rm galaxy}$ values, where the solid (dotted) lines represent the 13,801 (4,449) AP identifications with $f_{\rm clst+gal} \geq 0.1$ (0.5). The bimodality of this distribution shows the tendency for AP classifiers to strongly differentiate between clusters and galaxies. Bottom: Histogram of $f_{\rm galaxy}$ values for AP identifications cross-matched to the Year 1 galaxy (red dashed line) and high-quality cluster (blue solid line) catalogs. An $f_{\rm galaxy}$ threshold of 0.3 divides clusters and galaxies with minimal classification errors.

maximizes the number of clusters identified while minimizing the number of non-cluster contaminants. As the bottom panel of Figure 3.4 shows, these are directly competing goals; decreasing the $f_{\rm cluster}$ threshold to include a greater number of high-quality clusters necessarily introduces additional contaminants as well.

To evaluate how our choice of $f_{\rm cluster}$ threshold affects the resulting cluster catalog, we calculate completeness and contamination fractions based on a comparison between the AP and Year 1 catalogs within their shared search footprint in the disk of M31. We exclude the bulge region (Brick 1) from our comparison as its classification results differ sufficiently from the rest of the survey (see Section 3.5.1 for further discussion). We define completeness as the fraction of high-quality Year 1 clusters accepted by the AP selection criteria. Contamination is quantified as the fraction of accepted AP clusters that were previously classified as non-clusters or galaxies by the Year 1 catalog, or are new AP-only objects not identified or classified during the Year 1 search.

We note that these definitions of completeness and contamination make an imperfect assumption that the Year 1 search is flawless, in which no worthy clusters escaped identification and every high-quality cluster tabulated deserves that distinction. While this expert-derived catalog serves as a useful standard against which we can compare, it is inevitable that the completeness and contamination fractions we calculate with respect to the Year 1 catalog are approximate: 100% completeness will not be attained, and we expect a modest, non-zero contamination fraction. To evaluate previously unidentified objects, we could perform an expert review to individually assess these possible contaminants, however this strategy cannot remove the element of researcher subjectivity. Instead, we adopt an explicitly conservative stance that affects the absolute values of the contamination fractions we derive, but which do not impact the analysis choices we make due to the relative nature of most of these decisions.

We calculate a completeness versus contamination curve with respect to the expert-derived Year 1 catalog, akin to a receiver operator characteristic (ROC) curve. By continuously lowering the $f_{\rm cluster}$ threshold for the definition of AP clusters, we increase the completeness of Year 1 objects identified (bottom panel of Fig. 3.6). However, the decreasing $f_{\rm cluster}$ threshold also increases the contamination, defined as the fraction of the cluster

catalog objects that are either Year 1 non-clusters or new AP-only clusters (top panel of Fig. 3.6). In addition to the initial uniformly-weighted set of object identifications, we construct completeness versus contamination curves assuming different user weighting schemes, as discussed in Section 3.3.2. We compare the result from uniformly-weighted inputs (red) to the range of results obtained from a grid of weighting systems (gray), including the curve derived for our optimal weighting scheme (black).

To choose a catalog cutoff, we seek a metric that identifies the $f_{\rm cluster}$ cutoff value for which the resulting catalog achieves a balance between completeness and contamination. We choose to work directly with the completeness versus contamination curve and define $d_{\rm optimal}$, the distance from each point along the curve to the optimal corner of the plot (completeness and contamination fractions are 1.0 and 0.0, respectively). We note that our choice of metric, which values the minimization of false positives and false negatives equally, is somewhat arbitrary; given a specific use-case, one might prefer a metric that optimizes for a greater number of classifications at the expense of additional contamination. Our choice to weight completeness and contamination equally is grounded in the goal of creating a general-purpose catalog. Also, when we considering the specific shape of the completeness versus contamination curves we are working with, we find that this metric also tends to select the approximate point of diminishing return, the limit beyond which relaxing the catalog threshold tends to add more contaminants than additional good objects. On the completeness versus contamination plot this limit corresponds to the point at which the curve is tangent to a line with a slope of unity. In addition, it is also comforting that our choice of metric also tends to approximately conserve the number of clusters within the Year 1 footprint, yielding a similar number of clusters as found in Paper I. Together, the similarity of these limits gives us confidence that our specific choice of cutoff is appropriate.

We use the $d_{\rm optimal}$ metric to identify an optimal completeness and contamination combination of 85.3% and 10.5%, respectively, for the case of uniform user weighting; the corresponding $f_{\rm cluster}$ cutoff is plotted in Figure 3.6, which is tabulated along with other corresponding information in Table 3.1. We improve sample completeness and contamination fractions using a user weighting system, as we discuss in Section 3.3.2.

We select a catalog of likely background galaxies using a combination of $f_{\rm clst+gal}$ and

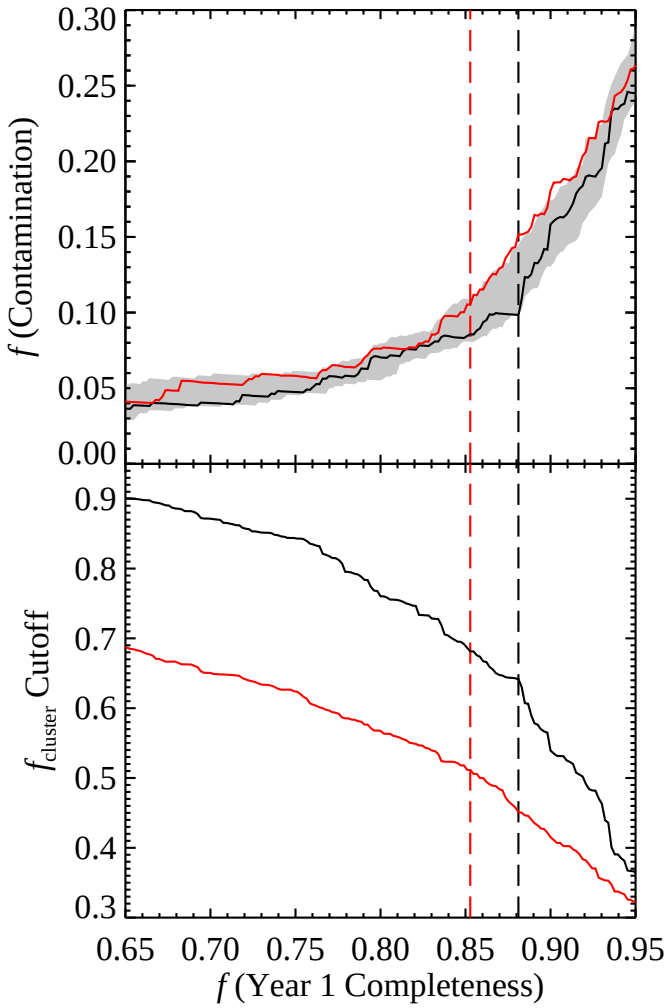

Figure 3.6 Top: Completeness versus contamination curves that result from uniform user weighting (red) and the optimal user weighting system (black). The gray shaded region denotes the parameter space covered by the sum of all curves derived for the grid of weighting systems we tested (see Section 3.3.2). Bottom: The $f_{\rm cluster}$ thresholds used by the uniform and adopted weighting systems as a function of Year 1 completeness. The vertical dashed lines in both panels denote the catalog limits adopted for each system based on the $d_{\rm optimal}$ metric.

Table 3.1. User Weighting Parameters

Name	Detection		Non-detection		$d_{optimal}$	Optimal	Optimal	$f_{\rm cluster}$
	$m_{\rm logistic}$	$b_{\rm logistic}$	$m_{\rm logistic}$	$b_{\rm logistic}$		Completeness	Contamination	Cutoff
Uniform Weights	···	···	···	···	0.1809	0.8528	0.1052	0.5114
Uniform Weights (Match Comp)	···	···	···	···	0.1928	0.8811	0.1518	0.4512
Uniform Weights (Match Cont)	···	···	···	···	0.1828	0.8453	0.0974	0.5214
Best Weights	16.0	0.6	39.0	1.1	0.1543	0.8811	0.0984	0.6416

$f_{\rm galaxy}$ selection criteria in a process similar to the one described here for the clusters. We document that analysis and its accompanying details in Section 3.8.2.

3.3.1 $f_{\rm cluster}$ Uncertainties & Robustness

To demonstrate the robustness of our $f_{\rm cluster}$ metric, quantify its associated uncertainties, and establish its consistency across two separate rounds of data collection, we carried out a repeatability experiment during the 2013 campaign. We selected 741 images (397 normal, 344 synthetic) searched during the 2012 campaign (Round 1; R1) that included highly-ranked cluster candidates and repeated data collection for these images during the 2013 campaign (Round 2; R2). We match the catalogs that emerge from each run and compare $f_{\rm cluster}$ scores for 1,241 objects whose R1 and R2 scores average to >0.35 (891 from normal data, 350 from synthetic data) to test the repeatability of $f_{\rm cluster}$ scores for likely clusters. We present the distribution of $f_{\rm cluster}$ differences between the two rounds in Figure 3.7. We model the $\Delta f_{\rm cluster}$(R1 − R2) scatter using an expression for the combined variance of two drawing experiments governed by the binomial distribution:

$$\sigma(p) = \sqrt{\frac{2p(1-p)}{N}}, \tag{3.3}$$

where $N = 88$, representing the median number of image views, and p is the averaged R1 and R2 $f_{\rm cluster}$ score for each object, representing our best estimate of an object's "true" $f_{\rm cluster}$ value. We plot 1, 2, and 3σ contours as predicted by our noise model, which accurately captures the scatter shown in the data.

These results demonstrate that image classifications collected during the 2012 and 2013

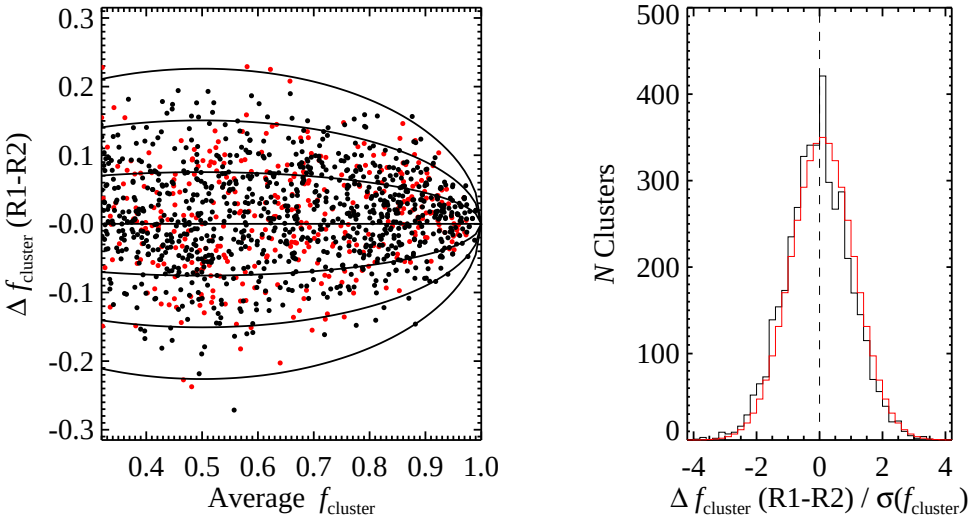

Figure 3.7 Left: Comparison of $f_{\rm cluster}$ for clusters derived from the 2012 campaign (R1) data versus those from the 2013 campaign (R2). Black points reflect measurements made in normal images, red points reflect measurements made in synthetic images. We plot 1, 2, and 3σ contours showing the scatter predicted by our noise model. Right: Histogram of $f_{\rm cluster}$ differences scaled by the expected dispersion. A Gaussian function with σ=1 and a peak value of 350 is overlaid for reference. The dispersion of the $f_{\rm cluster}$ differences between the two rounds matches the statistical expectation of the noise model.

campaigns are functionally equivalent, allowing us to easily combine data from the two rounds. This experiment also shows that our procedure of combining >80 image classifications from the pool of AP participants provides consistent $f_{\rm cluster}$ results with minimal systematic biases.

3.3.2 User Weighting

Up to this point, we have assumed that the abilities of all classifiers are equal on average. In this section we investigate whether weighting individual volunteers based on the quality of their classifications can improve the cluster sample. User weighting has been applied in several other citizen science projects (Lintott et al. 2008; Schwamb et al. 2012) and seems naturally applicable to our AP data. In line with these previous implementations, we calculate weightings based on the level of agreement between a participant's classifications

and the consensus opinion of all the volunteers. Individuals who agree with the consensus opinion are up-weighted, while those who disagree with the consensus opinion are down-weighted. Expanding beyond previous implementations, we vary the strength and form of weighting, evaluating the success of each iteration by comparing completeness versus contamination curves (derived through comparison to the Year 1 sample) to the unweighted case presented earlier in Section 3.3.

We could have chosen another way to assign weights, such as assessing a volunteer's performance with respect to expert-derived Year 1 results, or basing weights on a participant's recovery rate of synthetic clusters. One downside suffered by both of these alternative methods: resulting weights are based on only a fraction of the available classification data. Decreasing the volume of classifications considered by the weighting system leads to an increasing number of participants with little or no assessment information, and noisier ability estimates for every volunteer. Additionally, weighting systems tend to produce catalogs that resemble the data used for training and calibration. We were concerned that defining weights based on data that did not sample the variety and parameter ranges included in the full cluster sample might result in unwanted biases. Particularly in the case of the synthetic cluster data, which was specifically designed to characterize cluster recovery near the detection limit, these biases could be significant. To exploit the unique benefits provided by our crowd-sourced methodology, we utilize the unfiltered opinion of AP volunteers.

Figure 3.8 shows two quantities that we use to characterize the performance of our volunteers: the fraction of consensus clusters a volunteer identified, $f_{\rm consensus}$, and the mean $f_{\rm cluster}$ of all cluster identifications made by a volunteer, $\bar{f}_{\rm cluster}$. We define consensus clusters as objects that show a high degree of agreement among AP participants, where $f_{\rm cluster} > 0.6$ and $f_{\rm galaxy} < 0.2$; these limits provide a sample with a sufficient number of clusters to enable weighting of individual participants while ensuring that weights are not based on questionable candidates (see Figure 3.4).

Examination of Figure 3.8 reveals that there is wide variation of classification behavior among AP volunteers. Individuals that lie in the upper left part of the plot are conservative classifiers; everything they clicked was an obvious cluster, leaving many consensus clusters unmarked. On the other hand, participants in the lower right are liberal classifiers; they

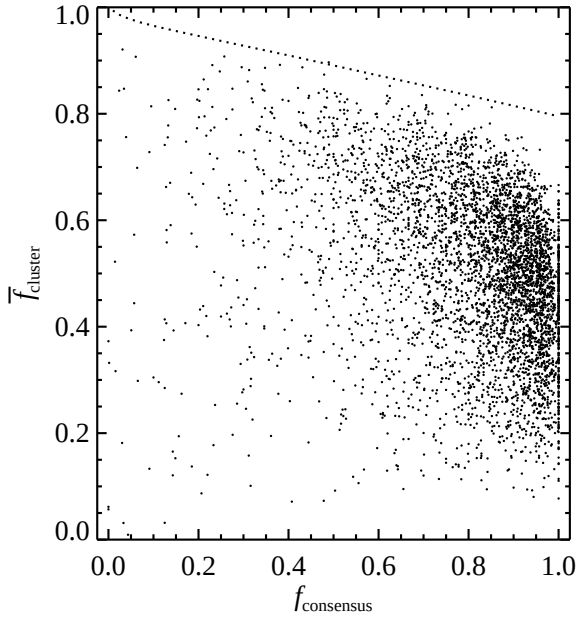

Figure 3.8 Performance metrics for 4,671 volunteers who classified ≥50 AP search images. The x-axis represents $f_{\rm consensus}$, the fraction of consensus clusters ($f_{\rm cluster} \geq 0.6$ and $f_{\rm galaxy} < 0.2$) identified by each participant out of the total number of consensus clusters they saw. The y-axis represents $\bar{f}_{\rm cluster}$, the average $f_{\rm cluster}$ value of all clusters identified by that volunteer. The dotted line represents an approximate ceiling to $\bar{f}_{\rm cluster}$ values as a function of $f_{\rm consensus}$, calculated by considering the intrinsic $f_{\rm cluster}$ distribution of the consensus cluster sample. Conservative classifiers, those that identify only the best cluster candidates, lie in the upper left portion of the plot. Liberal classifiers, those that identify most good clusters but also identify many low-ranked candidates, lie in the bottom right portion of the plot.

identified a large fraction of consensus cluster sample, but also identified many other low-ranked objects that are not likely clusters. Volunteers with scores that lie in the upper right portion of Figure 3.8 are desirable classifiers, obtaining high completeness but with little sacrifice to the overall quality of their identifications. We note that because of the intrinsic $f_{\rm cluster}$ distribution of the good clusters, volunteers who identify a large fraction of the good clusters cannot have an average $f_{\rm cluster}$ of 1.0; we compute the upper limit to average $f_{\rm cluster}$ based on the $f_{\rm cluster}$ distribution of good clusters and plot this envelope as a dashed line in Figure 3.8.

To make best use of classifications from both conservative and liberal cluster identifiers, we apply separate weightings to volunteer's detections and non-detections. Specifically, we weight a participant's detections based on the average $f_{\rm cluster}$ of clusters they identified, while non-detections are weighted based on $f_{\rm consensus}$, the fraction of consensus clusters the volunteer identified. As an example: in the case where a liberal classifier in the lower right corner of Fig. 3.8 did not click on a cluster, their non-detections are up-weighted because they are known to identify most good clusters. The detections from the same classifier, however, are down-weighted because this individual identifies many low-quality cluster candidates in addition to the high-quality ones.

We adopt a threshold number of subimage classifications above which we can assume we have adequately characterized a participant's classification behavior. Volunteers with fewer than 50 subimage classifications are distributed with greater randomness across the $\bar{f}_{\rm cluster}$ versus $f_{\rm consensus}$ plane, suggesting large uncertainties in the values of their performance metrics; we adopt 50 classifications as the threshold. Individuals who fall below this classification threshold are assigned mean detection and non-detection weights. Even when this limit is imposed, $\sim$90% of all image classifications are weighted using individually determined user weights. We note that anonymous accounts from unregistered users are treated in the same way as those from registered users for weighting purposes. Most of these users are assigned mean detection and non-detection weights due to the fact that they submit a small number of classifications (median number of classifications is 2); $\sim$5% of unregistered users surpass the minimum subimage threshold for individual weight assignment.

Next we determine how to translate performance metric scores into relative user weights.

We adopt a general form for the transformation based on the generalized logistic function. Favorable aspects of this functional form include its tunable scaling and that it allows for the saturation of weights at high and low input metric scores. Our "constrained" logistic function is defined as:

$$W(x) = B \times \left(A + \frac{1}{1 + e^{-m_{\text{logistic}}(x - b_{\text{logistic}})}} \right), \tag{3.4}$$

where x represents the input performance metric (either $\bar{f}_{\text{cluster}}$ or $f_{\text{consensus}}$) while m_{logistic} and b_{logistic} are the slope (growth rate) and the offset (position of maximum growth) of the logistic curve, respectively. The variables A and B are normalizations set such that W varies from 0 to 1 over the interval $x = [0, 1]$, providing the constrained aspect of this function. Once a set of logistic function parameters have been chosen for the detection and non-detection weighting functions, we apply user weightings to individual cluster votes on an image-by-image basis and recalculate weighted f_{cluster} values, $f_{\text{cluster,W}}$.

We vary the input logistic function parameters to search for a set of values that produce the best possible weighted catalog. We construct a grid of weighting systems by varying the values of the four free parameters: the slope and offset values for both the detection and non-detection weights. For each set of parameters, we calculate a completeness versus contamination curve and its corresponding minimum distance to the corner of optimal completeness and contamination, d_{optimal}. We gradually extended the weighting grid to include an increasing range of logistic parameter values until we identified a minimum d_{optimal} value that was unsurpassed. We defined the set of parameters that yielded this minimum d_{optimal} value as our optimal AP weighting system.

The range of completeness versus contamination curves is represented by the gray region in the top panel of Figure 3.6. We also plot the individual curve derived for the optimal weighting system and list its logistic function parameters in Table 3.1. The optimal weighting system provides a contamination fraction of 9.8% at a completeness of 88.1%. When compared to the uniform weighting results, applying user weighting decreases the number of contaminants by 36% (from $f_{\text{contamination}}$ of 0.152 to 0.098 at completeness of 88.1%), or alternately increases completeness from 84.6% to 88.1% (at $f_{\text{contamination}}$ of 0.098). While user weighting does not dramatically change the total number of cataloged clusters or the

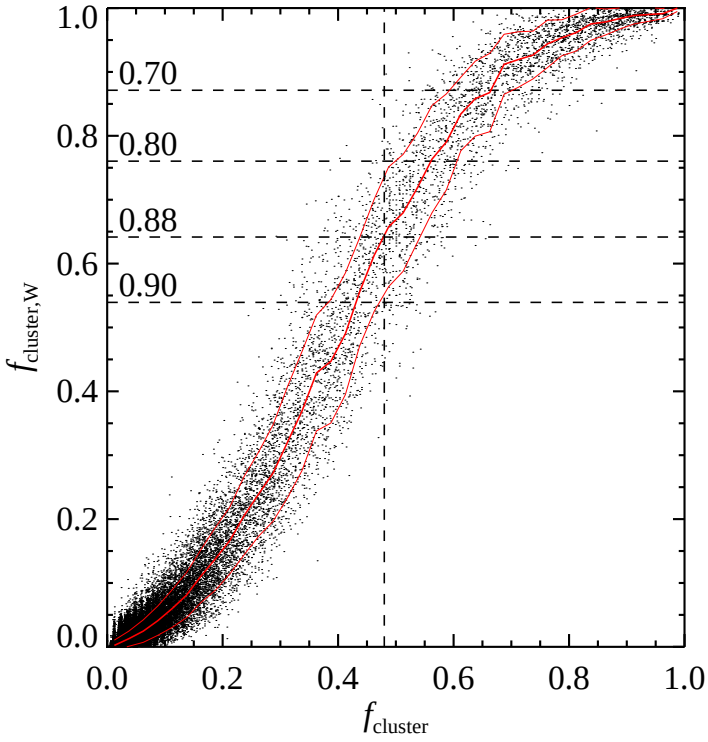

Figure 3.9 A comparison between optimally-weighted $f_{\rm cluster,W}$ scores and uniformly-weighted $f_{\rm cluster}$ values, showing the impact of user weighting on individual object scores. The red lines show the median trend and one standard deviation around the median. Horizontal dashed lines denote the $f_{\rm cluster,W}$ cutoffs corresponding to each of the printed Year 1 completeness fractions, while the vertical dashed line denotes the approximate $f_{\rm cluster}$ value that corresponds to the optimal $f_{\rm cluster,W}$ cutoff.

Year 1 completeness percentage, we are able to reduce the number of possible contaminants by a significant amount.

We compare original versus weighted $f_{\rm cluster}$ values to illustrate the impact of user weighting on individual clusters. Figure 3.9 shows that user weighting tends to increase the separation between high and low $f_{\rm cluster}$ objects, providing better differentiation at moderate $f_{\rm cluster}$ values that lie near the catalog cutoff. To visualize how the choice of $f_{\rm cluster,W}$ cutoff affects the output cluster catalog, we represent four different threshold values as horizontal lines in the figure, each labeled according to its corresponding Year 1 completeness fraction. We also plot a vertical line in Figure 3.9 representing the approximate $f_{\rm cluster}$ cutoff that best approximates the optimal $f_{\rm cluster,W}$ threshold.

The user weighting applied here enhanced final AP catalog results by achieving small but quantifiable improvements through a combination of decreased contamination and increased completeness. We note that we were fortunate to obtain a large number of classifications per image (>80) allowing us to account for variations in participant performance by averaging over a large number of classifications. Many citizen science projects cannot afford to collect a similar number of per-image classifications because they need to distribute effort over a larger volume of data, or because the project is working on time-sensitive tasks that cannot wait for additional input to be collected. In these cases, we expect that user weighting would play an essential role in obtaining high-quality results.

We utilize the $f_{\rm cluster,W}$ values as defined by the optimal user weighting system throughout the rest of the paper.

3.4 Catalog Completeness

We introduced our set of synthetic cluster tests in Section 3.2.2; here we present catalog completeness results derived from those tests, including how catalog completeness correlates with properties of the clusters and their surrounding fields.

The traditional method of characterizing the completeness of a cluster catalog is to identify the 50% completeness limit as a function of cluster luminosity. The two plots in the left column of Figure 3.10 show the behavior of the 50% completeness limit in F475W as a function of cluster age for the full sample of synthetic clusters. These plots show that while

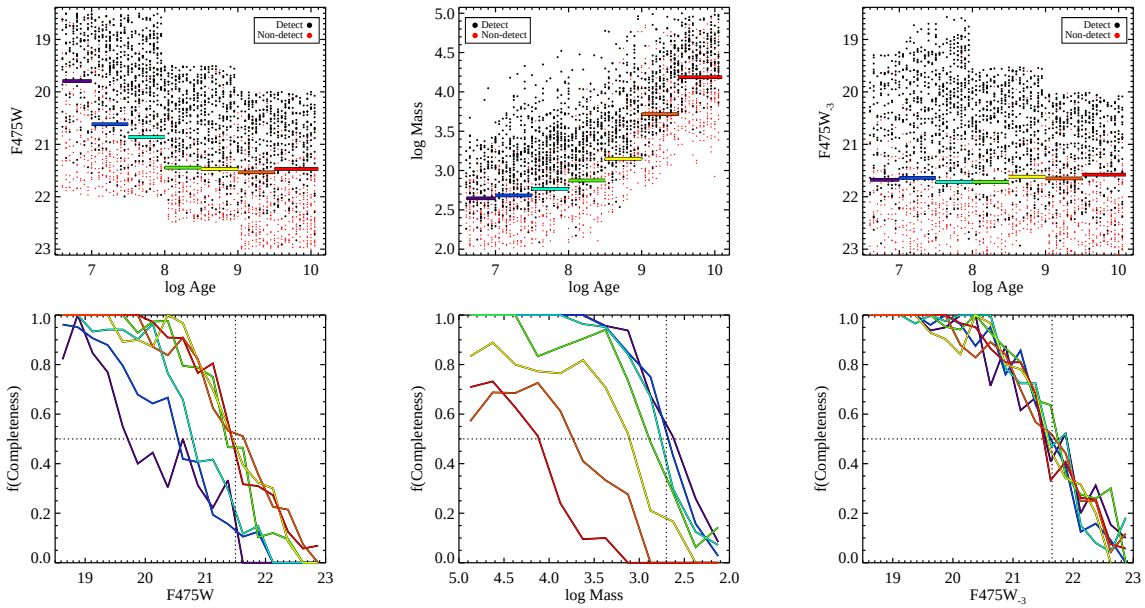

Figure 3.10 Completeness results from synthetic cluster analysis. Top Panels: Detection results for individual synthetic clusters (black = detected, red = not detected), as well as 50% completeness limits calculated for each age bin. Bottom Panels: Completeness functions for each age bin, color-coded to match their corresponding bin in the top plot. Results as a function of F475W magnitude, mass, and F475W$_{-3}$ magnitude are presented in the left, center, and right columns, respectively. F475W$_{-3}$ magnitudes represent the cluster flux that remains after subtracting the contribution of the cluster's three most luminous members.

the synthetic results at log(Age/yr) > 8.0 agree with a single, age-independent magnitude limit at F475W $\sim$ 21.5, there is an apparent age dependence at younger ages. This result conflicts with the standard assumption that luminosity-based completeness limits for cluster catalogs are independent of age.

To understand why we find brighter, non-constant completeness limits at ages <100 Myr, we examine our completeness results as a function of cluster mass, presented in the middle column of Figure 3.10. Under the assumption of an age-independent, constant luminosity completeness limit, we would expect a continuous increase in the 50% mass completeness with increasing age due to stellar evolution driven fading of the cluster's stars. In contrast to these expectations, we find a near-constant 50% completeness limit for log(Age/yr) < 8.0 of $\sim$500 $M_{\odot}$. It appears that catalog completeness correlates with cluster mass rather than luminosity at ages <100 Myr.

To explain the observed completeness behavior, it is important to note that nearly every synthetic cluster we tested with an age <100 Myr has a mass $<3\times10^3 M_{\odot}$. The integrated light of young low mass clusters is dominated by a small number of bright stars. This fact leads to large stochastic variations in the total integrated light for a sample of clusters with identical masses (see Fouesneau & Lançon 2010; Beerman et al. 2012; Popescu et al. 2012). In addition, cluster identification in HST imaging of M31 relies greatly on the presence of an over-density of individually resolved stars, such that the number of observable stars might correlate better with a cluster's detection probability than its total luminosity in this low-mass regime. In this case, the correlation between completeness and mass is explained by a strong correlation between mass and the number of bright, observable cluster members.

We conclude that there are two regimes for AP cluster catalog completeness: for ages <100 Myr, cluster detection is limited by the number of observable member stars; for ages >100 Myr, cluster detection is governed by the total cluster luminosity. To bridge these regimes, we devise a single cluster metric that correlates strongly with the 50% catalog completeness limit, independent of cluster age: F475W$_{-3}$, the F475W magnitude remaining after subtracting the flux contribution from the cluster's three brightest stars. By excluding the contribution of the three brightest cluster stars, we significantly reduce the stochastic variation in cluster luminosity that imprinted an age-dependence into the completeness

results. We experimented with the number of stars to exclude and found that three provided the best correction. The plots in the right column in Figure 3.10 show that our data are consistent with a single, age-independent 50% completeness limit at a F475W$_{-3}$ magnitude of 21.65, where the new metric successfully unifies the two completeness regimes.

Using the results derived from the full set of synthetic cluster tests as a baseline, we can test whether completeness depends on two other important factors: the spatial profile of the cluster and the characteristics of the field surrounding the cluster. At a fixed luminosity, we expect the completeness to worsen for larger, more extended clusters because the same total luminosity is spread over a larger area, causing the contrast between cluster and underlying background to decrease. Likewise, the cluster to background contrast also decreases as the background surface brightness and stellar density increase, which also leads to a prediction of brighter cluster luminosity completeness limits.

Contrary to the simple expectation, we observe non-monotonic behavior in the 50% completeness limit as a function of a cluster's effective radius ($R_{\rm eff}$; equivalent to the half-light radius), as shown in the top panel of Figure 3.11. While the 50% completeness limit reaches its faintest value at $\log(R_{\rm eff}/{\rm arcsec}) \sim -0.35$, detection limits worsen as clusters become more extended, the limits also worsen as clusters become more compact. Detection becomes more difficult at small $R_{\rm eff}$ due to the inability for an image classifier to distinguish between single sources and a compact collection of individual stars. This behavior was also seen by Silva-Villa & Larsen (2011) in their study of extragalactic clusters. The variation in $R_{\rm eff}$ can cause F475W$_{-3}$ 50% completeness limits to deviate by >0.5 mag from the baseline level, translating to a mass completeness difference of up to 0.15-0.2 dex.

Background stellar density, on the other hand, shows the expected behavior that higher stellar density makes cluster detection more difficult. We quantify local stellar density by counting the number of red giant branch (RGB) and main sequence (MS) stars that lie within the search images (36.25×25 arcsec) that host each synthetic cluster. These counts are based on the survey-wide 6-band GST photometric catalogs (Williams et al. 2014), where we define RGB stars as sources with F110W-F160W > 0.5 and F160W < 21.0, and MS stars as sources with F475W-F814W < 1.0 and F814W < 25.0. The middle and bottom panels of Figure 3.11 show 50% catalog completeness limits as a function of N(RGB) and

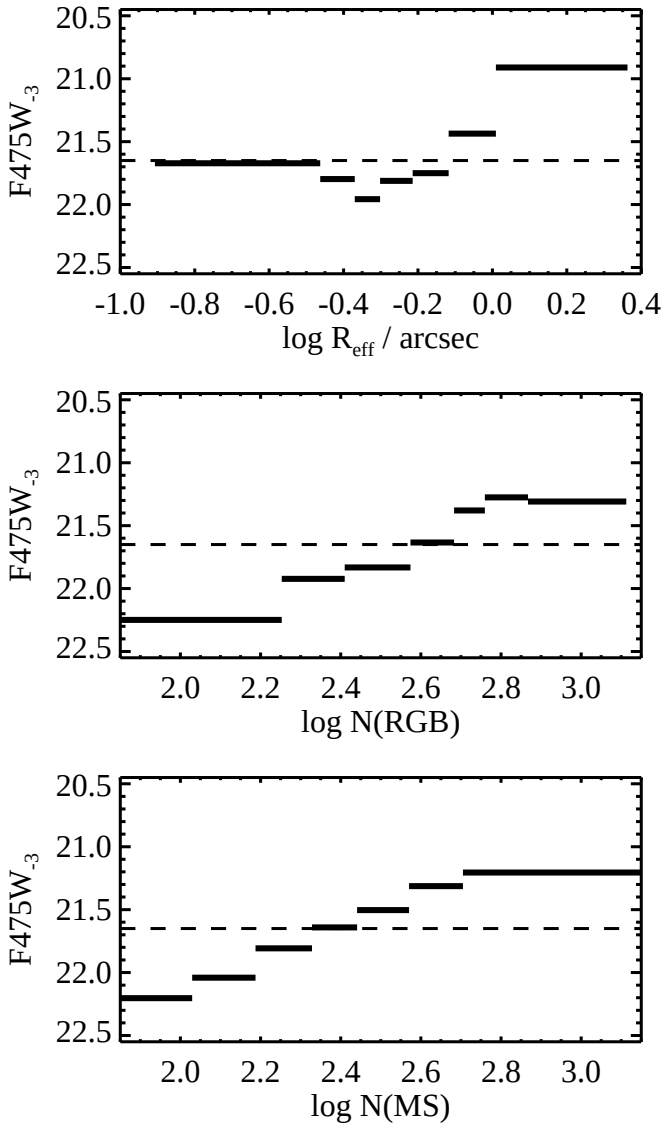

Figure 3.11 Deviations from average completeness in F475W$_{-3}$ magnitude as a function of $R_{\rm eff}$, N(RGB), and N(MS) in the top, middle, and bottom panels, respectively. The dashed line represents the baseline 50% completeness level of F475W$_{-3}$ of 21.65. Seven bins divide the synthetic cluster sample into equal parts ($N \sim 440$) as a function of each cluster variable.

N(MS). As a function of N(RGB) and N(MS), the F475W$_{-3}$ 50% completeness limits vary by ∼0.5 mag, translating to a mass completeness difference of up to 0.15-0.2 dex. This dependency affects the detection of PHAT clusters in the inner disk and bulge, as well as those within dense star forming regions – especially those located within the ∼10 kpc ring.

To supplement the above description of overall, sample-wide completeness behavior, we present a table of object-by-object completeness test results in Section 3.8.3. These results allow catalog users to calculate completeness functions for specific spatial regions or over a custom range in parameter space.

3.5 Results

3.5.1 AP Cluster Catalog

We apply the catalog construction techniques and user weighting methodology presented in Section 3.3 to define an AP cluster catalog, adopting final selection criteria of:

$$f_{\rm cluster,W} > 0.6416 \text{ AND } f_{\rm galaxy} < 0.3. \tag{3.5}$$

These criteria yield a sample of 2,714 clusters. We add two additional sets of clusters to these initial selections. First, we add 35 clusters to the sample that are located in the bulge-dominated region within ∼3 kpc of M31's center, as defined by an ellipse with a center of (10.684575, +41.268972), semi-major axis of 815 arcsec, semi-minor axis of 410 arcsec, position angle of 45 degrees, and bounded by the PHAT footprint. These objects are primarily globular clusters that were identified and confirmed by previous surveys. These objects suffer from systematically low $f_{\rm cluster,W}$ scores due to their atypical appearance (compact and smooth with few individually resolved stars), high-surface brightness backgrounds, and suboptimal search image scalings. We decided that the most straightforward solution to correct for these missed objects was to include all previously confirmed clusters (high-quality Year 1 or RBC flag of 1) that lie within the defined region and evaluate all candidate or possible objects. We confirmed by-eye that each of the previously confirmed objects has an appearance consistent with that of a cluster, and confirmed two additional candidate objects. Second, we include 4 additional expert cluster identifications from the B03 tooth images that were not included in the AP search due to delayed data availability.

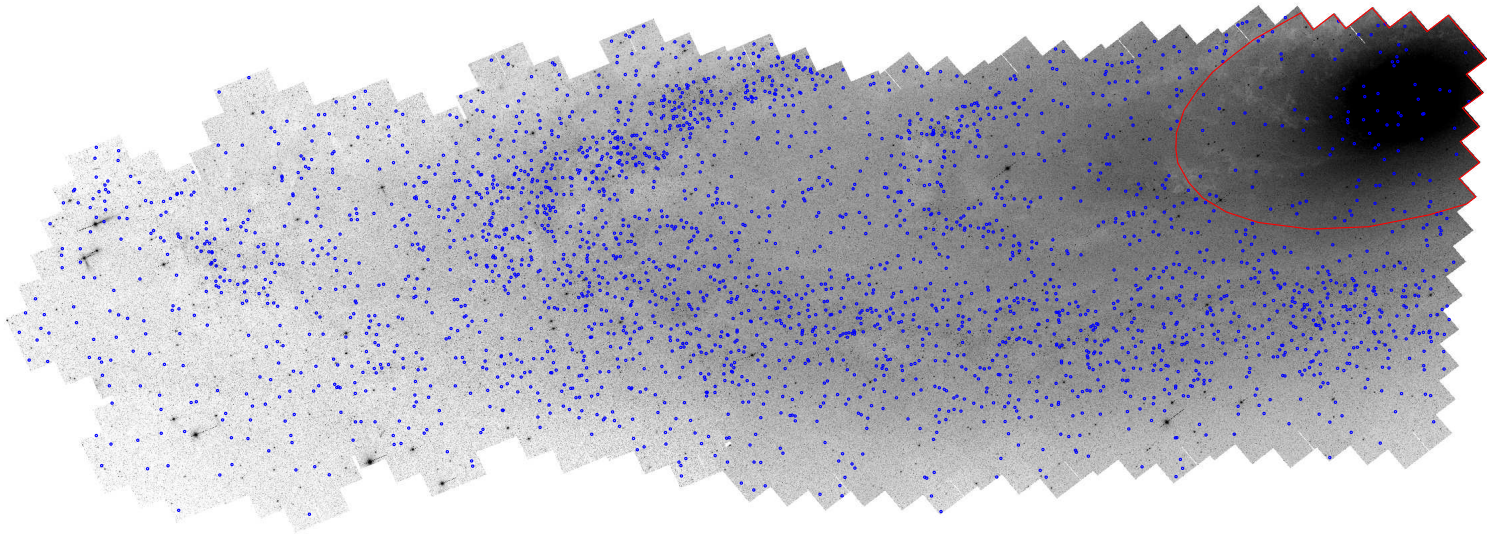

Figure 3.12 Spatial distribution of AP cluster catalog overlaid on the PHAT survey-wide F475W image. The red ellipse denotes the bulge region within which the catalog completeness and object recovery vary significantly from the rest of the survey.

Table 3.2. AP Cluster Catalog

AP ID	RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$R_{\rm eff}$ ($''$)	$m_{\rm ApCor}$[a]	F275W$_{\rm ap}$	σ	F336W$_{\rm ap}$	σ	F475W$_{\rm ap}$	σ
$f_{\rm clst+gal}$	$f_{\rm galaxy}$	$f_{\rm cluster,W}$	PC ID	Alt ID	Flags	F814W$_{\rm ap}$	σ	F110W$_{\rm ap}$	σ	F160W$_{\rm ap}$	σ
1	11.435516	41.698562	2.19	0.60	-0.01	20.12	0.04	19.16	0.01	18.77	0.01
1.0000	0.0000	1.0000	···	B374-G306	···	17.69	0.08	17.19	0.15	16.59	0.20
2	11.366514	41.701013	1.86	0.61	-0.04	20.91	0.10	20.25	0.02	20.01	0.10
0.9717	0.0083	0.9894	···	M085	···	19.05	0.21	>18.06	···	>17.06	···
3	11.471290	42.049246	1.95	0.88	-0.14	21.27	0.14	20.81	0.03	20.67	0.08
1.0000	0.0000	1.0000	···	···	···	20.07	0.31	>18.97	···	>18.07	···
4	11.474664	42.038351	2.87	0.98	-0.05	20.10	0.04	19.10	0.02	18.75	0.03
1.0000	0.0227	0.9909	···	B483-D085	···	17.78	0.08	17.29	0.16	16.66	0.21
5	10.991636	41.359328	1.46	0.40	-0.01	20.88	0.04	20.29	0.03	20.09	0.06
1.0000	0.0000	1.0000	···	M005	···	19.36	0.10	18.73	0.21	17.72	0.25

Note. — Table 3.2 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content. Three-sigma upper limits are denoted by a ">" symbol. PC ID and Alt ID refer to cluster identifiers from the Year 1 catalog and other literature sources, respectively. Flags: E = Extended Object (see Section 3.6.2); B = Bulge or B03 object manually added (see Section 3.5.1).

[a]Aperture Corrections are provided such that $m_{\rm Total} = m_{\rm ap} + m_{\rm ApCor}$.

The final AP cluster catalog includes 2,753 objects. Figure 3.12 shows the positions of the clusters within the PHAT survey footprint. Andromeda's ∼10 kpc star-forming ring is a prominent feature visible in the cluster's spatial distribution. We assign identifiers in descending order of their maximum per-image, uniformly-weighted $f_{\rm clst+gal}$ score. Positions, aperture sizes, and other relevant catalog metadata are presented in Table 3.2.

All AP candidates with $f_{\rm clst+gal} \geq 0.1$ that are not included in the cluster catalog (or galaxy catalog; see Section 3.8.2) are listed in an ancillary table in Section 3.8.3. We include information on these additional candidates to allow other workers the opportunity to make different choices concerning catalog selection.

3.5.2 Comparison to Previous Cluster Catalogs

We cross-match our AP cluster catalog with a selection of previously published catalogs: the Year 1 catalog, the RBC, Caldwell et al. (2009), and the HKC. We include alternate identifiers for previously classified objects in Table 3.2 and summarize the high degree of consistency between the AP catalog and previous results below.

By design, the AP catalog bears a strong resemblance to the Year 1 catalog. When we consider the portion of the AP catalog that lies within the Year 1 imaging footprint (including B01, differing slightly from the Section 3.3 analysis), we count 688 clusters, which is a 14.5% increase over the 601 object Year 1 catalog. The agreement between the two samples is good: the AP catalog includes 88.5% (532/601) of the good Year 1 clusters, and 91% of the AP cluster catalog were previously classified as high-quality or possible Year 1 objects. The AP catalog includes 39 Year 1 catalog rejections and 22 objects not classified in the Year 1 search. While the majority of object-by-object classification differences are caused by clusters with $f_{\mathrm{cluster,W}}$ scores that lie near the catalog cutoff, we discuss a number of meaningful systematic differences between the two catalogs in Section 3.6.2.

Comparison of the AP catalog to the RBC and the Caldwell et al. (2009) catalog provides an opportunity to cross-reference with commonly cited sources, linking our present work to a wealth of ancillary information about these clusters, including a great deal of follow-up spectroscopy. These ground-based catalogs do not reach the faint objects accessible to the PHAT imaging, therefore the following comparison mostly consists of verifying or discarding previously unconfirmed candidates that lie at the middle or bright end of the AP sample.

Cross-matching the AP catalog with the RBC, we find that 260 previously confirmed, candidate, or controversial clusters (RBC flag = 1, 2, or 3) match to AP clusters, while 42 AP classifications conflict with those from the RBC (40 AP clusters are not RBC clusters, 2 RBC clusters are not AP clusters), and 18 additional RBC candidate or controversial classifications were rejected. PHAT's high spatial resolution imaging is often used as a definitive tool for classifying objects, so we defer to AP classifications for these conflicting cases. We also find good agreement between the AP and the Caldwell et al. (2009) catalogs. Only 18 conflicts arise from the Caldwell catalog (8 AP clusters are not Caldwell clusters, 10 Caldwell clusters are not AP clusters), while 232 cluster classifications are common to both the Caldwell and the AP catalogs.

Finally, we compare the AP catalog with the HKC catalog compilation. These clusters represent the low-mass additions to previous ground-based catalogs provided by early targeted HST observations, and therefore include many objects that lie at or near completeness limits. As such, a direct comparison shows 156 previously identified clusters confirmed by

our AP classifications, while 57 are not confirmed. This 73% yield is nearly identical to the 72% yield we found for the Year 1 catalog during a similar comparison exercise. A vast majority of HKC objects that were not confirmed by the AP catalog are borderline, marginal candidates where there is a subjective difference in opinion between the HKC authors and the consensus judgement of AP volunteers; rejected objects are distributed uniformly in $f_{\rm cluster,W}$, such that half of these rejected objects have $f_{\rm cluster,W} > 0.3$.

Overall, the comparison between the AP catalog and previous non-PHAT M31 cluster catalogs shows good agreement with few conflicting classifications. A total of 733 unique, previously cataloged objects (both cluster and non-cluster classifications) match to AP candidates; 468 were previous (confirmed) cluster classifications, of which 404 were confirmed by the AP catalog. Within the PHAT survey footprint, we have increased the sample of confirmed clusters by a factor of $\sim$6 (from 468 to 2753). The HST-based AP catalog provides improvement in terms of catalog completeness and quality, and builds upon the firm foundation laid by these previous works. Commentary on individual classification differences can be found Section 3.8.4.

3.5.3 Integrated Photometry

We perform integrated aperture photometry for each of the AP catalog entries. Our photometry procedures are described in Paper I; we summarize the main ideas here, but refer the reader to that paper for additional details. We use the mean center and median radius of an object's merged classifications to define the position and radius ($R_{\rm ap}$) of the photometric aperture. The sky background is calculated within an annulus ten times the size of the photometric aperture, extending from 1.2 $R_{\rm ap}$ to $\sim$3.4 $R_{\rm ap}$. Photometric uncertainties are dominated by uncertainties in the sky background determination; this source of uncertainty is often ignored in extragalactic cluster photometry. Identical apertures (constant angular size) are employed across all six PHAT images. Aperture magnitudes for significant detections (S/N $\geq$ 3 with respect to the variation in the sky background) are listed for each photometric passband in Table 3.2; 3σ upper limits are provided for non-detections, and blank entries denote incomplete image coverage.

Table 3.3. Passband Photometric Quality Comparison for Cluster Sample

Passband	N(Detections)
F275W	1733 (62.9%)
F336W	2481 (90.1%)
F475W	2717 (98.7%)
F814W	1871 (68.0%)
F110W	1209 (43.9%)
F160W	1035 (37.6%)
F336W+F475W	2464 (89.5%)
F475W+F814W	1867 (67.8%)
F336W+F475W+F814W	1701 (61.8%)

We obtain photometric $R_{\rm eff}$ estimates by interpolating radial flux profiles. These values are then used to derive aperture corrections, which help account for cluster light that falls outside of the photometric aperture. We compare synthetic cluster input luminosities to measured magnitudes and find that this effect causes losses on the order of 0.1–0.3 mag. Corrections assume a King (1962) profile with a concentration ($c = R_{\rm tidal}/R_{\rm core}$) of 7, scaled to match the cluster's photometrically determined F475W $R_{\rm eff}$, then extrapolated to radii beyond $R_{\rm ap}$ to obtain a magnitude correction, $m_{\rm ApCor}$. Aperture corrections can be applied to raw aperture magnitudes to obtain total magnitude[7] estimates. These estimates accurately recover the photometry of synthetic clusters with no bias at brighter magnitudes (<19) and <0.2 mag bias for fainter clusters (see Sec. 4.2 in Paper I).

We summarize the photometric measurements in Table 3.3 where we tabulate the number of detections in each band, as well as the number of objects with detections spanning various combinations of photometric bands.

[7] $m_{\rm Total} = m_{\rm ap} + m_{\rm ApCor}$

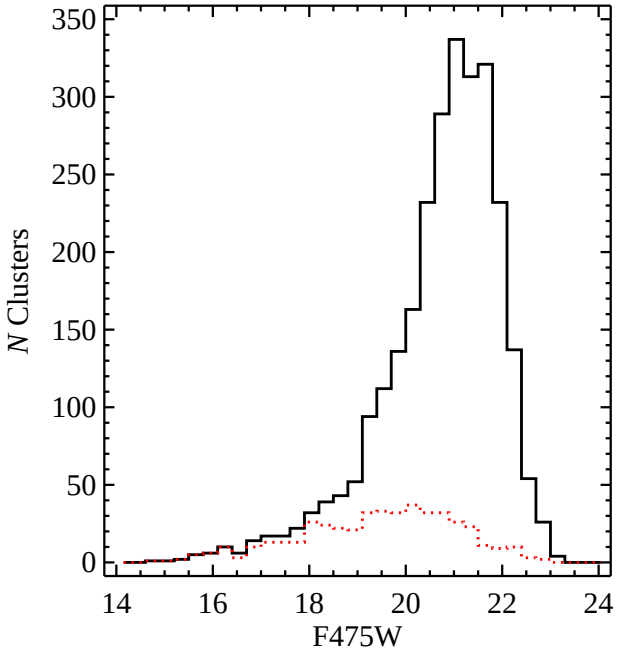

Figure 3.13 Histogram of F475W integrated magnitudes for 2,717 AP clusters (out of 2,753 total). The red dotted histogram represents the distribution of luminosities for 401 previously known clusters confirmed by the AP catalog (out of 404 total) that lie within the PHAT footprint, showing the vast improvement in cluster identification provided by the PHAT data.

We found that accounting for the presence of image artifacts was critical to obtain accurate cluster photometry in the F275W, F336W, and F110W filter bandpasses. Images in these three wavelengths proved problematic due to their small number of repeat observations and minimal spatial overlap between neighboring images, hindering typical artifact rejection techniques that require three or more exposures. Interpolating over pipeline-rejected pixels in the F110W images was relatively straightforward, however detecting and rejecting UV cosmic ray image artifacts was more difficult. We conservatively identify F275W and F336W artifacts by flagging bright, single-passband objects by comparing flux ratios of F275W, F336W, and F475W images. This method allows us to reject hundreds of artifacts that tend to bias measurements to brighter magnitudes, however we caution that some uncorrected artifacts may continue to affect our UV photometry.

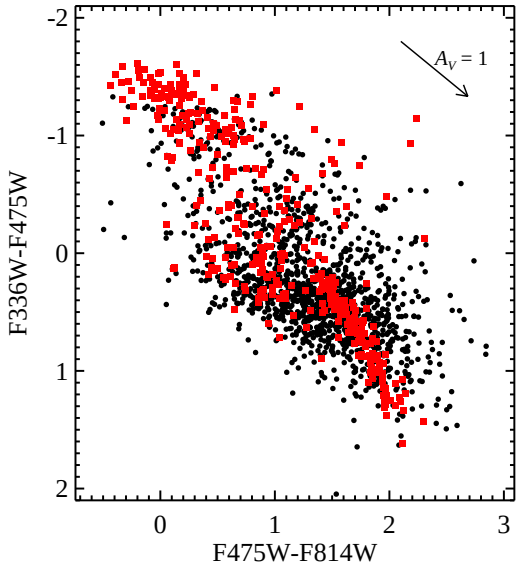

Figure 3.14 Color-color diagram of 1,701 clusters with F336W, F475W, and F814W photometric detections. The 378 clusters with F475W < 19.5 are distinguished as red squares. The color-color sequence of luminous globular clusters (see text in Sec. 3.5.3) is prominent in the sample of bright clusters.

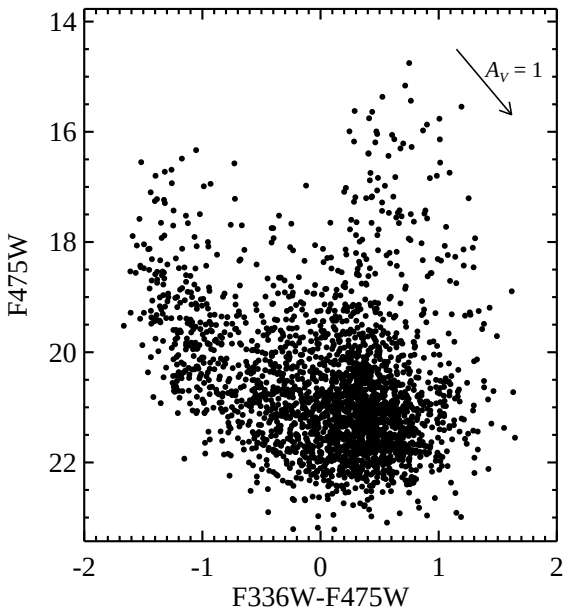

Figure 3.15 Color-magnitude diagram of 2,464 clusters with F336W and F475W photometric detections.

In Figure 3.13, we compare the distributions of F475W magnitudes for previously known (and confirmed) clusters in the PHAT footprint and the new AP catalog. The factor of ∼6 increase in the number of clusters shows the staggering improvement made possible by PHAT's high spatial resolution imaging. The ability to differentiate between single bright stars and compact clusters allows us to include fainter, less massive clusters in the AP catalog. Ground-based imaging limited previous efforts to clusters brighter than F475W ∼ 19.5, and while the HKC pushed that limit faint-ward, the amount of HST imaging available to those authors was significantly less than what is now available through PHAT.

Figures 3.14 and 3.15 show the color-color and color-magnitude distributions of AP clusters, providing a glimpse into the age composition of the catalog. While the clusters span a wide range of colors that reflect a diversity in ages, a dominant portion of the catalog lies within the following color and magnitude range: 20<F475W<22, 0<F336W−F475W<1, and 1<F475W−F814W<2. The specified region of color and magnitude parameter space points to a dominant population of $\sim 10^3$ $M_\odot$, ∼200–400 Myr old clusters that dominate the catalog by number, consistent with the age distribution found for the Year 1 sample (Fouesneau et al. 2014). This population dominates the cluster catalog because it represents a relatively large linear age range (leading to large number of clusters for a near constant formation history) where catalog completeness is still relatively high (50% complete to ∼1,000 $M_\odot$ at 300 Myr). We note that the large color dispersion shown in Figure 3.14 agrees with predictions of stochastically-sampled cluster models (see Fig. 4 in Fouesneau et al. 2014). In addition, the vertical sequence spanning 15<F475W<19 with a color range of 0.2<F336W−F475W<1.3 in Figure 3.15 represents the old (10-14 Gyr), massive ($> 10^5$ $M_\odot$) globular clusters. These massive, luminous systems also form a well-defined sequence of bright clusters in Figure 3.14, running from (F475W−F814W, F336W−F475W) coordinates of approximately (1.4,0.2) to (2.1,1.3), corresponding to a metallicity sequence running from -2.5<[Fe/H]<0.0 for these systems.

We fit luminosity functions to the cluster photometry using a simple power law ($N \propto L^{-\alpha_L}$); we plot the results in Figure 3.16. Notably, when we remove objects that lie within the previously defined bulge region (see Section 3.5.1), we find that luminosity functions steepen significantly. As we argued in Paper I, old massive globular clusters dominate

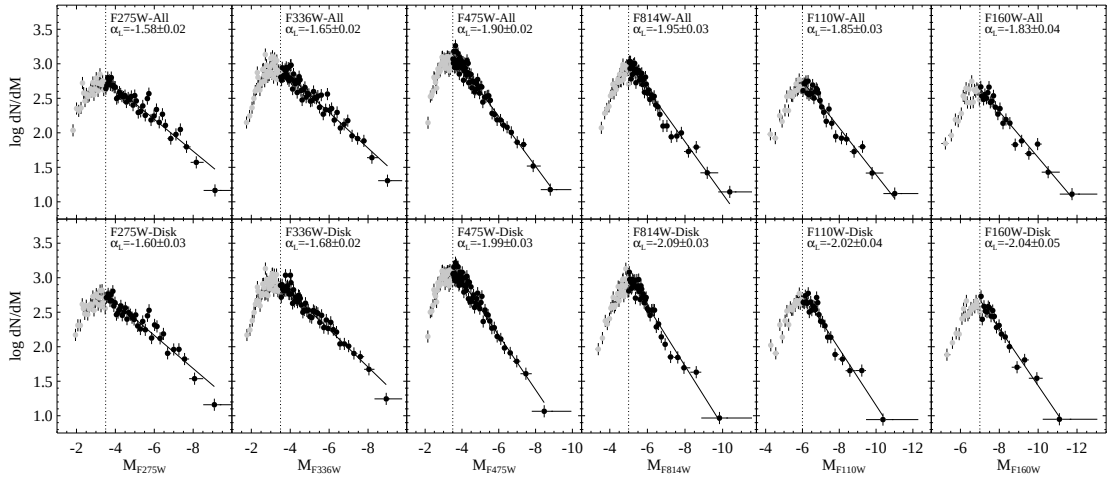

Figure 3.16 Luminosity functions across six PHAT passbands. Each plotted point represents an equal number of clusters ($N = 25$) and linear fits are made to points above the adopted completeness limit (dotted line). Top panels show results for the full AP cluster sample, while the bottom plots constrain the sample to those objects that lie within the disk, outside the inner bulge.

the bright end of the luminosity function; removing these objects, which reside primarily in the galaxy's bulge, allows us to examine the luminosity function behavior of younger ($\lesssim$3 Gyr) cluster populations. The observed population-dependent variations in luminosity function indices affirm that factors such as the underlying cluster formation history, the intrinsic cluster mass function, and the stochastic conversion from mass to luminosity for less massive clusters all play a role in determining the overall distribution of cluster luminosities. Untangling these various effects for the PHAT sample is possible through direct age and mass determinations of the individual clusters; we will perform this analysis as part of future work (Beerman et al., in prep.).

3.6 Discussion

3.6.1 Comparing the M31 Cluster Catalog to Extragalactic and Galactic Samples

To place the PHAT catalog of M31 star clusters into context, first we compare the luminosity distribution of our sample to those from three nearby star-forming galaxies: M83 (Bastian

et al. 2012b), M33 (San Roman et al. 2010), and the LMC (Hunter et al. 2003; Popescu et al. 2012). We choose these three galaxies because they are well-known extragalactic cluster targets that have publicly-available cluster catalogs; we compare our sample to the much more heterogenous Milky Way catalog in the next subsection. We compare the luminosity distributions of each sample in the left panel of Figure 3.17, where we convert from PHAT's F475W to V-band apparent magnitudes using the following empirical relation:

$$m_V = m_{\rm F475W} - 0.363(m_{\rm F475W} - m_{\rm F814W}) - 0.111. \quad (3.6)$$

Completeness limits for the three samples scale as a function of distance: M83 has the brightest completeness limit at $M_V \sim -6$, followed by M33's limit at $M_V \sim -5.5$, and the LMC's limit at $M_V \sim -4.5$. The M31 detection limit of $M_V \sim -3.5$ leads to the inclusion of many more clusters, particularly those of moderate mass and intermediate ages: 10^3–10^4 $M_\odot$ between 100 Myr and 1-3 Gyr (Fouesneau et al. 2014). As a result, the PHAT sample contains ∼3 times more clusters than any of the other extragalactic samples compared here.

At bright magnitudes ($M_V < -6$) where all four cluster samples are complete, we can compare the number of luminous blue ($B - V < 0.5$, or equivalently F475W−F814W < 1.1) clusters in each sample. This provides a first-order comparison of the young cluster populations captured by the catalogs of our set of comparison galaxies. We show in the right panel of Figure 3.17 that the M83 catalog includes the largest number of blue clusters, followed in order by M33, the LMC, and M31. Differences in the star formation rate (SFR) for the galaxy regions surveyed explain the differences observed in blue cluster populations. The cluster sample from the starburst galaxy M83 corresponds to a SFR of 1.3 $M_\odot$ yr^{-1} (coverage fraction of 2/5 applied to galaxy-wide SFR of 3.3 $M_\odot$ yr^{-1}; Boissier et al. 2005), while M33's SFR is 0.45 $M_\odot$ yr^{-1} (Verley et al. 2009) and the LMC's SFR is 0.25 $M_\odot$ yr^{-1} (Whitney et al. 2008). Within the PHAT footprint, the current SFR is much lower at ∼0.1 $M_\odot$ yr^{-1} (coverage fraction of 1/3 applied to galaxy-wide SFR of 0.25 $M_\odot$ yr^{-1}; Ford et al. 2013). Larger SFRs correlate with larger numbers of blue clusters; the relatively low number of luminous blue clusters in the AP catalog are a consequence of M31's relative quiescent SFR.

Next, we compare our PHAT cluster catalog to the sample of known Galactic clusters.

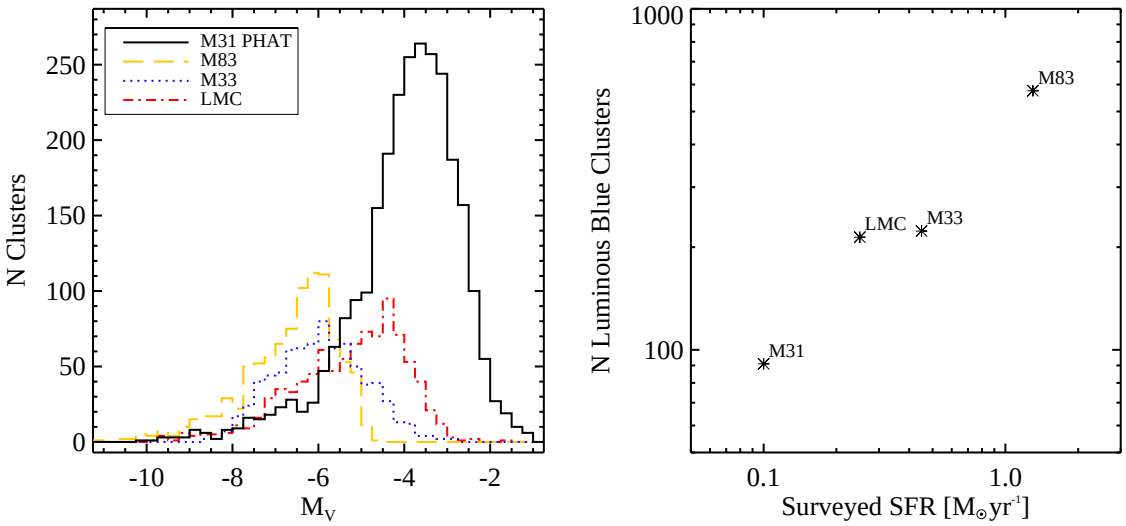

Figure 3.17 Left: A comparison of luminosity functions for four extragalactic samples of star clusters: M31 (this work), M83 (Bastian et al. 2012b), M33 (San Roman et al. 2010), and the LMC (Hunter et al. 2003; Popescu et al. 2012). This plot shows the relative difference in detection limits and total number of clusters for each catalog. Right: A comparison of the number of luminous blue clusters ($M_V < -6$, $B - V < 0.5$) in each galaxy sample as a function of SFR for the region that was surveyed in each galaxy. This plot shows that the AP sample includes fewer luminous blue clusters due to the relatively low SFR found in the PHAT survey footprint.

Without question, observations of Milky Way star clusters provide rich, detailed datasets for individual clusters and their member stars that cannot be matched in an extragalactic setting. The ability to measure star-by-star proper motions, detect and resolve stars down to the hydrogen burning limit, and efficiently obtain detailed abundance information through spectroscopy of individual members are all major advantages of studying clusters in the Galaxy. However, it is interesting to explore how Galactic cluster samples compare on galaxy-integrated scales. Our Sun's position within the disk of the Milky Way, surrounded by obscuring gas and dust along the Galactic mid-plane, does not provide the optimal vantage point for observing the distribution of clusters throughout our galaxy. In fact, we argue below that extragalactic samples provide a better assessment of overall cluster populations, due to the uniformity of selection and the ability to survey a wide range of galactic environments.

The recent catalog of Milky Way clusters by Kharchenko et al. (2013) contains 2547 clusters[8], similar to the number of entries in the PHAT cluster catalog. But while the sample sizes are comparable, the uniformity and selection function of the Milky Way clusters differ significantly from the AP clusters in M31. The sample of Milky Way clusters is compiled from a heterogenous set of literature sources, including earlier work of (Dias et al. 2002), leading to an ill-defined selection function and catalog completeness that is difficult to characterize. Assuming a constant surface density of clusters, Kharchenko et al. (2013) suggest that the sample is complete to a radius of $\sim$1.8 kpc around the sun thus covering an area of $\sim$10 kpc^2. Not only is this area more than an order of magnitude smaller than the physical region covered by PHAT, but the surveyed region of Galaxy is limited to the Solar neighborhood. Most of the area within 1.8 kpc of the Sun lies within a Galactic inter-arm region, limiting the amount of on-going star formation and range of environments one can study.

According to estimates compiled in a recent review by Portegies Zwart et al. (2010) (based on the sample of Dias et al. 2002), the young (excluding globular clusters) Milky Way cluster sample includes objects that range in mass from 25 $M_\odot$ to 5×10^4 $M_\odot$. Within

[8]This total excludes associations, moving groups, and remnant cluster classifications from the catalog's 3006 overall entries.

a radius of 1.8 kpc, the complete cluster sample includes a mass range that varies over <3 orders of magnitude, up to 4000 $M_{\odot}$, the mass of the Orion Nebula Cluster. The proximity of the Milky Way clusters allows for the inclusion of low mass objects that remain undetected in M31, however the accurate understanding of mass completeness and catalog selection for PHAT, along with the number and variety of clusters included, makes the AP catalog the best available resource for a wide range of cluster science studies: cluster dissolution, mass functions, cluster formation efficiency, and how cluster properties vary with environment.

3.6.2 Catalog Differences: Year 1 & AP

The cluster definition we use for the AP catalog, as described in Section 3.1.1, is more liberal than the one used in our previous Year 1 catalog. In Paper I, we excluded three categories of candidate clusters that we do not explicitly reject from the AP catalog:

1. *Loose Associations* — Defined by their lack of centrally concentrated stars, these objects are likely to be gravitationally unbound due to their large spatial extents and low stellar densities, and were therefore rejected from inclusion in the Year 1 catalog. The AP search yielded many high-significance candidates that were not identified during the Year 1 effort.

2. *Emission Line Regions* — Compact, high surface brightness HII regions show up prominently in F475W imaging ([OIII] and Hβ emission lines lie within the F475W bandpass) and tends to enhance the visual appearance of associated clusters. While line emission on its own does not provide explicit evidence for or against the presence of a cluster (non-cluster HII regions and line emitting clusters both exist), we find that cluster candidates associated with emission line flux are accepted more frequently into the AP catalog than by the expert-based Year 1 search. We document this tendency because it reveals a possible systematic affecting catalog completeness that is not captured by our synthetic cluster tests: low mass clusters that produce line emission may be systematically overrepresented in the AP catalog with respect to the completeness function derived in Section 3.4.

3. *Small Clusters* — While we emphasized a liberal, inclusive approach to cluster identification in Paper I, small candidate clusters were often discarded, with a loosely-defined limit requiring 3-4 spatially correlated stars to trigger inclusion in the catalog. For the AP search, no star count limit was ever discussed.

These three categories of objects represent systematic differences between the Year 1 and AP catalogs. Of these three, the loose associations represent the most conspicuous difference: the number of bright blue objects (F336W−F475W < -0.5 and F475W < 19.75) identified within the Year 1 footprint more than doubled, from 15 to 35 clusters, many of which appear extended and poorly concentrated. In an effort to clearly identify these uncertain and controversial AP clusters, we flag objects that match the following criteria as possible associations: bright (F475W < 19.75), blue (F336W−F475W < -0.5), and spatially extended. A cluster is characterized as spatially extended either through its light profile, according to its half-light radius, or its profile of resolved main sequence stars, according to the radius that contains 60% of the cluster's main sequence stars ($R_{0.6N(\mathrm{MS})}$). We adopt the following criteria for spatial extension: $R_{\mathrm{eff}} > 1.05$ arcsec (4 pc), or $R_{0.6N(\mathrm{MS})} > 0.5R_{\mathrm{ap}}$ for stars with F475W−F814W < 1 and F475W < 24. The combined color, magnitude, and spatial extension criteria identify 64 association-like objects; flags identifying these objects are included in Table 3.2. These extended candidates are the most likely examples of regions hosting spatially correlated star formation, but where the stars may not have ever been gravitationally bound to one another. As such, one should carefully evaluate the possibility that these candidates may not be the young progenitors of the older clusters we identify in this catalog.

3.7 Summary

We presented our methodology for transforming crowd-sourced effort into cluster catalogs for the AP-based analysis of the PHAT survey data. We show the validity of our crowd-sourced cluster identification methodology and show good consistency between our results and expert-derived by-eye catalogs. In addition, we present a thorough analysis of the resulting completeness characteristics of our cluster catalog, an essential component to any

study of galaxy-wide star cluster populations. Our completeness tests demonstrate that our PHAT cluster catalog is mass-limited and 50% complete to $\sim$500 $M_{\odot}$ up to an age of 100 Myr, at which point the catalog becomes luminosity-limited at F475W $\sim$21.5.

The final cluster catalog includes 2753 entries, spanning more than three orders of magnitude in F475W luminosity. Our use of HST imaging provides access to systems spanning the range from massive globular clusters to low-mass ($< 10^3$ $M_{\odot}$) clusters in the disk, similar to Milky Way open clusters. Analysis of this sample provides a unique and unmatched opportunity to obtain a comprehensive understanding of star cluster populations within a large spiral galaxy. The AP catalog serves as the definitive base data product that will enable an array of stellar cluster studies within M31.

3.8 Supplementary Information

3.8.1 Catalog Construction Procedure

AP catalog construction occurs in two phases: merging identifications on an image-by-image basis, followed by the merging of per-image catalogs into a single survey-wide catalog. Throughout this description, we use the terms "click" and "identification" interchangeably to represent image markers placed by AP volunteers. We begin by describing the first phase, which consists of three steps:

1. *Create Candidate List* — From the set of all cluster and galaxy identifications recorded for a given image, we construct a list of initial candidate objects by grouping center positions using a matching radius of 20 pixels (equivalent to 1 arcsec or 3.81 pc). Our choice of matching radius was tuned such that clicks representing the same object were merged together, but distinct neighboring objects were not merged. We observed that the positioning of marker centers are quite precise; the distribution of user-determined centers for well-defined image features can be described as a 2D Gaussian with σ=2 pixels (equivalent to 0.1 arcsec or 0.4 pc). We iterate through the list of identifications, sorted from smallest to largest radius under the assumption that small radii identifications will have the most accurate center positions. Each resulting

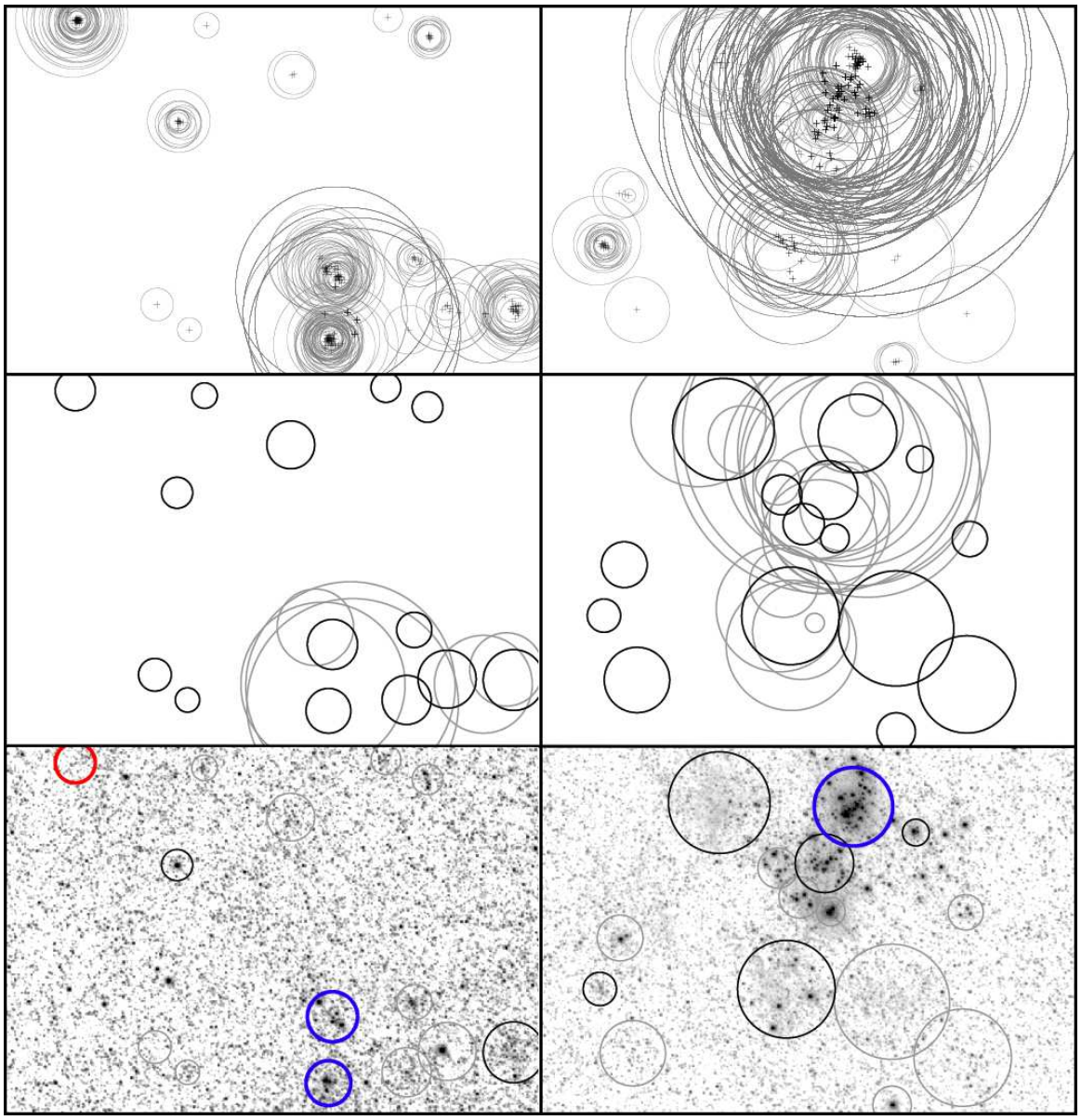

Figure 3.18 Catalog construction examples, featuring identifications from B02-F11_22 (left) and B06-F16_22 (right). Top panels: We plot all individual cluster and galaxy markings along with their centers. Middle panels: We plot all initial object candidates that result from the grouping of center positions. Gray circles represent candidates that were pruned, while black circles represent candidates that go on to become final catalog entries. Bottom panels: Each final candidate is shown, color coded by its final status: clusters (thick blue), galaxies (thick red), ancillary candidates with $f_{\mathrm{clst+gal}} \geq 0.1$ (thin black), and low significance candidates with $f_{\mathrm{clst+gal}} < 0.1$ (thin gray).

candidate consists of a central position and radius, represented by the mean X and Y image coordinates and the median radius of each set of merged clicks.

2. *Prune Candidates* — Next we prune duplicate objects from the initial set of candidate objects. Here, a duplicate object is a candidate that corresponds to the same image feature as another in the initial list, but the component clicks were not merged during the previous step. This process begins by iterating through the initial candidate list in order of decreasing $f_{\mathrm{clst+gal}}$. For each iteration, we define the candidate in question as the primary object, and search for secondary objects, which are any other initial candidates whose circular boundary encloses the primary's center. If we identify any secondary objects with a $f_{\mathrm{clst+gal}}$ less than that of the primary, the secondary is dropped from the candidate list. If a secondary candidate has a higher $f_{\mathrm{clst+gal}}$ than the primary, the primary is dropped from the candidate list. Once we've iterated through all initial candidates, the result of this pruning procedure is a list of spatially-unique candidates.

3. *Re-associate Identifications with Final Candidates* — To calculate final hit-rate statistics for each surviving candidate, we identify all original identifications where the candidate aperture encloses the identification center and vice versa and use these clicks to calculate $f_{\mathrm{clst+gal}}$, f_{cluster}, and f_{galaxy}values. However, candidates retain their previous center and radius values. Finally, we remove any candidates with only one associated identification (i.e., single click candidates), while remaining multi-click candidates go on to join the final per-image catalog.

We present two image examples that show how our catalog construction algorithm works. The top row of Figure 3.18 shows all object identifications and their associated centers for each image. The second row shows the full list of merged candidates that result from the first step described above, where those that survive the pruning process are highlighted in bold. Finally, we show the final list of candidates that survive the candidate pruning and subsequent single-click cut overlaid on top of the field's single-band F475W image, where the most significant detections ($f_{\mathrm{clst+gal}}$ is $\geq$0.1) are shown in red. The left column of Figure

3.18 shows the B02-F11_22 sub-image, a field that consists primarily of well-separated, well-defined candidates. The right column shows B06-F16_22, which highlights a challenging case with many non-unique, overlapping feature identifications.

The B06-F16_22 image example presents a particularly informative example of our cataloging algorithm in action. The transition from the raw identification data in the top panel to the initial candidate list in the middle panel shows that our methodology for merging clicks (using a small 20 pixel matching radius) is quite conservative, insuring that nearby objects are not incorrectly combined. Next, this initial candidate list is pruned to remove true duplications, cutting the first set of candidates down to those plotted in black in the middle panel. This operation takes the significance of each candidate into account (according to $f_{\mathrm{clst+gal}}$ scores, reflecting total numbers of clicks), and yields a final list of objects that are spatially unique. Identifications associated with the dropped duplicate candidates are not discarded, as most are re-associated during the final step of per-image processing. Finally, the bottom panel shows the output of catalog processing, showing reasonable results even for this complex set of inputs. While the low and moderate significance identifications (gray and black circles, respectively) are not included in the AP catalog published in Table 3.2, these objects are all recoverable due to their inclusion in the publicly available ancillary catalog presented in Section 3.8.3.

The primary AP base data product is produced in the second phase of the construction process: merging per-image catalogs into a final survey-wide catalog. We perform this merge in a two-step process:

1. *Match Per-Image Candidates* — We compile a list of all sub-image catalog entries, and iterate through each entry in order from highest $f_{\mathrm{clst+gal}}$ to lowest. During each iteration, we define the candidate in question as the primary object, and search for secondary objects, which are any other candidates whose circular boundary encloses the primary's center and vice versa. If we identify any secondary objects, these matches are immediately removed from the list. When complete, the resulting list of surviving objects represents our final list of spatially-unique catalog entries.

2. *Merge Candidate Properties* — To determine the properties of each final catalog ob-

ject, consider each entry and its set of associated secondary entries. From this set of per-image objects, identify those that lie completely within the bounds of their host sub-image (whose radius is less than the distance to the closest image edge) and merge their positions (in RA/Dec coordinates) and radii using the mean of their individual values, and assign final $f_{\rm clst+gal}$, $f_{\rm cluster}$, and $f_{\rm galaxy}$ values and using a mean weighted by the number of total sub-image views. Excluding objects that do not fall completely within their host image allows us to limit the influence of edge effects and biases on the final cataloged properties. If none of the merged per-image entries pass this edge criteria, we adopt the properties of the entry that lies furthest from an image edge.

3.8.2 PHAT Background Galaxy Catalog

To define an AP galaxy sample, we base our selection on a combination of $f_{\rm clst+gal}$ and $f_{\rm galaxy}$ criteria. Utilizing the $f_{\rm clst+gal}$ metric allows for better recovery of moderate and high $f_{\rm clst+gal}$ objects with $f_{\rm galaxy}$ scores that lie on the tail (0.3–0.8) of the distribution. Adopting an $f_{\rm galaxy}$ threshold of 0.3, as discussed in Section 3.3, we construct a completeness curve for the galaxy sample to choose an appropriate $f_{\rm clst+gal}$ cutoff. Similar to our cluster analysis, we compare the AP sample to the Year 1 galaxy sample. The Year 1 galaxy sample was not a focused effort to identify all possible galaxies, therefore we do not categorize AP identifications that do not match Year 1 galaxies as contaminants, but study the behavior of the relative completeness fraction of these objects in a similar way. We do not pursue the application of user weights for these results, but derive a single unweighted curve presented in Figure 3.19.

We observe a transition in the behavior of the completeness curve at a Year 1 completeness of $\sim$0.67. The slope of the completeness curve becomes steeper; quantitatively, this transition represents the point of diminishing returns, where more non-Year 1 objects are being added to the sample than previously identified Year 1 objects. We choose this transition point as a suitable limit for catalog inclusion. Therefore, we define the AP galaxy sample using the following selection criteria:

$$f_{\rm clst+gal} > 0.37 \text{ AND } f_{\rm galaxy} \geq 0.3 \tag{3.7}$$

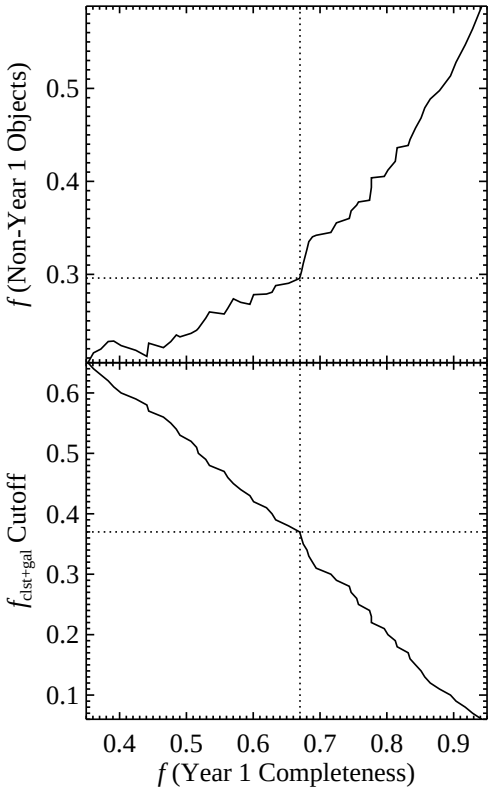

Figure 3.19 Top: Completeness versus contamination curve for galaxy sample. Bottom: Completeness versus $f_{\mathrm{clst+gal}}$ cutoff values.

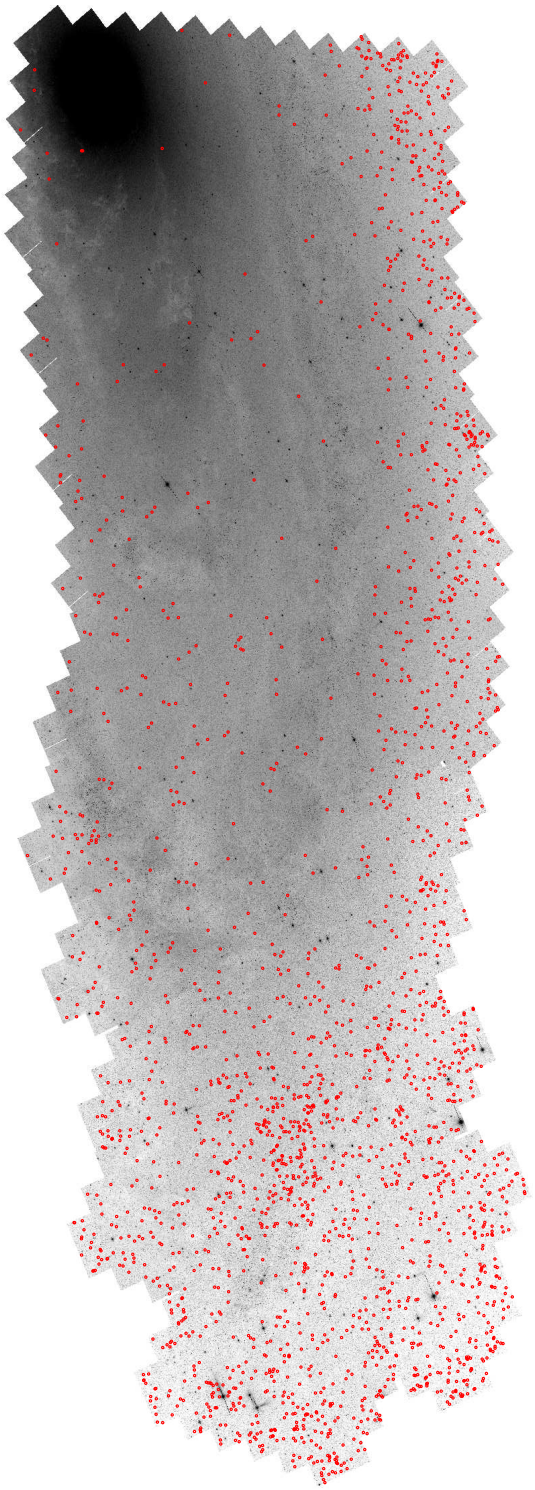

Figure 3.20 Spatial distribution of AP background galaxy catalog overlaid on the PHAT survey-wide F475W image.

Table 3.4. AP Background Galaxy Catalog

AP ID	RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$f_{\rm clst+gal}$	$f_{\rm galaxy}$	F814W$_{\rm ap}$	σ
8	11.447226	42.268672	2.75	1.0000	0.9884	18.34	0.05
10	11.922144	42.102526	3.53	1.0000	1.0000	18.53	0.11
20	11.911096	42.076717	2.00	0.9902	0.9604	20.07	0.29
21	11.585498	41.726941	3.82	0.9901	0.9900	16.04	0.02
22	11.270065	41.312829	2.38	0.9901	0.9500	18.67	0.11

Note. — Table 3.4 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

These criteria select a sample of 2,270 background galaxies. The catalog is presented in Table 3.4 and their spatial distribution is shown in Figure 3.20.

3.8.3 Ancillary Catalog Data

Additional Candidate Catalog: In addition to the AP clusters presented in Table 3.2 and background galaxies presented in Table 3.4, we present Table 3.5 containing 8775 other candidate object identifications with $f_{\rm clst+gal} \geq 0.1$ that were not included in either of the other data tables.

Synthetic Cluster Results: We present cluster-by-cluster synthetic recovery results in Table 3.6 to allow for custom completeness analyses. The table includes input cluster parameter information (i.e., age, mass, $R_{\rm eff}$, position, etc.) as well as AP catalog metadata.

3.8.4 Commentary on Existing Cluster Catalogs

To supplement the broad comparison to previously published cluster catalogs that we presented in Section 3.5.2, here we provide commentary on conflicting M31 object classifications. We summarize these differences numerically in Table 3.7 and present classification updates in Table 3.8.

Table 3.5. AP Ancillary Catalog

AP ID	RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$R_{\rm eff}$ ($''$)	$m_{\rm ApCor}$[a]	F275W$_{\rm ap}$	σ	F336W$_{\rm ap}$	σ	F475W$_{\rm ap}$	σ
$f_{\rm clst+gal}$	$f_{\rm galaxy}$	$f_{\rm cluster,W}$	PC ID	Alt ID	Flags	F814W$_{\rm ap}$	σ	F110W$_{\rm ap}$	σ	F160W$_{\rm ap}$	σ
1706	11.393493	41.774981	1.25	0.36	-0.02	>21.62	···	21.23	0.24	21.21	0.20
0.3415	0.0000	0.5252	···	···	···	>20.08	···	>19.39	···	>18.73	···
2073	11.701786	41.963523	1.09	0.46	-0.11	23.85	0.39	23.21	0.07	22.57	0.16
0.7738	0.1538	0.6389	···	···	···	20.95	0.33	20.09	0.34	19.16	0.31
2149	11.133017	41.395088	2.51	1.69	-0.38	16.76	0.04	17.28	0.15	18.80	0.04
0.5294	0.0000	0.4364	···	···	···	>19.62	···	>18.12	···	>16.93	···
2486	11.554857	41.873578	1.81	0.23	-0.00	19.33	0.11	19.23	0.11	>20.12	···
0.3372	0.0345	0.3793	···	···	···	>19.57	···	>18.62	···	>17.77	···
2532	10.915584	41.487991	2.08	1.08	-0.21	21.71	0.34	20.63	0.17	20.06	0.16
0.7263	0.2754	0.5502	···	SK070A	···	18.14	0.09	17.46	0.11	16.78	0.22

Note. — Table 3.5 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content. Three-sigma upper limits are denoted by a ">" symbol. PC ID and Alt ID refer to cluster identifiers from the Year 1 catalog and other literature sources, respectively.

[a]Aperture Corrections are provided such that $m_{\rm Total} = m_{\rm ap} + m_{\rm ApCor}$.

Table 3.6. Synthetic Cluster Results

ID	log(Mass/$M_\odot$)	log(Age/yr)	Z	A_V	$R_{\rm eff,in}$ ($''$)	F475W$_{\rm in}$	F814W$_{\rm in}$	F475W$_{-3,\rm in}$
RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$f_{\rm clst+gal}$	$f_{\rm galaxy}$	$f_{\rm cluster,W}$	N(MS)	N(RGB)	Detected
1	3.17	7.30	0.019	1.612	0.319	20.75	20.06	21.07
11.012636	41.181335	1.39	0.9333	0.0000	0.9997	169	374	T
2	4.51	9.60	0.019	0.253	1.397	20.19	18.34	20.21
11.003787	41.184849	1.99	0.7326	0.0159	0.9386	192	418	T
3	2.17	8.10	0.019	0.230	0.467	21.18	21.18	22.36
10.985614	41.192447	1.42	0.5222	0.0000	0.6968	145	468	T
4	3.92	10.05	0.0001	1.366	0.343	22.92	21.02	23.22
11.004518	41.190121	1.16	0.1786	0.0000	0.1575	200	398	F
5	4.47	10.05	0.004	0.370	0.683	20.72	18.82	20.87
10.990636	41.195372	1.47	0.7126	0.0806	0.9198	199	497	T

Note. — Table 3.6 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content. N(MS) and N(RGB) values are evaluated per search image, as defined in Section 3.4.

Table 3.7. Summary of Existing Cluster Catalog Classifications and Revisions

Catalog	Clusters	Candidates[a]	Non-Cluster[b]
	# Accepted as AP Cluster (# Rejected as Not AP Cluster)		
Year 1	532 (69)	95 (142)	39
RBC	232 (2)	28 (18)	40
Caldwell	232 (10)	···	8
HKC	156 (57)	···	···

[a]The "candidate" classification refers to PHAT Year 1 possible cluster table and RBC flags 2 and 3.

[b]The "non-cluster" classification refers to galaxy, star, HII region, and other non-cluster classifications.

Table 3.8. RBC Flag Revisions and Commentary

AP ID	RBC Name	New Flag	Old Flag	Comments
55	SK102A	1	6	···
68	SK213B	1	2	···
77	SK182B	1	6	···
89	M065	1	2	···
91	M004	1	6	···

Note. — Table 3.8 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

Table 3.9. Archival AP Cluster Catalog

AAP ID	RA (J2000)	Dec (J2000)	R_{ap} ($''$)	$f_{\rm clst+gal}$	$f_{\rm galaxy}$	$f_{\rm cluster,W}$	Alt ID
1	10.522610	41.435868	2.42	0.9903	0.0294	0.9730	B069-G132
2	10.541597	40.907603	2.82	0.9804	0.0000	0.9905	···
8	10.509294	40.896004	1.97	0.9800	0.0408	0.9908	SK041A
9	10.753904	41.656852	1.42	0.9126	0.0000	0.9915	···
11	10.521275	40.885136	1.66	0.9712	0.0594	0.9828	···

Note. — Table 3.9 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

3.8.5 Archival Image Search and Catalogs

As part of the second round of AP data collection, we included additional, non-PHAT ACS images obtained from the HST archive. These images were obtained by a program (PID: 10273, PI: Crotts) that observed four contiguous stripes within M31, each composed of 8 side-by-side ACS fields. Please see Figure 3.2 for the locations of the strips with respect to the PHAT survey footprint. This program utilized a F555W, F814W filter combination and exposure times that are shorter than those of PHAT: F555W varying between 81 and 413 sec, F814W varying between 457 and 502 sec. We divided each of the 32 ACS images into 54 sub-images, yielding a total of 1,728 search images.

Following the same catalog construction procedures (see Section 3.3) and selection criteria (see Section 3.5.1) used for the PHAT classification data, we construct catalogs of clusters and background galaxies. We present the cluster sample in Table 3.9, the background galaxy sample in Table 3.10, and compile an ancillary sample of all other identifications with $f_{\rm clst+gal} \geq 0.1$ in Table 3.11.

Table 3.10. Archival AP Background Galaxy Catalog

AAP ID	RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$f_{\rm clst+gal}$	$f_{\rm galaxy}$
3	10.846248	41.040394	1.99	0.9804	0.8900
4	10.536274	41.444508	5.10	0.9802	0.9596
5	10.595225	40.953062	2.44	0.9802	0.9192
6	10.465582	41.411634	4.50	0.9802	0.9495
7	10.463831	41.416378	4.79	0.9802	0.9697

Note. — Table 3.10 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

Table 3.11. Archival AP Ancillary Catalog

AAP ID	RA (J2000)	Dec (J2000)	$R_{\rm ap}$ ($''$)	$f_{\rm clst+gal}$	$f_{\rm galaxy}$	$f_{\rm cluster,W}$	Alt ID
238	10.491200	41.439039	1.01	0.7788	0.2716	0.6317	KHM31-357
334	10.750010	41.001666	1.35	0.6923	0.1389	0.5875	KHM31-453
389	10.651615	41.552907	1.36	0.6400	0.2188	0.5784	···
399	10.546207	41.509361	1.41	0.6341	0.0385	0.6197	KHM31-367
400	10.950410	41.192429	1.36	0.6337	0.2031	0.5965	···

Note. — Table 3.11 is published in its entirety in the electronic edition of the *Astrophysical Journal*. A portion is shown here for guidance regarding its form and content.

Chapter 4

STAR CLUSTER FORMATION EFFICIENCY AND THE CLUSTERED FRACTION OF YOUNG STARS

We make spatially resolved measurements of star cluster formation efficiency, the fraction of young stars formed as members of long-lived star clusters (Γ), across the PHAT survey footprint in M31. We derive robust constraints for Andromeda's cluster and field populations over the last $\sim$300 Myr through color-magnitude diagram analysis of individually resolved stars. We find that 3–6% of young stars (10–100 Myr old) are star cluster members. This fraction varies across the galaxy disk and tends to increase in correlation with total gas and star formation rate surface densities ($\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$). These Γ measurements expand the range of well-studied galactic environments, providing high quality constraints in an HI-dominated, low intensity star forming environment. The observed trends with $\Sigma_{\rm SFR}$ are broadly consistent with previous evidence for environmentally-dependent cluster formation efficiency derived at galaxy-integrated spatial scales. However, we find better agreement between observations and theoretical models if we account for known variations in gas depletion times ($\tau_{\rm dep}$) when calculating Γ predictions. This modification accounts for the qualitative shift in star formation behavior when transitioning from a predominately molecular gas phase to an atomic-dominated interstellar medium, yielding good agreement between our Γ observations and model predictions in most cases. We also demonstrate that we can explain Γ measurements of starburst systems within the same theoretical framework.

4.1 Introduction

The clustering behavior of stars is a direct, observable result of star formation physics. At the onset of star formation, young embedded stars inherit the highly structured spatial distribution of the molecular gas from which they form. The newly formed stars soon decouple from the gas, however, as it is expelled by stellar feedback processes. Because star formation is an inefficient process ($\sim$1% per free-fall time; Krumholz et al. 2012), gas dis-

persal removes most of a region's binding gravitational potential. This results in most stars expanding and dispersing, creating stellar associations and complexes with characteristic sizes of tens to hundreds of parsecs. In some cases, however, the concentration of stellar mass is high enough such that collections of stars remain gravitationally bound and tightly clustered beyond the initial gas embedded phase, creating long-lived ($\gtrsim$50–100 Myr) star clusters that we observe today.

While the shape of the stellar initial mass function (IMF) and measurements of the formation efficiency of individual stars are often used to test theoretical descriptions of star formation, observations of star clusters also provide a means to constrain these processes. Star cluster formation depends on the complex interplay of: star formation efficiency, which dictates how much of the gas reservoir is transformed into possible cluster members; feedback processes, which dictate the transition of stars from a gas-rich to gas-poor gravitational potential; and the energetics of the natal environment, which dictate the kinematics of stellar and gaseous components within the star forming region. As a result, accurately reproducing the observed behavior of star clusters, and young stellar distributions generally, is a key challenge for any theoretical star formation model.

We explore an important observational metric of star cluster formation in this work: star cluster formation efficiency, which is the fraction of stars born in long-lived star clusters (Γ; Bastian 2008, 2013). This quantity directly relates cluster formation to total star formation activity. Past measurements of cluster formation efficiency were obtained on galaxy-integrated scales for a wide range of galaxies (e.g., Larsen & Richtler 2000; Goddard et al. 2010; Silva-Villa & Larsen 2011; Adamo et al. 2011; Cook et al. 2012). These studies provided evidence that Γ varies systematically as a function of star forming environment, ranging from a few percent for galaxies with low star formation intensities (quantified using $\Sigma_{\rm SFR}$) up to $\sim$50% for high intensity galaxy mergers. Recently, studies have begun to explore Γ with increasing detail, performing spatially resolved analyses to better investigate the environmental dependence of cluster formation (Silva-Villa et al. 2013; Ryon et al. 2014; Adamo et al. 2015).

In addition to these observational studies, work from Kruijssen (2012) took an important first theoretical step in modeling and predicting Γ behavior. Motivated by the theoretical

work of Elmegreen (2008) and star formation simulations by Bonnell et al. (2008), Kruijssen (2012) presents a framework to predict Γ based on the idea that long-lived star clusters emerge from regions with high star formation efficiency. In this model, the densest portions of hierarchically-structured molecular clouds attain high star formation efficiencies because while the star formation efficiency remains constant per free-fall time (Elmegreen 2002), these regions progress through multiple short free-fall times. As a result, these regions become stellar-dominated before gas expulsion truncates star formation. Low gas fractions in these dense sub-regions prevent gas expulsion from dramatically changing the gravitational potential, leading to the formation of long-lived star clusters.

Here we measure star cluster formation efficiency across the Andromeda galaxy (M31) using data from the *Hubble Space Telescope* (HST) obtained by the Panchromatic Hubble Andromeda Treasury survey (PHAT; Dalcanton et al. 2012). M31 is an interesting target for a number of reasons. First, Andromeda hosts a low intensity star formation environment, characterized by low mean star formation rate surface densities ($\Sigma_{\rm SFR}$). In contrast to previous actively star forming targets, M31 provides a lower star formation intensity environment for studying Γ. Its global star formation rate (SFR) is only 0.7 $M_\odot$ yr^{-1}, with 0.3 $M_\odot$ yr^{-1} contained within the PHAT survey footprint that we analyze here (Lewis et al. 2015). Furthermore, star formation activity in M31 has a non-regular radial distribution, featuring a prominent long-lived star forming ring that contains a majority of the galaxy's star formation activity. This non-uniform spatial distribution produces distinct star forming environments that serve as interesting individual laboratories for observational study. M31's interstellar medium (ISM) also sets it apart from previous Γ analysis targets, as it is dominated by its atomic phase, unlike most massive star forming spiral galaxies that are molecular gas dominated. Finally, preliminary investigations show that Andromeda's cluster dissolution rate is low (Fouesneau et al. 2014), suggesting characteristic disruption timescales >100–300 Myr that leave its population of long-lived star clusters intact for study.

Our analysis in M31 benefits from a number of important advantages over previous extragalactic Γ studies. First, we use a robust catalog of 2753 clusters that were visually identified as part of the Andromeda Project citizen science project (Johnson et al. 2015b).

This cluster search was performed on uniform imaging from the PHAT survey, in which clusters appear as groupings of individually resolved member stars, reducing confusion and ambiguity in cluster identifications. In imaging of galaxies at larger distances (>1 Mpc), cluster members are blended together and cluster profiles are only marginally resolved, even with the resolving power of HST. Catalog completeness is well characterized and shows that the PHAT young cluster sample (<300 Myr old) is complete to 500–1000 $M_{\odot}$ (depending on age and galactic position), providing access to low mass clusters that are undetectable in most extragalactic surveys. Another critical benefit of resolving individual cluster member stars is the ability to derive cluster ages and masses by fitting color-magnitude diagrams (CMDs). This fitting provides stronger constraints than those obtained through multi-band SED fitting for young clusters, avoiding uncertainties caused by large stochastic variations in the integrated light of low mass clusters (see e.g., Fouesneau & Lançon 2010; Krumholz et al. 2015).

The benefits of studying cluster formation efficiency in M31 reach beyond the realm of cluster-specific observations. Star formation history (SFH) results derived from the PHAT observations of field star populations provide valuable spatially (∼100 pc scales) and temporally (Δ log Age/yr $\sim$ 0.1) resolved information about the total star formation activity across the disk of M31 (Lewis et al. 2015). These constraints are a considerable improvement over total SFR estimates made using luminosity-based methods (e.g., via Hα, FUV+24μm). In addition, the availability of HI and CO datasets that have high spatial resolution and sensitivity, allowing detailed characterization of the star forming ISM, even at low gas surface densities. These gas phase constraints provide rich ancillary information that allow us to map how differences in natal environments affect properties of emergent cluster populations.

In this paper, we take advantage of the superior quality of data provided by the PHAT survey to perform a high precision, spatially resolved investigation of Γ across a range of star forming environments in M31. Our work complements previous observational studies by providing a high quality anchor to extragalactic Γ measurements in more distant galaxies where the level of detail available with respect to characterizing both clusters and field populations is limited by available spatial resolution.

This paper is organized into six sections. We begin with a description of our observational

data in Section 4.2, followed by a presentation of star cluster and field star characterization analysis in Sections 4.3.1 and 4.3.2. We calculate Γ and its associated uncertainties in Section 4.4.1. In Section 4.4.2, we calculate theoretical predictions for Γ using the Kruijssen (2012) model and compare these to our observations. In Section 4.5, we compare our results to previous observations, discuss the validity of key assumptions, and follow-up on interesting aspects of our results and their broader implications. We finish with a summary of our work in Section 4.6. Throughout this study, we assume a distance modulus for M31 of 24.47 (785 kpc; McConnachie et al. 2005), for which 1 arcsec corresponds to a physical size of 3.81 pc.

4.1.1 Γ and Cluster Definitions

We begin our study of Γ by defining the measurement we pursue in this work, and in doing so clarify points of confusion that have arisen in the literature. The quantity of interest here is the fraction of stellar mass born in long-lived clusters relative to the total coeval stellar mass formed; we refer to this value as Γ following Bastian (2008). We rewrite the original definition (the ratio of the cluster formation rate to the total star formation rate, CFR/SFR) to clarify that this measurement is made over a specific time interval. We define:

$$\Gamma = \frac{M_{\rm cl}(t_1, t_2)}{M_{\rm tot}(t_1, t_2)}, \tag{4.1}$$

where $M_{\rm cl}$ represents the integrated cluster mass, $M_{\rm tot}$ represents the integrated total stellar mass, and (t_1, t_2) represents the the age interval over which the masses are integrated. This ratio is equivalently referred to as the cluster formation efficiency.

Investigators typically restrict measurements of Γ to young ages ($t_2 < 10$–100 Myr) due to a number of considerations. First, SFR estimates obtained from broadband indicators (e.g., Hα, FUV+24μm) or from fitting shallow color-magnitude diagrams only provide constraints at young ages. Second, estimates of total cluster mass are increasingly reliant on mass function extrapolations and small number statistics with increasing age due to evolutionary fading and worsening mass completeness limits. Finally, mass loss and cluster dissolution are smallest at young ages, while at older ages observations of Γ may no longer reflect its initial value. With the PHAT data, we can measure Γ out to older ages ($t_2 = 300$ Myr) thanks to deep optical imaging that allows detection of main sequence turnoffs in both

clusters and the field at $m_{\rm turnoff} \sim 3\ M_{\odot}$.

A number of different factors control the appropriate choice of the young age limit, t_1. In principle, t_1 could extend to ages as low as 1-3 Myr, when clusters transition out of their embedded phase and become detectable in optically-selected samples. In practice, however, it is difficult to differentiate between long-lived, gravitationally bound star clusters and unbound, expanding stellar associations that are still compact at young ages. The two classes of objects are both formed in dense configurations that are difficult to distinguish using stellar densities and spatial distributions until the stars evolve dynamically. Nevertheless, we aim to select a consistent sample of objects over the full time interval of interest and exclude dissolving associations that would contaminate the sample and artificially boost the number of young clusters.

Fortunately, the dissolution of associations occurs on short timescales. A study by Gieles & Portegies Zwart (2011) demonstrated that by 10 Myr, the distinction between clusters and associations is clear. The ratio of a cluster's age to its crossing time, represented as $\Pi \equiv (\text{Age} \ / \ T_{\rm cross})$, increases to values >1 over this time span for systems that remain compact. In contrast, associations continually expand leading to $\Pi \leq 1$ at all ages. We adopt t_1=10 Myr for our study to avoid subjective determinations about the boundedness and longevity for indeterminate young objects. From a practical standpoint, this choice does not have a large effect on our analysis; due to M31's near-constant formation rate of stars and clusters over the past 300 Myr, this age limit only excludes a small fraction of the stellar mass accessible for study. Furthermore, in Section 4.3.1 we calculate Π for all clusters in our resulting sample to eliminate any possibility of persistent contamination.

We choose to measure Γ over two age ranges: 10–100 Myr and 100–300 Myr. The minimum and maximum values t_1=10 Myr and t_2=300 Myr are set by the limitations of the dataset as discussed above, while the division between these two bins at 100 Myr allows us to look for signs that Γ evolves with time. We primarily refer to the 10–100 Myr Γ measurements throughout our study due to the relatively higher time resolution and precision in cluster and field star age information at these ages, as well as the compatibility of this age range with those used by previous studies. In addition, measurements in this younger age bin should correspond better to present-day ISM properties and Kruijssen

(2012) predictions.

For our analysis, we assume no cluster dissolution over the relevant 10–300 Myr time interval and make no correction for this effect. Our analysis of cluster age distributions observed for the PHAT cluster sample (L. Beerman et al., in preparation; Fouesneau et al. 2014) appear consistent with little or no cluster destruction for the young cluster population. Under this assumption, the value of Γ should not change with time, and therefore:

$$\Gamma = \Gamma_0 = \Gamma_{10-100} = \Gamma_{100-300} \tag{4.2}$$

where Γ_0 represents an initial, intrinsic cluster formation efficiency, and Γ_{10-100} and $\Gamma_{100-300}$ represent clustered stellar fractions over age ranges of 10–100 Myr and 100–300 Myr. We assess the validity of this assumption and how cluster dissolution would affect our inference of Γ_0 in detail in Section 4.5.4.

4.2 Data

4.2.1 PHAT Observations and Photometry

The PHAT survey imaged 1/3 of the disk of M31 in six passbands spanning ultraviolet to near-infrared wavelengths. The survey provides resolved stellar photometry of 117 million sources that we use to determine the properties of both the cluster and field populations, with completeness limits that allow the detection of individual main sequence stars down to $\sim$3 $M_\odot$. Here we provide an overview of the crowded field stellar photometry derived for PHAT; full details are found in Dalcanton et al. (2012) and Williams et al. (2014).

All PHAT resolved stellar photometry is derived using the DOLPHOT software package, an updated version of HSTPHOT (Dolphin 2000). In this work, we use only the optical wavelength Advanced Camera for Surveys (ACS) data, obtained in the F475W (g) and F814W (I) passbands.

To fit the SFH for the field populations, Lewis et al. (2015) used two-band optical photometry catalogs from first generation survey photometry. These `gst` catalogs include high-quality detections that pass criteria for signal-to-noise ratio (S/N), crowding, and sharpness, using photometry parameters described in Dalcanton et al. (2012).

We characterized clusters using photometry catalogs that differ from the field star catalogs in two ways. First, the two-band optical photometry was calculated using new photometry parameters described in Williams et al. (2014). Second, we adopted a relaxed set of quality cuts for the cluster catalogs relative to the field `gst` catalogs: S/N > 4 in both passbands, $(\texttt{Sharp}_{\rm F475W}+\texttt{Sharp}_{\rm F814W})^2 \leq 0.1$, and no crowding cut. The resulting cluster photometry catalogs are similar to the survey's `st` catalog.

The photometry catalogs are supplemented by large numbers of artificial star tests (ASTs) that are used to quantify catalog completeness, photometric biases, and uncertainties. These ASTs are an essential component of the analysis for both field stars and clusters, enabling the accurate simulation of stellar populations across the CMD. Details concerning the cluster and field ASTs are discussed individually in Section 4.3.

4.2.2 Spatial Analysis Regions

To measure Γ and investigate its variation across the disk of M31, we divide the PHAT survey footprint into seven distinct regions. We define these regions according to three key considerations: the cumulative amount of total stellar mass required to make a statistically significant measurement of Γ ($\gtrsim 10^6$ $M_\odot$); the physical scales associated with dissolving stellar structure; and the variation of galactic environments in M31.

We show the adopted layout for our seven analysis regions in Figure 4.1. These regions were defined to isolate the distinctive 10 kpc star forming ring (Region 2) from the inner disk (Region 1; $R_{\rm gc}$ <10 kpc) and outer disk (Region 3; $R_{\rm gc}$ >13 kpc), and to divide the mass formed over the 10–100 Myr age range equally between the regions. As a result, each region hosts $\sim$3–5 $\times$ 10^6 $M_\odot$ of star formation during the 10–100 Myr epoch, and in addition to isolating the inner and outer disk environments, our subdivision of the 10 kpc star forming ring isolates two prolific star forming regions: OB54 (Region 2e) and OB30/31 (Region 2a), referencing region identifications from van den Bergh (1964). In addition to the seven primary analysis regions, we also report results for the 10 kpc ring as a whole by combining its five individual components (2a–2e), and also report average survey-wide results by integrating over all seven analysis regions.

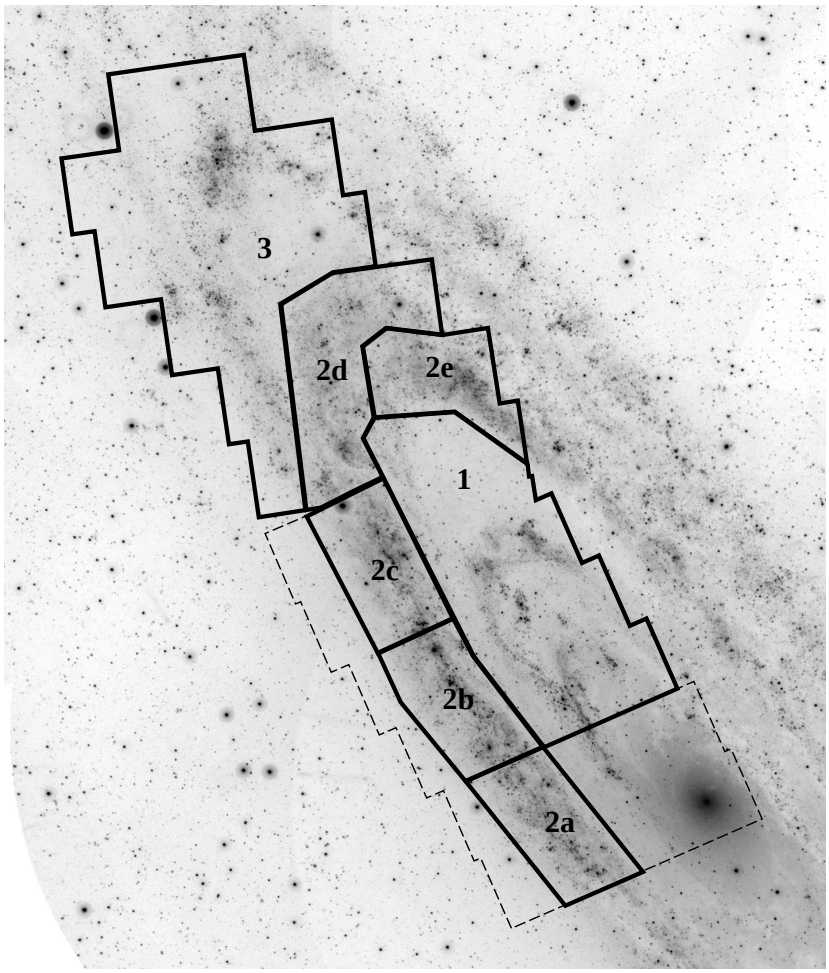

Figure 4.1 Spatial distribution of M31 analysis regions with labels. The underlying GALEX NUV image highlights young star forming regions. North is up and east is left in the image.

We omit two PHAT "bricks" from our analysis that lie in the central bulge-dominated regions of the galaxy (B01 and B03) due to increased levels of crowding that degrade the effective depth of the data and make SFH derivations more uncertain (Lewis et al. 2015). We also exclude the outer disk region outside the eastern portion of the star forming ring because we cannot cleanly separate the ring and outer disk components due to projection effects. Excluding these regions do not impact our results due to the negligible number of young clusters and total recent star formation that we omit.

4.2.3 Ancillary Data and ISM Properties

In addition to the PHAT survey data, we make use of HI observations from WSRT/GBT (Braun et al. 2009) and ^{12}CO(1-0) observations from IRAM (Nieten et al. 2006) to obtain estimates of total gas surface density ($\Sigma_{\rm gas}$=$\Sigma_{\rm HI}$+$\Sigma_{\rm H_2}$) across the disk of M31. These HI and CO datasets have angular resolutions of 30 arcsec and 23 arcsec, respectively. We refer the reader to the primary references for full details concerning the observations and image reduction. The IRAM CO data coverage does not extend beyond the star forming ring, therefore we supplement our knowledge of molecular gas in the outer disk of M31 using high resolution (5 arcsec) interferometric observations of ^{12}CO(1-0) from CARMA (A. Schruba et al., in preparation) obtained for a 300 arcsec diameter region in the vicinity of the OB102 star forming complex. We use these observations to measure the molecular gas fraction, $f_{\rm mol} \equiv \Sigma_{\rm H_2}/(\Sigma_{\rm HI}+\Sigma_{\rm H_2})$, and we obtain a value of 2%. The missing estimates of $\Sigma_{\rm H_2}$ will therefore have a negligible effect on $\Sigma_{\rm gas}$ values given that the outer disk is vastly dominated by its atomic gas component.

We derive basic properties of M31's ISM within each of the spatial analysis regions using these HI and CO datasets. For the HI data, we convert column density maps derived by Braun et al. (2009) directly to deprojected atomic gas surface density ($\Sigma_{\rm HI}$) assuming an inclination angle of 77 degrees. We measure molecular gas surface densities ($\Sigma_{\rm H_2}$) using CO maps making the same inclination as for the HI and adopting a CO-to-H_2 conversion consistent with observational constraints from the Milky Way (Bolatto et al. 2013): $\alpha_{\rm CO}$ = 4.35 $M_\odot$ pc^{-2} (K km s^{-1})$^{-1}$, which assumes $X_{\rm CO} = 2 \times 10^{20}$ cm^{-2} (K km s^{-1})$^{-1}$.

To calculate total gas densities, $\Sigma_{\rm gas}$, we combine $\Sigma_{\rm HI}$ and $\Sigma_{\rm H_2}$ maps and make a factor 1.36 correction to account for helium mass. The ISM throughout M31 is dominated by its atomic component; we find $f_{\rm mol}$ values in the range of 0.02–0.20. As a result, $\Sigma_{\rm gas}$~$\Sigma_{\rm HI}$, which varies between 1–10 $M_\odot$ pc^{-2}. We calculate mass-weighted average atomic, molecular, and total gas surface densities within our analysis regions; we discuss the merits of mass-weighted averages in detail in Section 4.3.2.1 and Section 4.7.1 in regards to $\Sigma_{\rm SFR}$ calculations.

In addition to measuring gas surface density, we estimate $\sigma_{\rm gas}$ values using maps of HI

Table 4.1. ISM Observational Data

Region ID	Region Name	$\Sigma_{\rm HI}$ ($M_\odot$ pc^{-2})	$\Sigma_{\rm H_2}$ ($M_\odot$ pc^{-2})	$\Sigma_{\rm gas}$ ($M_\odot$ pc^{-2})	$f_{\rm mol}$	$\sigma_{\rm gas}$ (km s^{-1})	$\overline{R}_{\rm gc}$[a] (kpc)	$v_{\rm c}$ (km s^{-1})	Ω (Myr^{-1})	Q
1	Inner Disk	1.85	0.77	2.62	0.176	7.94	6.61	200	0.031	9.64
2a	Ring-OB30/31	5.74	0.74	6.48	0.084	9.73	11.45	250	0.022	3.39
2b	Ring-OB39/40/41	6.93	0.90	7.83	0.084	9.46	12.14	250	0.021	2.60
2c	Ring-OB48	6.92	0.60	7.52	0.052	8.03	12.11	250	0.021	2.30
2d	Ring-Spur	6.08	0.61	6.69	0.054	8.16	12.16	250	0.021	2.63
2e	Ring-B15	5.49	0.78	6.27	0.108	8.12	11.14	250	0.023	3.06
3	Outer Disk	3.75	0.14	3.89	0.018	7.17	15.83	250	0.016	3.03

[a]Mass-weighted mean galactocentric radius.

non-thermal velocity dispersion from Braun et al. (2009). We find little spatial variation in mass-weighted $\sigma_{\rm gas}$ measurements, which span a range from 7–10 km s^{-1}. We report region-by-region $\sigma_{\rm gas}$ values, along with derived gas surface densities in Table 4.1.

4.3 Analysis

In this section, we describe how we compute the two key components of Γ, $M_{\rm cl}$ and $M_{\rm tot}$. We discuss cluster CMD fitting used to measure ages and masses in Section 4.3.1 and discuss CMD fitting used to measure total SFHs in Section 4.3.2.

4.3.1 Cluster Properties

4.3.1.1 PHAT Clusters: Catalog and Completeness

We draw our cluster sample from the Andromeda Project (AP) cluster catalog (Johnson et al. 2015b). This catalog includes 2753 star clusters that lie within the PHAT survey footprint, covering a wide range of ages and masses. These clusters were identified through visual inspection of optical (F475W, F814W) images by volunteer citizen scientists, facilitated through the Zooniverse's Andromeda Project website. Each image was examined >80 times, providing robust classification statistics for each cluster candidate. The final sample of clusters was selected according to the fraction of user-weighted cluster identifications using a catalog threshold that maximizes completeness and minimizes contamination with

respect to the expert-derived PHAT Year 1 cluster catalog (Johnson et al. 2012).

Young star clusters appear in PHAT imaging as collections of individually resolved member stars, as seen for four example clusters presented in Figure 4.2. For ages <300–500 Myr, the stellar main sequence (MS) is readily detectable, providing robust age constraints for young clusters. At older ages, red clump and red giant branch (RGB) member stars are still individually resolved, but these features have limited age sensitivity decreasing the precision of CMD-based constraints for these clusters.

Completeness information for the cluster catalog is derived from a suite of 3000 synthetic clusters that characterize catalog selection as a function of cluster properties, as well as its variation across the survey footprint. Detailed properties of the artificial test sample are discussed in Section 2.2 of Johnson et al. (2015b). These synthetic clusters were injected into AP search images and passed through the same cluster identification processing as all the real data. Input positions for the test objects were distributed uniformly throughout the survey footprint to assess cluster detection across the full range of galactic environments.

A critical component to estimating Γ is the extrapolation from the observed cluster mass to a total cluster mass. We use the synthetic cluster results to derive completeness functions (in terms of cluster mass) for each analysis region as the first step towards accounting for cluster mass that falls below the observational detection limit.

We calculate completeness functions for each analysis region, averaged over the two age ranges of interest, 10–100 and 100–300 Myr. For each region, we select a subsample of synthetic clusters whose input ages and local RGB stellar densities fall within each of the two age bins and the observed range of stellar densities found within the analysis region. This selection accounts for the fact that cluster detection not only depends on cluster mass, but also on age and local stellar background density. Due to the structure of M31's stellar disk, the RGB stellar density selection is roughly equivalent to one based on galactocentric radius. Next, the selected synthetic clusters are assigned weights according to their local SFR densities to transform the intrinsic distribution of background MS stellar densities to a version that better reflects the relative distribution of clusters and completeness in the region. Using these weighted results, we derive a completeness function in terms of cluster mass for each region and model the behavior using a logistic function. We report fitted 50%

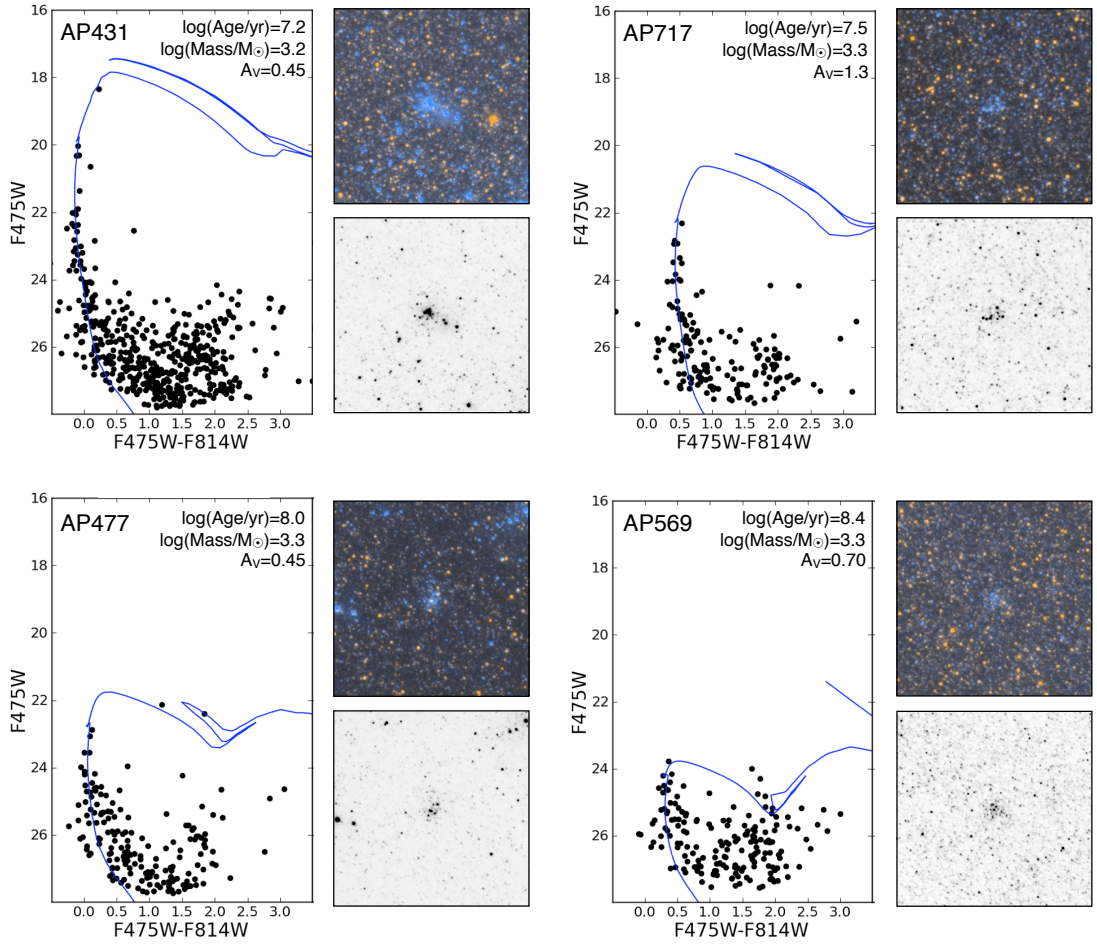

Figure 4.2 Four example clusters sampling the age range of interest for our Γ analysis. Each cluster was chosen to be $\sim 2 \times 10^3$ $M_\odot$ and logarithmically spaced in age between 10–300 Myr; fitted parameters for each cluster are listed in the figure. Isochrones from the Padova group (Marigo et al. 2008; Girardi et al. 2010) representing the best fit age and A_V from MATCH are overlaid on cluster CMDs. The color cutout is a F475W+F814W composite, the B/W cutout is an inverted version of a F475W image, and both are 15 arcsec (∼60 pc) on a side.

completeness limits for each of the analysis regions in Table 4.2. These limits range from 520–950 $M_{\odot}$ for the 10–100 Myr age bin and from 650–1250 $M_{\odot}$ for the 100–300 Myr age bin, depending on position within M31.

4.3.1.2 Determining Cluster Ages and Masses

CMD fitting of individually resolved member stars provides valuable constraints on a cluster's age and mass. We use the MATCH software package to analyze cluster CMDs following techniques described in Dolphin (2002). This software models observed CMDs by simulating stellar populations. The code populates theoretical isochrones according to input parameters that define the age, total mass, and dust attenuation of the population, as well as its distance, metallicity, stellar IMF, and binary fraction. We assume a simple stellar population model, such that each of these parameters is single valued when fit. These synthetic populations are convolved with a model of observational errors derived from ASTs and combined with a background model (here, representing non-cluster field populations) to produce a simulated CMD. This simulated CMD is then compared to the observed CMD, where the fit quality is evaluated according to a Poisson likelihood function. The software iterates through a series of synthetic CMDs to estimate the relative likelihood of different combinations of input parameters.

For cluster fitting, we adopt an M31 distance modulus of 24.47, a binary fraction of 0.35, a Kroupa (2001) IMF for masses from 0.15 to 120 $M_{\odot}$, and stellar models from the Padova group (Marigo et al. 2008) that include updated low-mass asymptotic giant branch tracks (Girardi et al. 2010). We limit the metallicity range to $-0.2 < [M/H] < 0.1$, matching $\sim Z_{\odot}$ present day gas phase metallicity observations within M31 (Zurita & Bresolin 2012; Sanders et al. 2012). A small variation in metallicity is allowed to provide systematic flexibility in the shape and location of the isochrones; metallicity is treated as a nuisance parameter and marginalized over when calculating constraints on the parameters of interest: age, mass, and dust attenuation.

We fit CMDs composed of stars that lie within a cluster's photometric aperture ($R_{\rm ap}$), adopting values from Johnson et al. (2015b). We characterize the underlying non-cluster

background population using stars that lie in an annulus between ∼1.2-3.2 $R_{\rm ap}$, which spans an area 10× the size of the cluster aperture. We perform $5{\times}10^4$ ASTs for each cluster to ensure accurate characterization of photometric recovery as a function of CMD position and cluster radius. Input positions for cluster ASTs are distributed radially according to the cluster's luminosity profile, ensuring that the overall completeness function accurately reflects the blend of photometric completeness and errors that affect the cluster-wide CMD.

We compute CMD fits for a grid of age and dust attenuation (A_V) values, and obtain mass determinations from the best-fit CMD model scaling at each grid point. We use relative likelihoods derived across the age-attenuation grid to obtain marginalized probability distribution functions (PDFs) for each of these parameters. We adopt the age, A_V, and mass of the best fit model and assign uncertainties to these values based on 16th and 84th percentiles of the marginalized 1D PDFs. We publish a full catalog of cluster parameter determinations, demonstrate that these results provide reliable results using synthetic cluster tests, and compare the CMD fitting results to those derived from integrated light fitting in L. Beerman et al. (in preparation).

The fitting identifies 1249 clusters with ages between 10–300 Myr that range in mass from 300 to 20,000 $M_\odot$; the age-mass distribution of the sample is shown in Figure 4.3. The sample's age distribution is consistent with a near-constant formation history with little or no cluster destruction, in agreement with initial PHAT results presented in Fouesneau et al. (2014). The median age uncertainty is 0.2 dex and the median mass uncertainty is 0.04 dex. We adopt 10% minimum uncertainties for all cluster mass estimates reflecting limits in precision due to systematic uncertainties.

With cluster ages and masses in hand, we can check for contamination from unbound associations. Following Gieles & Portegies Zwart (2011), we calculate the ratio of cluster age to crossing time, Π, using newly derived age and mass determinations and photometric half-light radii (equivalent to effective radius, $R_{\rm eff}$) from the AP catalog (Johnson et al. 2015b) to compute $T_{\rm cross}$:

$$T_{\rm cross} = 10 \left(\frac{R_{\rm eff}^3}{GM} \right)^{1/2}. \tag{4.3}$$

As discussed in Section 4.1.1, long-lived gravitationally bound clusters should retain short

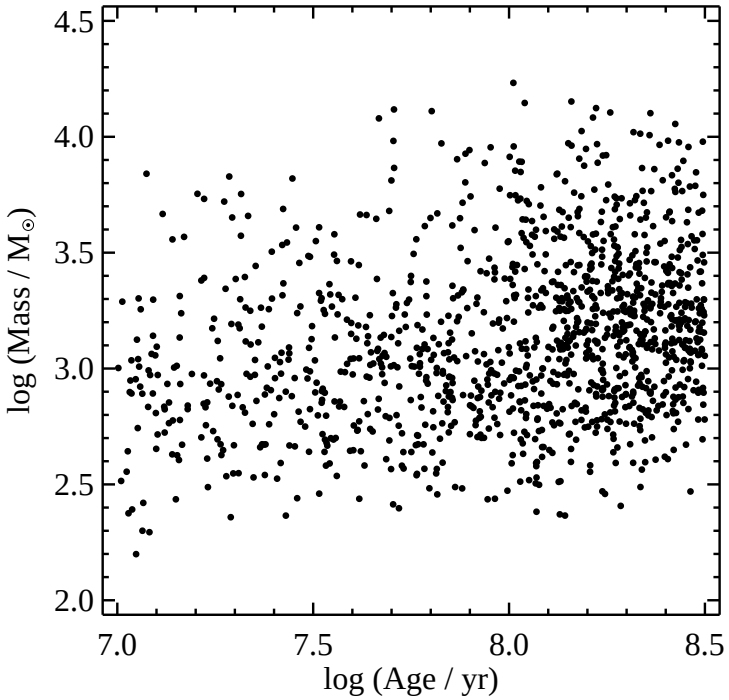

Figure 4.3 The age-mass distribution for 1249 PHAT/AP clusters in the 10–300 Myr age range. Random deviations of 0–0.1 dex in age are added to the 0.1 dex grid results to aid visibility.

Table 4.2. Cluster and SFH Observational Data

Region ID	Region Name	$M_{\rm cl,10-100}$ (10^4 $M_\odot$)	$M_{\rm tot,10-100}$ (10^6 $M_\odot$)	$M_{\rm cl,100-300}$ (10^4 $M_\odot$)	$M_{\rm tot,100-300}$ (10^6 $M_\odot$)	log $\Sigma_{\rm SFR}$ ($M_\odot$ yr^{-1} kpc^{-2})	$M_{\rm cl,lim}$[a] ($M_\odot$)
1	Inner Disk	11.36 ± 0.35	3.58 ± 0.04	34.38 ± 0.68	8.77 ± 0.14	-2.96	946
2a	Ring-OB30/31	16.77 ± 0.30	4.62 ± 0.09	12.05 ± 3.43	10.29 ± 0.28	-2.45	687
2b	Ring-OB39/40/41	7.45 ± 0.19	3.49 ± 0.07	16.33 ± 0.31	7.90 ± 0.21	-2.59	749
2c	Ring-OB48	7.16 ± 0.22	3.12 ± 0.06	16.47 ± 0.31	7.26 ± 0.18	-2.61	697
2d	Ring-Spur	8.62 ± 0.19	4.79 ± 0.07	34.50 ± 0.66	10.72 ± 0.23	-2.65	721
2e	Ring-B15	12.92 ± 0.21	3.77 ± 0.08	24.73 ± 0.55	9.06 ± 0.23	-2.48	830
3	Outer Disk	6.65 ± 0.18	3.49 ± 0.04	18.96 ± 0.64	7.36 ± 0.10	-3.13	522

[a]The observational 50% mass completeness limit for the 10–100 Myr age bin.

crossing times as their ages increase, and thus should have $\Pi > 1$ at ages $\geq$10 Myr.

We find that only 33 out of 1249 total 10–300 Myr old clusters have values of $\Pi<2$. This result suggests that contamination from associations is small, even when adopting a liberal defining threshold (versus canonical $\Pi=1$); these candidate associations together make up 4% of the total cluster mass in the 10–100 Myr age bin. Due to the ambiguity in defining a distinct threshold between clusters and associations based on observed Π values, and the small effect that excluding these objects would have on our final result, we opt to retain the full cluster sample and make no selection based on Π. The small fraction of possible contaminates suggests that adopting a minimum age of 10 Myr for our Γ analysis already successfully removed any significant population of potentially unbound stellar associations.

We calculate total observed cluster masses in each of the seven analysis regions for the 10–100 Myr and 100–300 Myr age bins ($M_{\rm cl,10-100,obs}$ and $M_{\rm cl,100-300,obs}$) by summing best fit masses. We derive uncertainties on these quantities by adding individual cluster mass uncertainties in quadrature. Region-by-region results are provided in Table 4.2.

4.3.2 Star Formation Histories

The second ingredient for calculating Γ is an estimate of the total stellar mass formed during the same time interval as the stellar clusters characterized in Section 4.3.1. We use recent SFH results calculated in Lewis et al. (2015); we refer the reader to that work for complete

details of the analysis, but provide a high-level overview of the analysis and results.

SFHs were derived from CMDs using the same MATCH software that was used for cluster fitting. The SFR is allowed to vary as a function of time for full SFH fitting (fit here with 0.1 dex resolution in logarithmic age), in contrast to cluster fitting that adopts the strong assumption of a simple stellar population. There are two other differences between the technique for computing extended SFHs rather than cluster SSPs. First, metallicity is allowed to vary but is restricted to increase with time. Second, dust attenuation is implemented using a two-parameter top hat model, defined by a minimum attenuation level and a differential spread. Other than these differences, assumptions for distance modulus, IMF, binary fraction, and stellar evolution models match those used for cluster analysis.

Lewis et al. (2015) present SFH results derived independently for ∼9000 individual regions that span the survey footprint, each measuring 24×27 arcsec (100×100 pc in projection). Input photometry used in this fitting was extracted from PHAT `gst` photometry catalogs described in Section 4.2.1. Accompanying ASTs were compiled from a master catalog of results constructed from individual sets of 10^5 ASTs that were run in each of the survey's ∼400 ACS fields of view. Each SFH analysis region uses $\sim 5\times10^4$ ASTs drawn from a 5×5 grid of adjacent regions, measuring 120×135 arcsec in total.

Random uncertainties associated with the SFHs are computed using a hybrid Monte Carlo (HMC) process (Dolphin 2013), producing 10^4 posterior samples of SFH parameter values. The 1σ uncertainties are calculated by identifying the region of parameter space with the highest probability density, containing 68% of the samples. In addition to these random uncertainties, there are possible sources of systematic uncertainties due to the adopted dust model parameters and the choice of stellar evolution models. For the purpose of our Γ analysis, we ignore both of these sources of uncertainty. First, the systematic uncertainty due to dust is negligible compared to the random uncertainties. Second, although there are non-trivial uncertainties and biases associated with adopting a specific set of stellar evolution models (see Dolphin 2012), our conclusions are based on relative SFRs and cluster masses that we derive self-consistently using a single set of model assumptions. Because any systematic offset is shared between the cluster and field results, we also omit this component of uncertainty from the error budget.

We spatially combine best-fit SFHs from Lewis et al. (2015) according to our analysis regions and integrate over 10–100 Myr and 100–300 Myr time intervals to obtain total masses, $M_{\mathrm{tot},10-100}$ and $M_{\mathrm{tot},100-300}$. We derive uncertainties for these integrated masses using an additional Monte Carlo sampling analysis. We draw 1000 realizations of the SFH for each original spatial region fit by Lewis et al. (2015) according to the confidence intervals established by the HMC analysis. We spatially and temporally combine these realizations in the same manner as the best fit results and adopt the relative scatter for the integrated mass from the Monte Carlo realizations as our combined uncertainty. We note that this method will tend to overestimate uncertainties on age-integrated masses due to significant covariance between neighboring age bins at high time resolution. However, we find that the derived uncertainties are already sufficiently small such that they are not a dominate component in the ultimate Γ error budget; any additional decrease in the total stellar mass uncertainty would have little or no effect on subsequent constraints. The resulting masses and uncertainties are presented in Table 4.2.

4.3.2.1 Calculating Σ_{SFR}

Previous work has shown a strong correlation between Γ and Σ_{SFR}, making the latter a valuable observable quantity. We use Σ_{SFR} to measure star formation intensity and it serves as our primary metric for comparing star forming environments. We calculate smoothed estimates of Σ_{SFR} using a deprojected 0.5 kpc^2 measurement kernel (an ellipse with major and minor axes of $\sim$100 and 23 arcsec, respectively), and masses derived from the full spatial resolution SFH analysis of Lewis et al. (2015) integrated over 10–100 Myr and 100–300 Myr age bins. The use of smoothing improves signal-to-noise and provides symmetric measurements in the deprojected spatial plane, while still maintaining good overall spatial resolution and the ability to characterize the range of star formation intensity found within each of our analysis regions.

We present the resulting Σ_{SFR} maps in Figure 4.4. The OB54 and OB30/31 regions stand out among the star forming complexes in the 10 kpc ring, and ring/arm features are visible in the inner and outer disk. The spatial distribution of star formation in M31 is

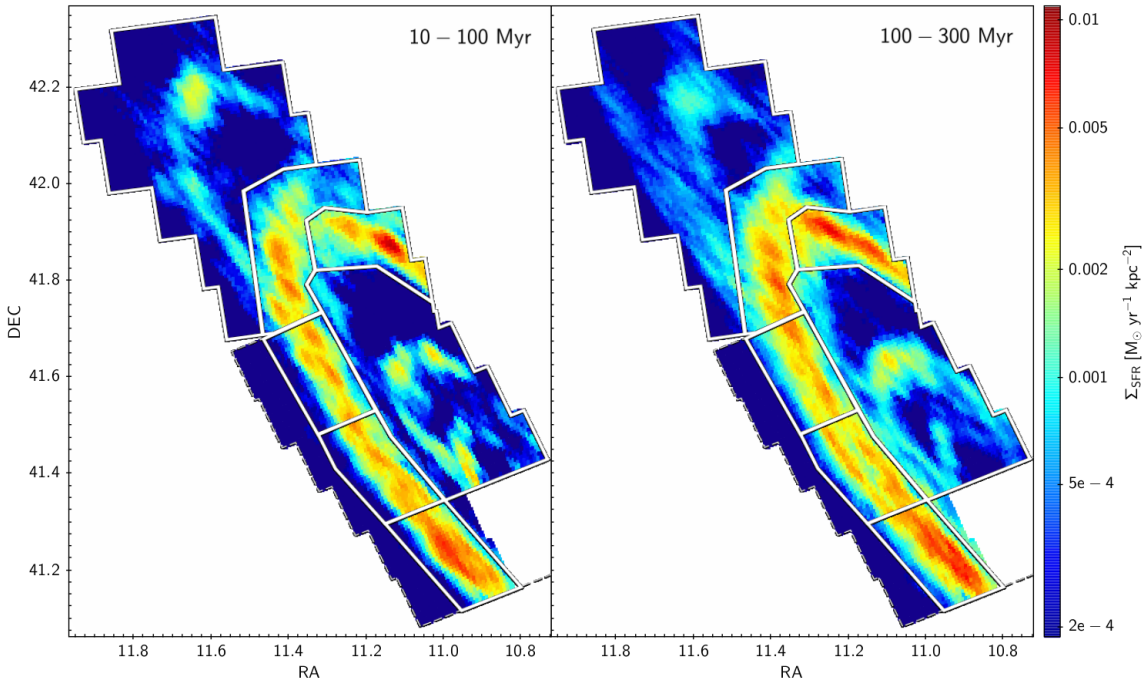

Figure 4.4 Maps showing $\Sigma_{\rm SFR}$ for 10–100 Myr (left) and 100–300 Myr (right) age bins, calculated using a deprojected 0.5 kpc^2 kernel.

distinctively not smooth nor uniform, which must be taken into account when assigning characteristic $\Sigma_{\rm SFR}$ values to each of the seven Γ analysis regions.

For each region, we calculate mass-weighted $\Sigma_{\rm SFR}$ values to characterize local star formation intensity. We average smoothed estimates weighted according to the integrated SF it represents. We report these mass-weighted mean values of $\Sigma_{\rm SFR}$ in Table 4.2. We also use the distribution of individual mass-weighted $\Sigma_{\rm SFR}$ values to characterize the spread in star formation intensity contained within each analysis region; see Section 4.7.1 for further discussion of these $\Sigma_{\rm SFR}$ distributions.

One advantage of this approach is its relative insensitivity to our choice of region boundaries. For example, the mass-weighted $\Sigma_{\rm SFR}$ estimate for the inner disk does not significantly change based on our decision to include or exclude the quiescent region that lies just inside the 10 kpc ring. Previous Γ analyses typically used simple area normalizations that produced unweighted $\Sigma_{\rm SFR}$ estimates, where the derived value is sensitive to exactly how much

low intensity star formation area was included in the aperture (e.g., in outer disk regions on the galaxy periphery). Inclusion of large areas with low or negligible SFRs can drive global $\Sigma_{\rm SFR}$ estimates to artificially low values, even in the case where all the SF within a given region takes place in a small, high intensity subregion. While the adoption of an unweighted area-normalization was often out of necessity (e.g., when SFR estimates were not available at higher spatial resolutions), these area-weighted values are susceptible to biases, particularly in the case of non-uniform, clumpy spatial distributions. We provide a comparison of mass-averaged and area-averaged $\Sigma_{\rm SFR}$ calculations in Section 4.7.1.

4.4 Results

Using results from the previous section, we calculate Γ and compare to theoretical predictions. We outline our probabilistic Γ analysis technique and present results in Section 4.4.1. Then, in Section 4.4.2 we compare our observational findings to model predictions from Kruijssen (2012).

4.4.1 Deriving Γ Constraints

We combine cluster mass results from our cluster characterization work with total stellar masses derived from the SFH analysis to obtain determinations of the fraction of stars born in long-lived clusters, Γ, over 10–100 Myr and 100–300 Myr age ranges. Here we introduce a forward modeling approach for transforming measurements of cluster mass and total stellar mass into Γ constraints, accounting for unobserved cluster mass and discrete sampling of the cluster mass function.

Our methodology uses two primary observational inputs: the observed cluster mass, $M_{\rm cl,obs}$ and the observed total stellar mass, $M_{\rm tot}$. However, note that Γ is defined as:

$$\Gamma = \frac{M_{\rm cl}}{M_{\rm tot}} = \frac{M_{\rm cl,tot}}{M_{\rm tot}}, \tag{4.4}$$

where $M_{\rm cl,tot}$ is the total mass in clusters. We transform between $M_{\rm cl,obs}$ and $M_{\rm cl,tot}$ as part of our modeling, extrapolating the cluster mass function beyond our observational completeness limit, represented here as $M_{\rm cl,lim}$, the 50% mass completeness limit derived from synthetic cluster results. We assume a power law mass function for the extrapolation,

where $\mathrm{d}N/\mathrm{d}M \propto M^{-2}$, over the range $10^2 < M/M_\odot < 2{\times}10^4$. We adopt a maximum mass of $2{\times}10^4$ $M_\odot$ based on the mass distribution of our 10–300 Myr cluster sample. We adopt a minimum mass of 100 $M_\odot$ based on short evolutionary time-scales for less massive clusters that would lead to their destruction on short timescales (<10 Myr; Moeckel et al. 2012) and to provide consistency with previous Γ studies. These assumed upper and lower limits of the mass function affect the absolute scaling of the Γ results, motivating our choice of values that are consistent with previous studies. The power law form of the cluster mass function with index -2 is well established as a good approximation for young cluster populations in M31 (Fouesneau et al. 2014), as well as other galaxies (e.g., Zhang & Fall 1999; Gieles et al. 2006a; Portegies Zwart et al. 2010). Strong evidence exists for a turnover at the high mass end of the mass function (e.g., Gieles et al. 2006a; Larsen 2009) where the exponential decline is well described by a Schechter (1976) distribution. For the purpose of our Γ analysis, we choose a simple power law model with an observationally-motivated maximum mass truncation rather than using a Schechter functional form; this choice has no significant effect on our Γ results.

Our modeling also accounts for the discrete sampling of the cluster mass function and its effect on our Γ constraints. Briefly, discrete sampling of the cluster mass function acts as a source of statistical noise in the transformation between $M_{\rm tot}$ and predicted values of $M_{\rm cl,obs}$; even when the intrinsic value of Γ is constant, random variations in the distribution of individual cluster masses can cause subsequent values of $M_{\rm cl,obs}$ to vary. This effect dominates the uncertainty in Γ here due to our tight constraints on $M_{\rm cl}$ and $M_{\rm tot}$ and the limited number of clusters contained in each region per age bin (∼80–100).

In terms of explicit calculations, we formulate a model that predicts an observed cluster mass, $M_{\rm cl,obs}$, from a set of input parameters, θ, which include Γ and $M_{\rm tot}$. The model begins by calculating a total cluster mass, $M_{\rm cl,tot}$, from the input parameters Γ and $M_{\rm tot}$. Next, a random seed value, X, is used to initiate a uniform random draw of discrete cluster masses from the power law cluster mass function described above. Finally, we apply an observationally-derived mass cut to the sampled masses, parameterized by the 50% completeness limit, $M_{\rm cl,lim}$, and obtain a prediction for the observed cluster mass, $\hat{M}_{\rm cl,obs}$.

The probability distributions for $M_{\rm cl,obs}$ are well described by a Gaussian, therefore

we use the following likelihood function to quantify the agreement between observed and predicted model quantities:

$$P(M_{\rm cl,obs}|\theta) = \frac{1}{\sqrt{2\pi}\sigma_{\rm cl}} \exp\left[-\frac{(M_{\rm cl,obs} - \hat{M}_{\rm cl,obs}(\theta))^2}{2\sigma_{\rm cl}^2}\right], \tag{4.5}$$

where θ represents the set of model parameters, $\{\Gamma, M_{\rm tot}, M_{\rm cl,lim}, X\}$. Using Bayes's theorem, we express the posterior probability of the model parameters in terms of the likelihood function:

$$P(\theta|M_{\rm cl,obs}) \propto P(M_{\rm cl,obs}|\theta)P(\theta). \tag{4.6}$$

The remaining term, $P(\theta)$ represents the priors on the model parameters. We adopt a flat prior for Γ ($0 \leq \Gamma \leq 1$), and use observationally derived means and associated Gaussian uncertainties as priors on $M_{\rm tot}$ and $M_{\rm cl,lim}$, listed in Table 4.2.

We use a Markov Chain Monte Carlo (MCMC) technique to efficiently sample the posterior probability distribution. Specifically, we use the `emcee`[1] package (Foreman-Mackey et al. 2013) and its implementation of an affine invariant ensemble sampler from Goodman & Weare (2010). For our fitting, we use 400 walkers, each producing 2000 step chains, of which we discard the first 100 burn-in steps. After completing the MCMC computation, we compute a marginalized posterior probability distribution for Γ, $P(\Gamma|M_{\rm cl,obs})$. We adopt the median value of the distribution as our primary Γ result and report the 16th to 84th percentile range as our 1σ confidence interval.

Throughout this paper we assume that cluster dissolution has a negligible effect over the adopted age range. As a result, we make no adjustment or corrections to the observed cluster mass other than the mass function extrapolation to $M_{min} = 100\ M_{\odot}$. If cluster disruption was significant, the true value of Γ would be larger than the result we obtain. We discuss the justification for this assumption in detail in Section 4.5.4.

We conclude here with a brief review of the advantages of our probabilistic modeling approach for calculating Γ constraints. Our main motivation for pursuing probabilistic fitting is its natural ability to derive robust confidence intervals on our derived values of Γ. The lack of robust uncertainty analysis has been a serious shortcoming of previous work.

[1] `http://dan.iel.fm/emcee/`

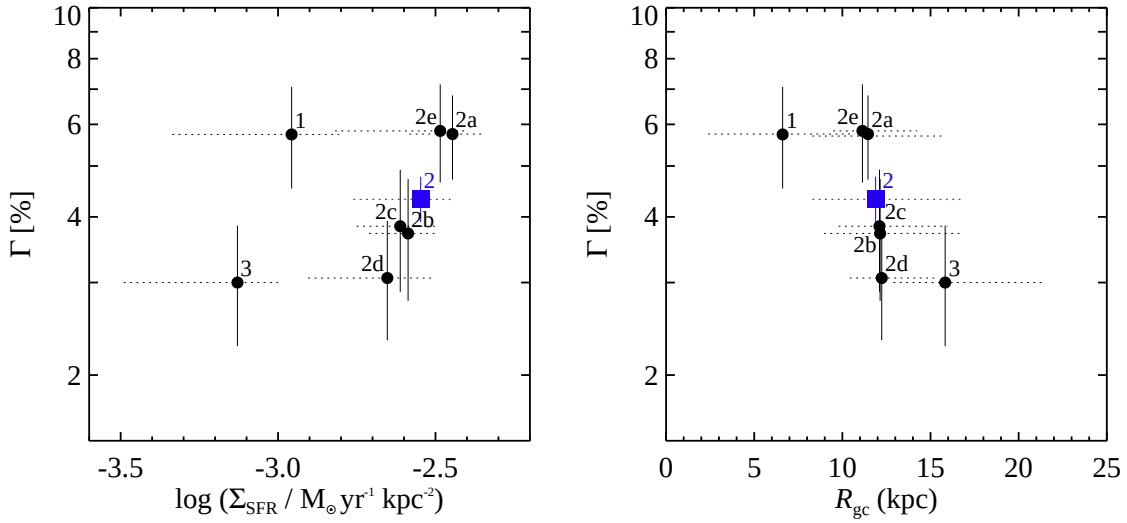

Figure 4.5 Γ results for the 10–100 Myr age bin, computed for each analysis region (black circles). We also show an aggregated data point (blue square) representing the combined result for the five regions in the 10 kpc star forming ring (those with $\log \Sigma_{\rm SFR} > -2.8$, and $10 < R_{\rm gc} < 13$). Left: Points are plotted at the median value of the Γ PDF and the region's mass-weighted mean $\Sigma_{\rm SFR}$. The solid vertical bars represent the 16th–84th percentile range of the Γ PDF, and the dotted horizontal bars represent the 25th–75th percentile range of the region's $\Sigma_{\rm SFR}$ distribution. Uncertainties for the mean $\Sigma_{\rm SFR}$ values are on order the size of the markers. Right: Γ results are plotted as in left panel, but now as a function of $R_{\rm gc}$. Dotted horizontal bars represent the full $R_{\rm gc}$ range of each analysis region.

We note, however, that recent studies have improved in this regard; for example, statistical variations due to discrete cluster mass function sampling were accounted for by Ryon et al. (2014) and Adamo et al. (2015), as well as Cook et al. (2012) in a limited sense. Within a probabilistic framework, we self-consistently combine constraints on individual input parameters while simultaneously accounting for extrapolation and stochastic sampling of the cluster mass function. Finally, our forward modeling approach allows a straightforward way to incorporate our empirically-derived cluster completeness limits, allowing greater usage of the entire observed cluster population.

Table 4.3. Γ Results and Predictions

Region ID	Region Name	Γ_{10-100} (%)	$\Gamma_{100-300}$ (%)	$\Gamma_{\rm predict}$ (%)
1	Inner Disk	$5.7^{+1.3}_{-1.2}$	$6.9^{+1.0}_{-0.9}$	2.0
2a	Ring-OB30/31	$5.8^{+1.1}_{-1.0}$	$2.3^{+0.9}_{-0.8}$	5.1
2b	Ring-OB39/40/41	$3.7^{+1.0}_{-0.9}$	$4.1^{+0.9}_{-0.8}$	6.1
2c	Ring-OB48	$3.8^{+1.1}_{-1.0}$	$4.3^{+0.9}_{-0.8}$	5.7
2d	Ring-Spur	$3.1^{+0.9}_{-0.7}$	$5.4^{+0.7}_{-0.7}$	4.6
2e	Ring-OB54	$5.8^{+1.3}_{-1.2}$	$5.2^{+0.9}_{-0.8}$	4.8
3	Outer Disk	$3.0^{+0.9}_{-0.7}$	$4.1^{+0.7}_{-0.7}$	2.9

4.4.1.1 Γ Results

We derive the probability distribution function of Γ for each of our spatial analysis regions, plot the results in Figure 4.5, and report our findings in Table 4.3. The precision of our cluster and total stellar mass determinations are on the order of a few percent, and as a result our reported confidence intervals for Γ reflect uncertainties due to stochastic sampling of the cluster mass function. This differs from previous studies where observational mass constraints are typically the dominant source of uncertainty and the contribution of sampling uncertainty is often ignored.

We find that Γ varies from 3–6% across the disk of M31. This result states that only a small fraction (<10%) of stars are born in long-lived star clusters, which follows the expected trend observed in previous studies given Andromeda's relatively quiescent level of star formation.

We find evidence for environmentally-dependent variations among the analysis regions. We observe that Γ varies in a broad sense with galactocentric radius ($R_{\rm gc}$), with cluster formation efficiencies in the outer disk region that are a factor of ∼2 lower than in the inner disk, and the mean result for the 10 kpc ring sits at an intermediate value. Yet, Figure 4.5 also shows that the behavior of Γ in M31 is more complex than a simple radial trend. Within the 10 kpc ring we find variations in Γ that span the full 3–6% observed range in spite of all five regions lying at approximately the same $R_{\rm gc}$. Furthermore, while the inner disk has a high cluster formation efficiency, it does not show an accompanying high $\Sigma_{\rm SFR}$.

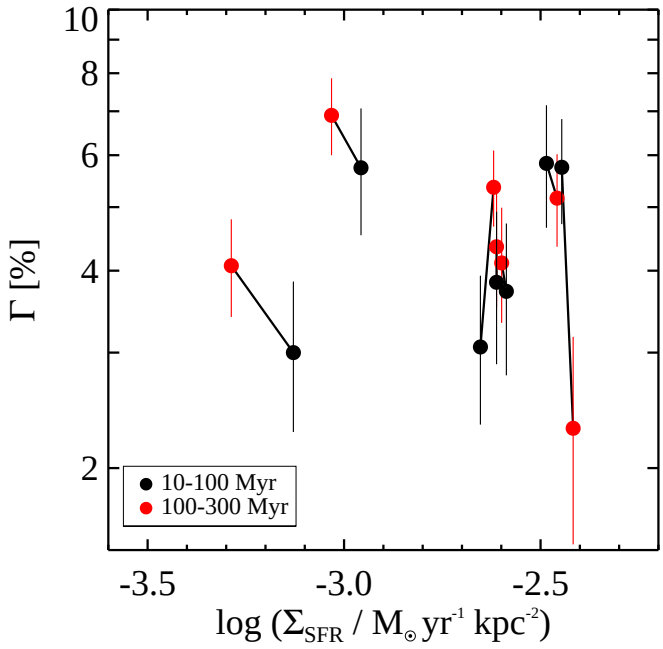

Figure 4.6 Γ results, comparing derived quantities for the analysis regions measured over two age ranges, 10–100 Myr (black points) and 100–300 Myr (red points).

This behavior conflicts with the expectation from a simple Γ-Σ_{SFR} correlation, given that the inner disk falls towards the bottom of the observed Σ_{SFR} range. These results show the richness of behavior captured by spatially-resolved studies that might otherwise be averaged out in galaxy-scale analyses of cluster formation efficiency.

We compare the values of Γ obtained for the 10–100 Myr and 100–300 Myr time bins in Figure 4.6. In general, we find good agreement between the Γ values derived for the two age bins. This result provides good evidence that cluster dissolution is negligible over the full 10–300 Myr age range, and gives additional confidence in the results we derive. The lone exception to this consistency with age is the southern-most subregion in the 10 kpc ring that hosts the OB30/31 star forming complex (Region 2a). While there is no clear explanation for this atypical behavior, we explore the possible influence of cluster dissolution in Section 4.5.4 as well as a number of other possible hypotheses.

4.4.2 Theoretical Γ *Predictions*

The theoretical framework presented in Kruijssen (2012) makes predictions for the fraction of stars formed in long-lived stellar clusters. This model is based on the idea that bound star clusters naturally arise from a hierarchically structured ISM, where clusters result from star formation that occurs in the high-density tail of the lognormal gas density distribution. In these high gas density regions the free-fall time is short, allowing time-integrated efficiencies calculated over the total duration of star formation (until it is truncated due to feedback processes or gas exhaustion) to reach high values, therefore increasing the likelihood of star cluster formation. Kruijssen (2012) assembles a self-consistent framework by combining: a model of a turbulent ISM within a gaseous disk that obeys hydrostatic equilibrium; a model of star formation that dictates a specific efficiency per free-fall time (Elmegreen 2002; Krumholz & McKee 2005); a model for the efficiency of initial cluster formation; and "cruel cradle" tidal destruction of structure during the embedded phase (<3-5 Myr).

This model of cluster formation efficiency depends primarily on total gas surface density, but also varies as a function of two additional input parameters: Toomre Q and angular velocity (Ω). Together, these three input parameters characterize the environmental conditions of star formation. In addition, Kruijssen (2012) introduces a simplified, fiducial Γ model, expressed in terms of $\Sigma_{\rm SFR}$. Observations of $\Sigma_{\rm SFR}$ are typically easier to obtain than determinations of $\Sigma_{\rm gas}$, Q, and Ω, thus this fiducial function serves an easy-to-use reference that has become the standard for theoretical Γ comparisons throughout the literature. The fiducial relation is calculated assuming a $\Sigma_{\rm gas}$-to-$\Sigma_{\rm SFR}$ conversion that follows from the Schmidt-Kennicutt star formation relation and typical environmental parameter values (Q=1.5, Ω derived from correlation with $\Sigma_{\rm gas}$ and therefore $\Sigma_{\rm SFR}$).

We derive detailed theoretical predictions for Γ using code[2] published as part of Kruijssen (2012) rather than simply comparing our observations to the simplified fiducial relation. We use region-specific observations of $\Sigma_{\rm gas}$, Q, and Ω as inputs, allowing us to compute tailored Γ predictions for each of our analysis regions. There are other adjustable parameters that allow the user to modify underlying assumptions for the star formation prescription, the

[2]Available at `http://www.mpa-garching.mpg.de/cfe/`

state of the gas and GMCs, feedback mechanisms from star formation processes, and the timescales for termination of star formation. We adopt default choices for these parameters, including an Elmegreen (2002) star formation prescription that dictates a single fixed star formation efficiency per free-fall time, and a SN-driven feedback prescription.

In addition to the gas surface density and velocity dispersion measurements presented in Section 4.2.3, we calculate the two remaining input parameters: angular velocity (Ω) and Toomre Q. We calculate Ω for each analysis region using a mass-weighted mean $R_{\rm gc}$ and assuming a flat rotation curve with a circular velocity of 250 km s^{-1} (Corbelli et al. 2010), except for the inner disk where we adopt a circular velocity of 200 km s^{-1}. We calculate the Toomre Q parameter for the gas disk using the expression

$$Q \equiv \frac{\kappa \sigma_{\rm gas}}{\pi G \Sigma_{\rm gas}} \approx \frac{\sqrt{2}\Omega \sigma_{\rm gas}}{\pi G \Sigma_{\rm gas}}, \tag{4.7}$$

where κ is the epicyclic frequency, $\sigma_{\rm gas}$ is the 1D velocity dispersion of the gas, Ω is the angular velocity within the galaxy, and where the second equality assumes that the rotation curve is flat within the disk region of interest. We present Ω and Q values for each region in Table 4.1.

4.4.2.1 Model Results

Predictions for Γ are compared to observations in Figure 4.7. The steady increase in the predicted values of Γ with gas surface density shows that the Kruijssen (2012) model for cluster formation efficiency is largely driven by changes in gas density, as expected for low to moderate $\Sigma_{\rm gas}$ environments. The "cruel cradle" tidal destruction component of the model only becomes important in high $\Sigma_{\rm gas}$ environments not found in M31.

The observed cluster formation efficiencies are a factor of $\sim$1.5 smaller than theoretical predictions for three of the analysis regions in the 10 kpc star forming ring, while the other two ring regions have Γ values that are slightly larger than the prediction. This discrepancy is likely due to the lack of correspondence between present day $\Sigma_{\rm gas}$ values and the properties of the progenitor gas that produced these 10–100 Myr old stellar populations. Analyses of the molecular gas component in nearby galaxies (e.g., Kawamura et al. 2009; Meidt et al. 2015), including M31 (L. Beerman et al., in preparation), have shown that

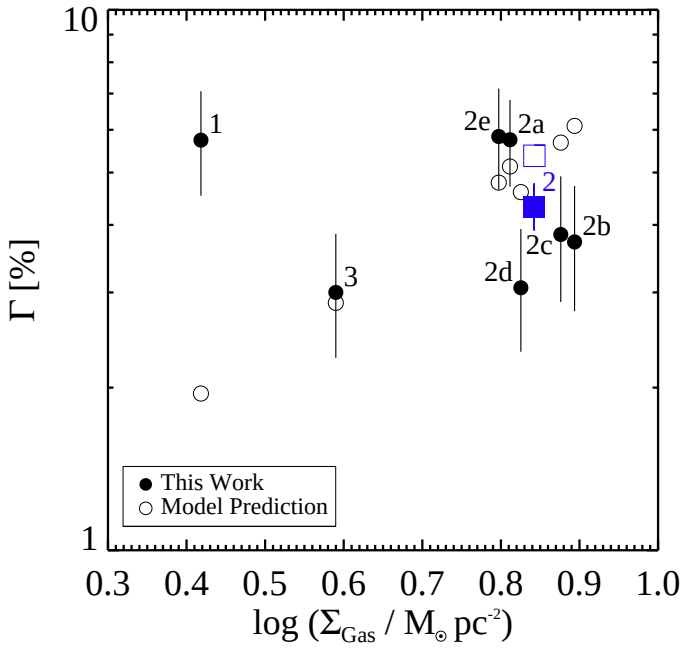

Figure 4.7 Γ results compared to predictions from the Kruijssen (2012) model, presented as a function of present-day $\Sigma_{\rm gas}$. Γ observations (filled symbols) and predictions (open symbols) for individual analysis regions are plotted as circles, and averaged Γ observations and predictions for the 10 kpc ring are plotted as squares. The observed inner disk observation of Γ ($\log \Sigma_{\rm gas} \sim 0.4\ M_\odot\ {\rm pc}^{-2}$) shows a large discrepancy with the predicted value; we discuss this disagreement in Section 4.5.1.

molecular cloud lifetimes are short — on the order of <20–50 Myr. Therefore, the cloud population responsible the 10–100 Myr populations we explore in this work are likely no longer in existence, having dispersed due to stellar feedback. Indeed, the OB30/31 region shows this particularly well given the marked absence of dense gas in the proximity of this ∼50 Myr old star forming complex.

Due to the amount of time that has passed since the epoch of formation, we do not expect the $\Sigma_{\rm gas}$ measurements to align with the 10–100 Myr old populations we are studying in detail on a region-by-region basis. However, the longevity of the 10 kpc star forming ring (Lewis et al. 2015) allows us to make the assumption that conditions in the gas remain similar on average at the integrated scale of the ring. From this point forward, we utilize ring-integrated estimates of $\Sigma_{\rm gas}$ (and associated values of $\tau_{\rm dep}$) as characteristic ISM values of all five 10 kpc ring regions during further analysis of the 10–100 Myr old cluster and stellar populations.

The largest disagreement between predicted and observed cluster formation efficiencies occurs in the inner disk analysis region. Despite its low gas surface density, this region has a Γ value of ∼6%; this value equals the highest Γ values obtained in the star forming ring where $\Sigma_{\rm gas}$ is a factor of ∼3 higher. In contrast, the model predicts a low cluster formation efficiency of only 2%. We remind the reader that this same inner disk region also appeared to show a relatively large value of Γ given its low $\Sigma_{\rm SFR}$ (Section 4.4.1.1). We discuss a possible explanation for this discrepancy next in Section 4.5.1, which leads to an interesting test of the Kruijssen (2012) model in a unique portion of parameter space.

4.5 Discussion

4.5.1 Inner Disk Γ Prediction: Accounting for Σ_*

We uncover an interesting situation in the inner disk of M31: the low $\Sigma_{\rm gas}$ observed for the region leads to a small predicted value for the cluster formation efficiency, however observations reveal a relative high cluster formation efficiency of ∼6% — a Γ value that is a factor of 2.5 larger than the prediction. The model presented by Kruijssen (2012) assumes a fixed characteristic ratio for the mid-plane pressures contributed by the vertical

distributions of gaseous and stellar disk components. The inner disk region of M31 has a relatively low gas surface density and is dominated by its stellar component, thus breaking normal assumptions about the relative contributions of gas and stars to the mid-plane pressure.

We alter standard model assumptions to account for the dominate contribution of the stellar component to the mid-plane pressure. The ϕ_P parameter is a dimensionless constant defined as the scaled contribution to the mid-plane pressure ($P_{\rm mp}$) from stars with respect to that from the gas. This factor is defined by Krumholz & McKee (2005) in their expression for mid-plane pressure as a sole function of $\Sigma_{\rm gas}$:

$$P_{\rm mp} = \phi_P \frac{\pi}{2} G \Sigma_{\rm gas}^2 \tag{4.8}$$

where ϕ_P is defined as

$$\phi_P = \phi_{\rm mp} f_{\rm gas}^{-1}. \tag{4.9}$$

Here, the constants $\phi_{\rm mp}$ and $f_{\rm gas}$ are defined as

$$\phi_{\rm mp} = \frac{\Sigma_{\rm gas}}{\Sigma_{\rm tot}} + \frac{\sigma_{\rm gas}}{\sigma_*}\frac{\Sigma_*}{\Sigma_{\rm tot}} = f_{\rm gas} + \frac{\sigma_{\rm gas}}{\sigma_*}(1 - f_{\rm gas}) \tag{4.10}$$

$$f_{\rm gas} = (\Sigma_{\rm gas}/\Sigma_{\rm tot}) \tag{4.11}$$

where Σ_* represents the stellar surface density, $\Sigma_{\rm tot} \equiv \Sigma_{\rm gas} + \Sigma_*$is the total mass surface density, and σ_* is the velocity dispersion of the stars. In Appendix A of Krumholz & McKee (2005), the authors argue that ϕ_P should have a constant value of $\sim$3 across a wide range of galactic environments.

Within the inner disk analysis region ($2.4 < R_{\rm gc}/{\rm kpc} < 10.7$), the contribution to the mid-plane pressure from the stellar disk (and the bulge/spheroid to a minor degree) is significant. We calculate ϕ_P explicitly for this region using observational constraints from previous studies of M31. We derive a deprojected stellar mass density of $\sim$80 $M_\odot$ pc^{-2}at the SF-weighted mean $R_{\rm gc}$ of 6.6 kpc using the galactic stellar mass profile from Tamm et al. (2012) and adopt σ_*=36 km s^{-1} from (Collins et al. 2011). We use $\Sigma_{\rm gas}$ and $\sigma_{\rm gas}$ values for the inner disk listed in Table 4.1 to compute $\sigma_{\rm gas}/\sigma_*$=0.22 and $f_{\rm gas}$=0.03. These values yield $\phi_{\rm mp}$=0.24 and ϕ_P=8.0 for the inner disk region. This newly derived value of ϕ_P

is significantly higher than the default value of 3, as expected due to the region's dominant stellar component. This value is sufficiently large that it lies outside the generous range of parameter space explored by Kruijssen (2012) in their study of parameter dependence.

We recompute the Kruijssen (2012) model prediction for the inner disk adopting the new, observationally-derived value of ϕ_P. We obtain a new prediction of 3.3% for the cluster formation efficiency, up from the original prediction of 2.0%. This increase improves the agreement between the Γ prediction and the observed value of $5.7^{+1.3}_{-1.2}$%, but a significant discrepancy remains.

In addition to its efficient cluster formation and high Σ_*, the inner disk region of M31 also shows a relatively high abundance of molecular gas ($f_{\rm mol}$=0.18), a large quiescent fraction of molecular clouds (L. Beerman et al, in preparation), and a relatively low total gas depletion time ($\tau_{\rm dep}$=2.4 Gyr) given its low value of $\Sigma_{\rm gas}$. Together, these observations suggest a unique set of conditions that provide a challenge for future Γ and star formation modeling efforts.

4.5.2 Galaxy-wide Γ *Results*

As discussed in the introduction, a growing body of evidence has revealed an environmental dependence of the fraction of stars born in long-lived star clusters. Beginning with Goddard et al. (2010), numerous studies have measured cluster formation efficiencies at galaxy-integrated scales, revealing a positive correlation between Γ and $\Sigma_{\rm SFR}$. Work by Goddard et al. (2010), Adamo et al. (2011), Silva-Villa & Larsen (2011), and Cook et al. (2012) each contribute galaxy-integrated measurements for small samples of galaxies. Additionally, studies by Annibali et al. (2011), Baumgardt et al. (2013), Ryon et al. (2014), Lim & Lee (2015), and Adamo et al. (2015) contribute results for individual galaxies. Together, these studies provide Γ measurements for a combined sample of >31 galaxies. Section 4.7.2 provides a detailed discussion concerning the curation of this set of results, and includes comments and caveats specific to individual studies.

We combine results from our seven analysis regions and compute a PHAT survey-averaged Γ measurement of 4.3±0.4%; we provide a full set of survey-wide results in Tables

4.2, 4.1, and 4.3. We plot this measurement, the curated set of galaxy-integrated literature values, and the predicted fiducial galaxy relation from Kruijssen (2012) in Figure 4.8. The PHAT survey-wide Γ measurement follows the established (but noisy) Γ-$\Sigma_{\rm SFR}$ trend previously observed, and lies near the predicted relation from Kruijssen (2012).

The current compilation of results shows an overall trend of increasing Γ with $\Sigma_{\rm SFR}$, but the scatter around the predicted relation is significant. This scatter could be due to physical differences in the observed galaxies, or due to observational heterogeneity and uncertainty. With regards to the possibility of actual physical differences, Kruijssen (2012) is clear in stating that his default Γ-$\Sigma_{\rm SFR}$ relation represents the characteristic expectation from the model, but that one should expect variations on the order of a factor of $\sim$2 due to real variation in star forming environments. Thus, plausible combinations of $\Sigma_{\rm gas}$, Toomre Q, and Ω exist that differ from the default values, leading to scatter in the predicted results.

Differences in analysis techniques, assumptions, and data quality among these heterogeneous Γ studies could also lead to the observed scatter. Authors differ in their methodology for deriving total SFRs (e.g., resolved stars versus Hα/FUV luminosity transformations), in the cluster age and mass ranges studied, and in the estimation of uncertainties (including cases where this analysis was not performed). In particular, we highlight the lack of reported uncertainties, and the underestimation of uncertainties in cases where these values are reported (i.e., not accounting for stochastic sampling of the cluster mass function), as a serious obstacle to differentiating between genuine Γ variation and observational scatter.

We find the overall level of agreement between multiple studies and with the model prediction quite impressive, but conclude that the heterogeneous nature of these galaxy-integrated observations fundamentally limit deeper interpretation of these results.

4.5.3 Spatially Resolved Γ Results and Star Formation Relation

While galaxy-integrated Γ observations provide excellent evidence for an environmentally dependent cluster formation efficiency, averaging a large range of star forming environments into a single constraint washes out interesting, revealing behavior. The advent of spatially resolved Γ measurements in M31 and M83 (Silva-Villa et al. 2013; Adamo et al. 2015)

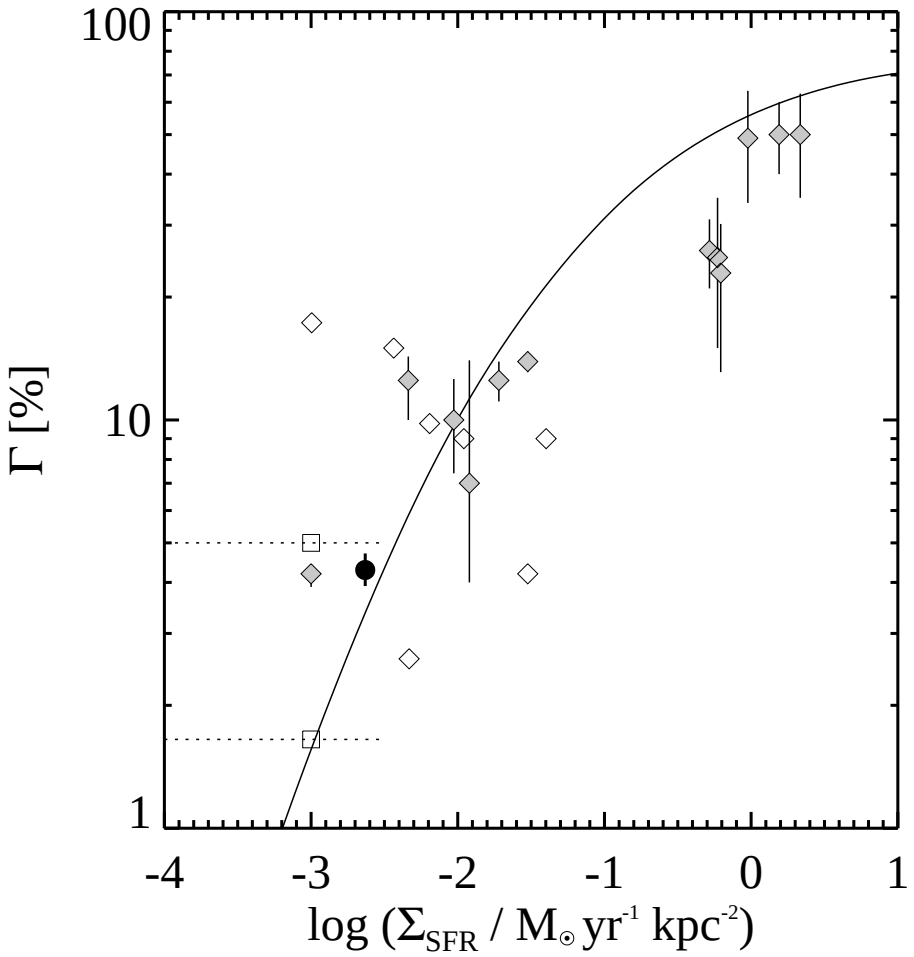

Figure 4.8 The survey-averaged Γ constraint for PHAT (black circle) is compared to previously published galaxy-wide results (diamonds; see text for references). Literature results that include uncertainties are plotted using filled gray symbols, while those without uncertainties are plotted using open symbols. The binned results from Cook et al. (2012) are plotted at their average values (open squares) with horizontal dotted lines representing the $\Sigma_{\rm SFR}$ bin width. The solid line represents the fiducial Γ relation from Kruijssen (2012). This compilation of galaxy-integrated Γ measurements show a positive correlation between Γ and $\Sigma_{\rm SFR}$ in good agreement with the predicted theoretical relation.

allows us to study cluster formation efficiency in greater detail. We focus on these two sets of results in particular due to the similarity of age range probed (using the 10–50 Myr equal-area region results from Adamo et al. 2015) and the availability of ISM observations.

When we consider spatially resolved Γ constraints, we must take care to adapt assumptions about star formation behavior that are calibrated at galaxy-wide scales. As discussed in Section 4.4.2, Kruijssen (2012) adopts the traditional assumption of a single, global star formation law ($\Sigma_{\rm SFR} \propto {\Sigma_{\rm gas}}^N$) defined by the Schmidt-Kennicutt relation (Kennicutt 1998) and its N=1.4 power-law slope when he derives a fiducial Γ-$\Sigma_{\rm SFR}$ relation. The often-used Schmidt-Kennicutt relation was defined using galaxy-integrated star formation observations of normal disk galaxies and starburst systems, but only includes a limited number of low intensity star formation environments. Spatially resolved studies of nearby galaxies, however, have shown that the relationship between $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ does not follow a simple single power law. Instead, star formation efficiencies are observed to decrease dramatically as environments transition from a molecular gas-dominated regime to an atomic gas-dominated regime. This behavior is shown clearly by Bigiel et al. (2008) using sub-kpc scale observations of nearby star forming galaxies. They find a linear relation between $\Sigma_{\rm H_2}$ and $\Sigma_{\rm SFR}$, but no correlation between $\Sigma_{\rm HI}$ and $\Sigma_{\rm SFR}$. These relations combine to yield a tight linear correlation between $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ at high gas densities that steepens with greatly increased scatter as the ISM becomes HI-dominated at $\Sigma_{\rm gas} < 10\ M_\odot\ {\rm pc}^{-2}$. This change in SF efficiency can be equivalently characterized as a change in the gas depletion time ($\tau_{\rm dep} \equiv \Sigma_{\rm gas}$ / $\Sigma_{\rm SFR}$), with inefficient SF at low gas density corresponding to long $\tau_{\rm dep}$.

In this section, we begin by defining a star formation relation based on spatially-resolved observations of nearby star forming galaxies. Next, we replace the original Schmidt-Kennicutt assumption from Kruijssen (2012) and use this new relation to translate Γ predictions calculated as a function of $\Sigma_{\rm gas}$ into $\Sigma_{\rm SFR}$ parameter space. Finally, we compare spatially resolved M31 and M83 Γ measurements to the newly derived theoretical predictions, and explore additional applications of the revised models to a wide range of Γ observations.

4.5.3.1 A Spatially Resolved Star Formation Relation: $\Sigma_{\rm gas}$ versus $\Sigma_{\rm SFR}$

Observations of nearby star forming galaxies provide evidence for (at least) two regimes of star formation behavior that each show a significant amount of intrinsic scatter. To account for these properties of spatially resolved star formation, we define a new star formation relation for use in conjunction with the Kruijssen (2012) Γ model. This new relation captures both the change in slope as well as the observed scatter, replacing the Schmidt-Kennicutt based transformation from $\Sigma_{\rm gas}$ to $\Sigma_{\rm SFR}$ with a flexible conversion. Namely, we allow an empirically-defined range of conversion factors, parameterized here by $\tau_{\rm dep}$, that are consistent with nearby galaxy observations.

We use observations from Bigiel et al. (2008) to define the two slope values of the new $\Sigma_{\rm gas}$-$\Sigma_{\rm SFR}$ relation, as well as an acceptable range of $\tau_{\rm dep}$ as a function of $\Sigma_{\rm gas}$. In Figure 4.9, we show that the distribution of Bigiel et al. (2008) observations[3] is well-characterized by a two-part star formation law, where $\Sigma_{\rm SFR} \propto {\Sigma_{\rm gas}}^N$ with N=1 at high H_2-dominated gas densities, and N=3.3 for low HI-dominated gas densities. This behavior differs significantly from the Schmidt-Kennicutt law, which we include for comparison. We characterize the intrinsic scatter using parallel upper and lower thresholds that define an envelope around the median relation, encompassing the $\tau_{\rm dep}$ variation observed in the data: 0.6 dex for high $\Sigma_{\rm gas}$, 1.6 dex at low $\Sigma_{\rm gas}$.

We take the ratio of $\Sigma_{\rm gas}$ and $\Sigma_{\rm SFR}$ measurements derived for the Γ analysis regions in M31 and M83 to compute $\tau_{\rm dep}$ values, and compare these observations to the Bigiel et al. (2008) dataset and our newly derived SF relation in Figure 4.9. We remind the reader that we only consider ring-integrated $\Sigma_{\rm gas}$ results for the 10 kpc ring region (see Section 4.4.2.1). Most of the M31 and M83 regions lie on the upper envelope of local observations, representing relatively short depletion times and high star formation efficiencies with respect to typical local galaxies. The innermost annulus in M83 and the inner disk region in M31 are exceptions to this overall trend: the innermost M83 region lies near the lower envelope

[3]We use as reference the distribution of $\Sigma_{\rm SFR}$(FUV+24μm) versus $\Sigma_{\rm gas}$ observations, represented by the contiguous portion of the orange contour (denoting a density of 2 samples per 0.05 dex-wide cell) from Figure 8 in Bigiel et al. (2008). This distribution is shifted by a factor of 1.36 in our work to account for the mass of helium that we include in $\Sigma_{\rm gas}$ that was not included in the original work.

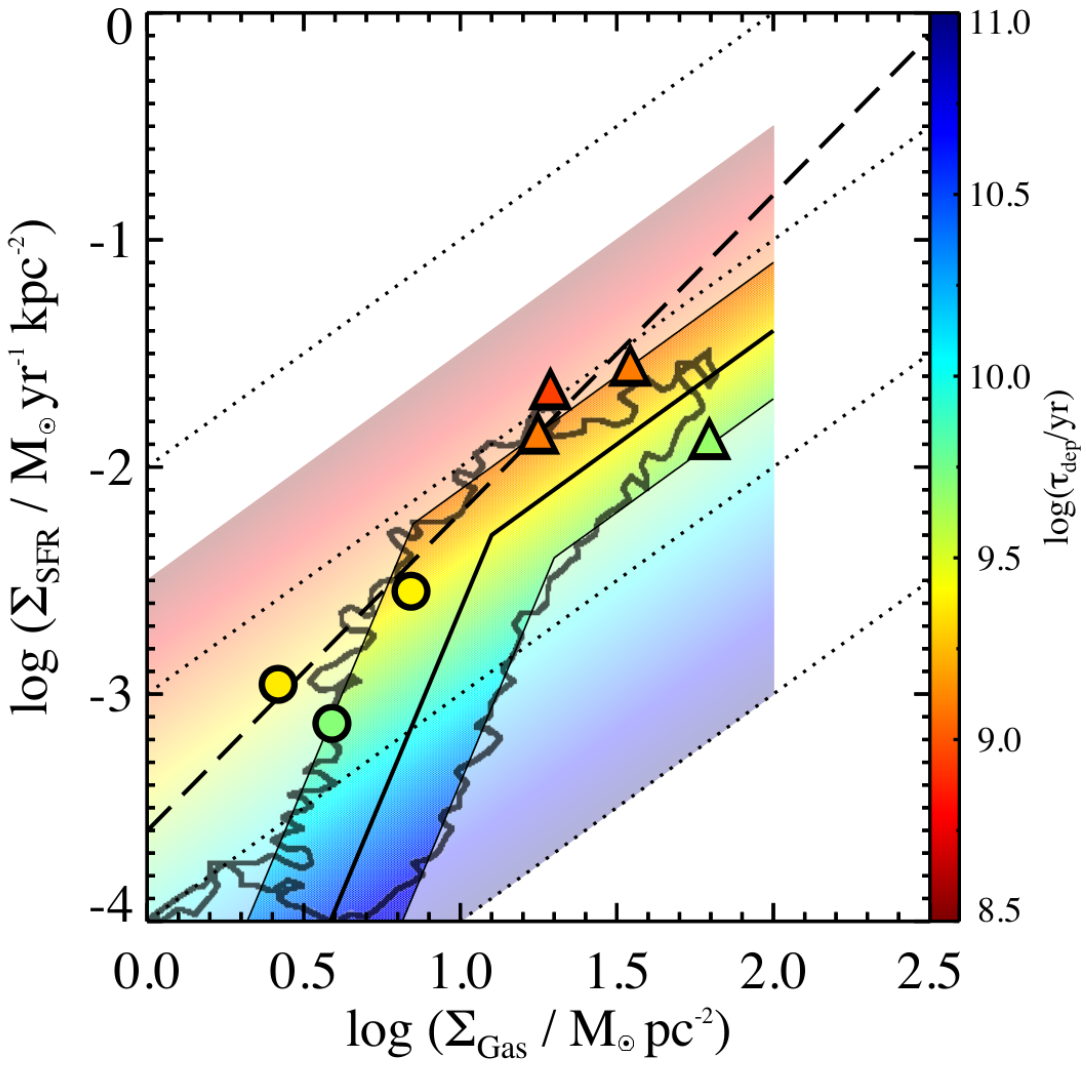

Figure 4.9 Comparison between the Schmidt-Kennicutt star formation relation (black dashed line) and spatially resolved observations from Bigiel et al. (2008) (gray contour). We include observations of M31 from this paper (circles; inner disk, outer disk, and 10 kpc ring), and observations from Adamo et al. (2015) of M83 (triangles; equal area annuli). We use a series of broken power laws to characterize the range of $\Sigma_{\rm SFR}$ (and thus $\tau_{\rm dep}$) as a function of $\Sigma_{\rm gas}$ that is consistent with the Bigiel et al. (2008) observations; we plot the median two-component star formation relation (thick black line) and its accompanying upper and lower envelopes (thin black lines). We also plot dotted lines that represent constant $\tau_{\rm dep}$ for log(yr) of 8, 9, 10, and 11 (from top to bottom), and include a background color gradient encoding $\tau_{\rm dep}$ values from $8.5 < \log({\rm yr}) < 11$.

of the Bigiel et al. (2008) observations and has a relatively large $\tau_{\rm dep}$, while the inner disk region of M31 lies well above the distribution of local star formation observations.

The inner disk region in M31 once again appears atypical, showing a relatively short $\tau_{\rm dep}$ given its low $\Sigma_{\rm gas}$. As we discussed in Section 4.5.1, the dominant contribution of Σ_* to the total mid-plane pressure could alter the normal behavior of $\tau_{\rm dep}$ as a function of $\Sigma_{\rm gas}$. The high stellar density would enhance the efficiency of star formation, lowering $\tau_{\rm dep}$ in the inner disk of M31 to 2.4 Gyr from an expected value of >10 Gyr given the region's $\Sigma_{\rm gas}$. This decreased value of $\tau_{\rm dep}$, and equivalent increase in total star formation efficiency, matches the conditions found in the 10 kpc ring, even though the inner disk hosts a factor of 2.7 lower gas surface density.

4.5.3.2 The $\tau_{\rm dep}$-dependent Spatially Resolved Γ Relation

We use our newly derived $\Sigma_{\rm gas}$–$\Sigma_{\rm SFR}$ relation based on the Bigiel et al. (2008) observational results to compute new Γ predictions in terms of $\Sigma_{\rm SFR}$ and $\tau_{\rm dep}$. We present these newly transformed Kruijssen (2012) model predictions in Figure 4.10. First, we derive a median predicted Γ–$\Sigma_{\rm SFR}$ relation using the median $\Sigma_{\rm gas}$-to-$\Sigma_{\rm SFR}$ relation in Figure 4.9. Then, we propagate the allowed scatter in the new star formation relation into our Γ–$\Sigma_{\rm SFR}$ predictions by defining upper and lower envelopes around the median relation that follow from the upper and lower thresholds defined in Figure 4.9. The parameter space bounded by these parallel envelopes represents the new preferred range of spatially-resolved Γ predictions as a function of $\Sigma_{\rm SFR}$. The range of preferred Γ values at a fixed value of $\Sigma_{\rm SFR}$ is defined by an empirically-defined range of $\tau_{\rm dep}$ values that are consistent with the Bigiel et al. (2008) observations. Therefore, given a specific pairing of $\Sigma_{\rm SFR}$ and $\tau_{\rm dep}$, the Kruijssen (2012) model provides a unique prediction of Γ; this aspect is represented in Figure 4.10 using a $\tau_{\rm dep}$ color coding of model predictions.

Adopting a spatially resolved star formation law yields new predictions for Γ that differ from the fiducial, galaxy-integrated relation from Kruijssen (2012). At fixed values of $\tau_{\rm dep}$ (curved bands of constant color), we find a steeper relation of Γ as a function of $\Sigma_{\rm SFR}$ with respect to the original relation; this results from flattening the slope of the underlying star

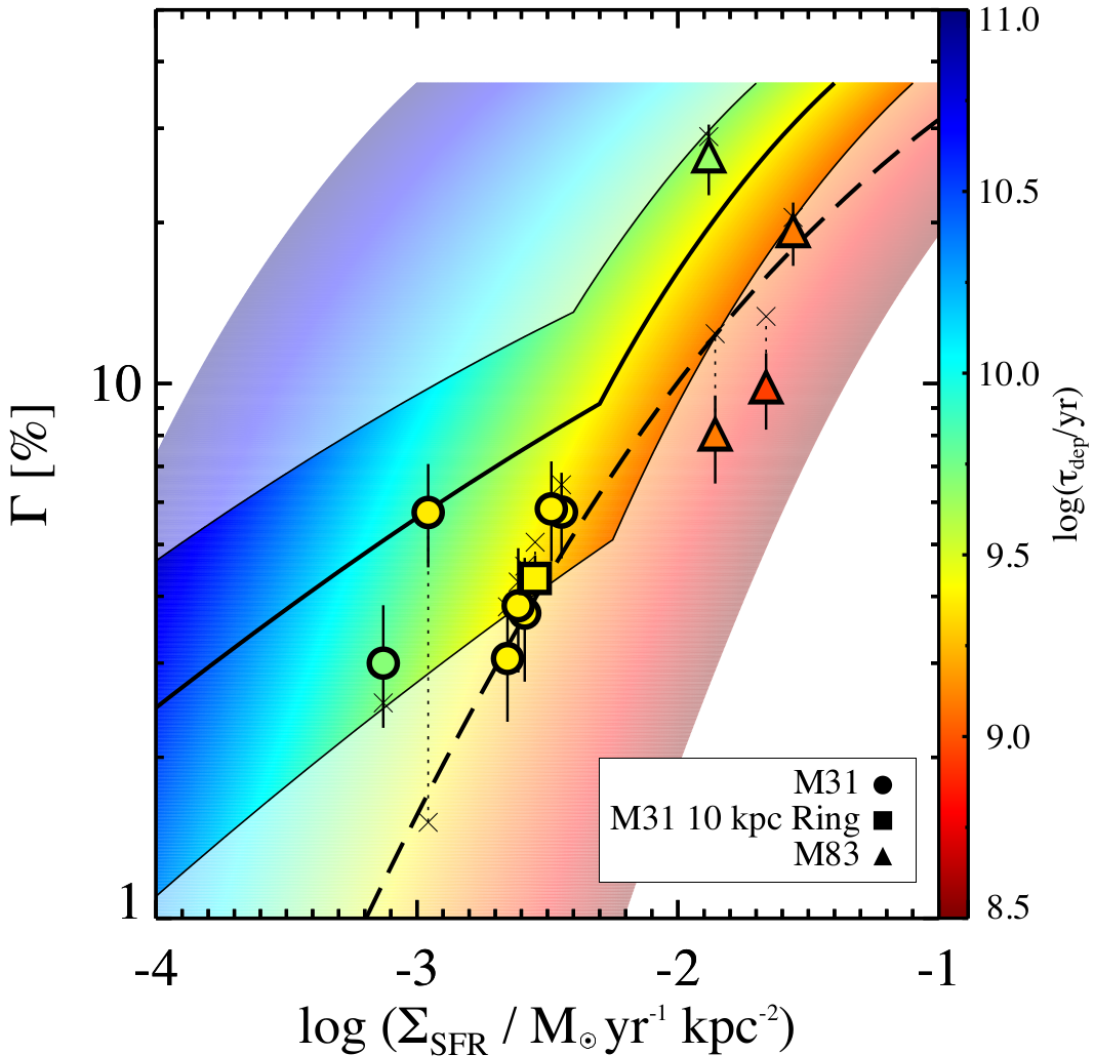

Figure 4.10 Spatially resolved Γ measurements for M31 (circles) and M83 (triangles). We plot Γ predictions from the Kruijssen (2012) model assuming the new spatially-resolved star formation law derived from the Bigiel et al. (2008) observations. Model predictions are plotted for $\Sigma_{\rm gas} < 100\ M_{\odot}\ {\rm pc}^{-2}$ using a color coding for $\tau_{\rm dep}$, ranging from log(yr) of 8.5 (red) to 11 (violet). We highlight the portion of model parameter space that is consistent with $\tau_{\rm dep}$ observations, as defined in Figure 4.9: the thin solid lines represent the upper and lower envelopes to the observed range; the thick solid line represents the median relation. The original galaxy-integrated Γ relation from Kruijssen (2012) is plotted as a dashed line. Data points for M31 and M83 are color coded according to observed $\tau_{\rm dep}$, where good agreement between the model and observations is represented by a color match between the data point and the underlying models. Positions in the plot representing a region's predicted Γ given observed values of $\Sigma_{\rm SFR}$ and $\tau_{\rm dep}$ are marked with Xs and connected to the corresponding observations by dotted lines.

formation law from N=1.4 to N=1.0. Also, as $\tau_{\rm dep}$ increases, the predicted Γ relation moves to the left in Figure 4.10 to lower $\Sigma_{\rm SFR}$ values. Finally, the transition from H_2-dominated to HI-dominated star formation creates a break in the predicted relation. As a result, we expect the distribution of Γ observations at low $\Sigma_{\rm SFR}$ to flatten due to the dramatic increase in $\tau_{\rm dep}$ at low $\Sigma_{\rm gas}$. In contrast to predictions from the steeply declining fiducial Γ relation, we expect these low density environments with $-4 < \log\ (\Sigma_{\rm SFR}\ /\ M_\odot\ {\rm yr}^{-1}\ {\rm kpc}^{-2}) < -3$ to form a small percentage of their stars (1–5%) in long-lived star clusters as opposed to negligible values of <1%.

Next, we compare the new theoretical relations for cluster formation efficiency to the observations of M31 and M83. We observed in Figure 4.9 that most of the M31 and M83 analysis regions had fast depletion times and fell in close proximity to the upper envelope of the observed range. Accordingly, a majority of the Γ observations lie near the upper $\Sigma_{\rm SFR}$ limit of the models in Figure 4.10, on the right/lower side of the predicted range. Moving beyond general sample-wide statements, the availability of $\tau_{\rm dep}$ measurements for each analysis region (represented by the color coding assigned to each point in Figure 4.9) allows us to individually evaluate the agreement between observations and theoretical predictions in all three relevant parameters (Γ, $\Sigma_{\rm SFR}$, and $\tau_{\rm dep}$). In the case of good agreement, we expect the data points in Figure 4.10 to match the color of the models located at the same position in the plot.

The most notable success of this $\tau_{\rm dep}$-dependent modeling is the fantastic agreement found for the 10 kpc ring regions. We find that the 0.2 dex variation in $\Sigma_{\rm SFR}$ leads precisely to the factor of ∼2 variation in Γ predicted by the model, following the steep line of constant $\tau_{\rm dep}$ in the plot (tracking the band of yellow models). The five individual regions, which we assume share a single characteristic value of $\tau_{\rm dep}$, form a tight sequence that is well-explained by the model at fixed $\tau_{\rm dep}$.

All but one data point in the combined M31/M83 dataset shows good agreement with the underlying $\tau_{\rm dep}$ model value, showing good overall consistency between the theoretical predictions and the observations for all three parameters: Γ, $\Sigma_{\rm SFR}$, and $\tau_{\rm dep}$. Although two of the M83 observations lie 2–3σ below their Γ predictions, we note that Adamo et al. (2015) use a conservative cluster catalog selection criteria (rejecting questionable "Class 2"

candidates) and state that their reported values could be low due to this cut. The only seriously discrepant observation is the inner disk region of M31, for the reasons discussed in Section 4.5.1.

One final point of discussion concerns our use of the default set of model parameters from Kruijssen (2012). Echoing the caveats outlined in that work, we remind the reader that the model predictions presented here are based on a set of assumptions that generally describe standard conditions in galaxy star forming environments. As the case of the inner disk region of M31 shows, specifying precise values for the full set of input parameters (i.e., $\Sigma_{\rm gas}$, Toomre Q, Ω, or even ϕ_P) or making other modifications to the standard set of assumptions (e.g., changing the prescription for stellar feedback timescales and mechanisms) could vary the model predictions by a factor of $\sim$2 or more. See Section 3.4, 7.1, and Appendix C in Kruijssen (2012) for further discussion of the sensitivity of Γ predictions due to model assumptions.

In conclusion, we find that adopting a $\tau_{\rm dep}$-dependent star formation relation provides improved predictions and interpretations of Γ modeling from Kruijssen (2012) for spatially resolved observations. In addition to successfully explaining Γ observations of M31 and M83, the new conversion from $\Sigma_{\rm gas}$ to $\Sigma_{\rm SFR}$ makes clear predictions for future Γ observations in low intensity, HI-dominated star forming environments.

4.5.3.3 A Starburst Γ Relation: Short $\tau_{\rm dep}$ or Radiative Feedback?

To close discussion of our $\tau_{\rm dep}$-dependent Γ relations, we expand beyond normal galaxies to consider more intense, starburst environments. The non-linear slope of the Schmidt-Kennicutt relation (N=1.4) indicates that for galaxy-integrated scales, $\tau_{\rm dep}$ decreases as $\Sigma_{\rm gas}$ increases. This result is rather intuitive, suggesting that gas tends to create stars more efficiently at higher gas densities (though remaining constant per free-fall time; Krumholz et al. 2012). These high star formation efficiencies are found for (U)LIRGs and other starbursting galaxies, but also in the dense central regions of otherwise normal galaxies.

From the current set of observations, three studies have placed constraints on the fraction of stars born in long-lived clusters within starburst environments ($\Sigma_{\rm SFR} > 0.1\ M_\odot\ {\rm yr}^{-1}$

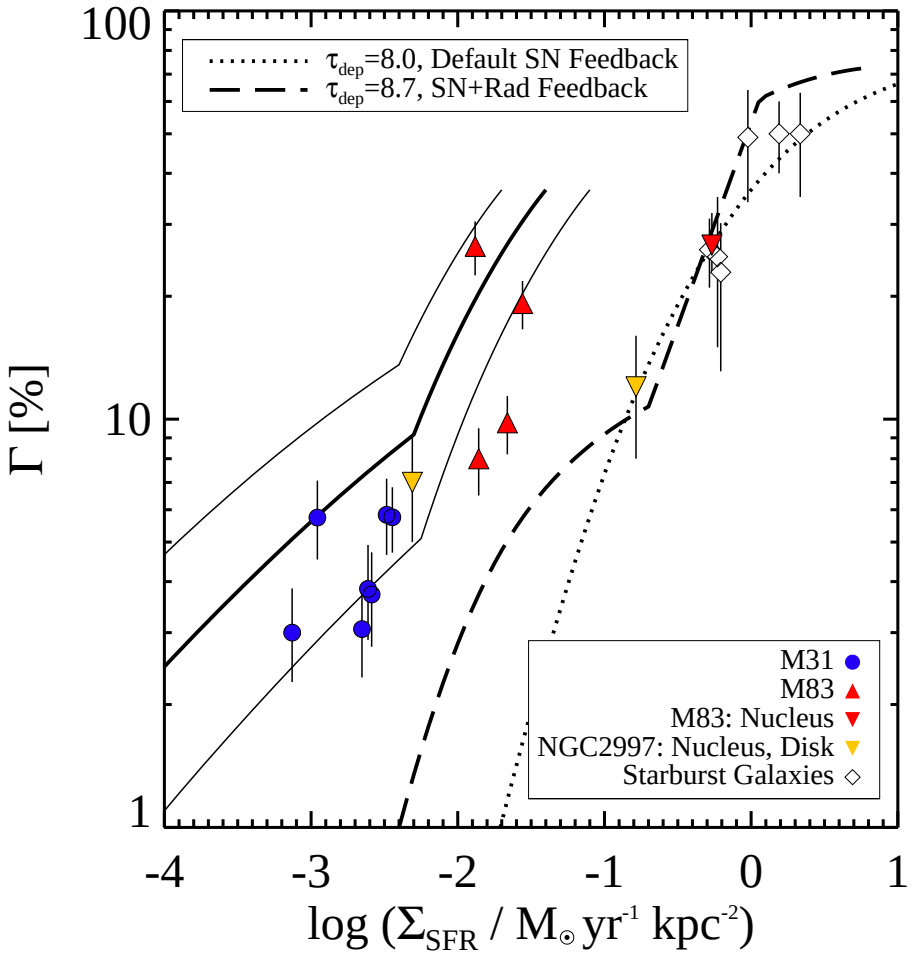

Figure 4.11 Revisiting starburst Γ observations considering a $\tau_{\rm dep}$-dependent star formation law. We plot spatially resolved measurements from M31 (circles; this work), M83 (upward triangles; Adamo et al. 2015; Goddard et al. 2010), and NGC2997 (downward triangles; Ryon et al. 2014), as well as integrated measurements from six starburst galaxies (Goddard et al. 2010; Adamo et al. 2011). Theoretical Γ relations from Kruijssen (2012): $\log(\tau_{\rm dep}/{\rm yr})$=8.0 with nominal SN-only feedback (dotted line), and $\log(\tau_{\rm dep}/{\rm yr})$=8.7 with an alternative SN+Rad combined feedback prescription (dashed line). The predicted parameter space for Γ observations in normal galaxies are represented by solid lines, as in Figure 4.10. We observe that regions with $\Sigma_{\rm SFR} > 0.1\ M_\odot\ {\rm yr}^{-1}{\rm kpc}^{-2}$ are consistent with our new Γ predictions when $\tau_{\rm dep} \sim 100$ Myr. Altering the model's stellar feedback prescription also leads to a satisfactory fit for a longer $\tau_{\rm dep}$ of ∼500 Myr.

kpc^{-2}): Goddard et al. (2010) analyzed NGC3256 and the nuclear region of M83, Adamo et al. (2011) analyzed a sample of five blue compact galaxies, and Ryon et al. (2014) analyzed the nuclear region of NGC2997. We plot these observations along with the spatially resolved M31 and M83 Γ results in Figure 4.11. Interestingly, we find that while values of Γ observed at lower Σ_{SFR} ($<$0.1 $M_{\odot}$ yr^{-1} kpc^{-2}) are well-explained by Γ relations with $\log(\tau_{\mathrm{dep}}/\mathrm{yr})$ between $\sim$9–10, the starburst environments appear well-matched with a Γ relation with $\log(\tau_{\mathrm{dep}}/\mathrm{yr})$ of 8.0 (100 Myr). This remarkable agreement between the predicted Γ relation and the observations for starburst environments would be an intriguing success for the Kruijssen (2012) theoretical framework if the value of τ_{dep} proves to be consistent with the intrinsic value for these systems.

In the case of the M83 nuclear region, we can compare the observed τ_{dep} to the short timescale of 100 Myr predicted by the model Γ relation. The Σ_{SFR} and Σ_{gas} measurements (as reported in Adamo et al. 2015) yield τ_{dep} of $\sim$1 Gyr, which is a factor of 10 larger than the prediction. This τ_{dep} measurement would be worth revisiting as it is based on low spatial resolution CO observations (Lundgren et al. 2004) and a SFR derived from a Hα luminosity. In addition, variation of the CO-to-H_2 conversion factor for the centers of galaxies could lead to an overestimation of $\Sigma_{\mathrm{H_2}}$ and τ_{dep} (Sandstrom et al. 2013; Leroy et al. 2013b).

If the observational τ_{dep} constraints for these high Σ_{SFR} environments are in fact longer than the 100 Myr value predicted in Figure 4.11, an alternative way to reproduce a steep Γ relation at high Σ_{SFR} is to include radiative pressure as an additional stellar feedback process. As Kruijssen (2012) explored in their Appendix C, adding (or substituting) radiative feedback to the nominal supernova feedback prescription produces a Γ-Σ_{SFR} relation with a different shape than obtained using supernova feedback alone — particularly at high Σ_{SFR}. Using an alternative set of assumptions allowed by the Kruijssen (2012) code, we calculate a Γ relation assuming combined feedback from supernova and radiative pressure (SN+Rad) and a characteristic value of $\log(\tau_{\mathrm{dep}}/\mathrm{yr})$=8.7 (500 Myr) and plot this relation for comparison in Figure 4.11. The plot shows that this alternative theoretical solution also agrees well with the distribution of starburst Γ observations. The downside to this solution is that the relation predicted for SN+Rad feedback does not agree with Γ and τ_{dep} observations in

non-starburst regions. Therefore, some tuning of the model would be required, such that the contribution from radiative feedback would need to increase as a function of $\Sigma_{\rm SFR}$.

Further observational work could help better understand cluster formation efficiency in starburst environments. Fortunately these two proposed scenarios predict $\tau_{\rm dep}$ that differ by a factor of $\sim$5, which should produce an observationally detectable difference. Additionally, further observations of Γ behavior at high $\Sigma_{\rm SFR}$ would provide a more complete picture of the variety and characteristics of long-lived cluster formation in starburst systems.

4.5.4 Possible Influence of Cluster Dissolution on Γ *Determinations*

The agreement between Γ determinations obtained for adjacent age ranges of 10–100 Myr and 100–300 Myr provide compelling evidence that cluster dissolution has little effect on our derived values of Γ. In the case of $\mathrm{d}N/\mathrm{d}M \propto t^{-1}$ cluster dissolution, we would expect the latter time bin to show a factor of $\sim$10 smaller Γ integrated over the two age ranges, but find no such difference.

While we conclude that cluster dissolution operates on sufficiently long timescales to leave our Γ determinations unaffected, this does not rule out significant cluster destruction occurring on longer timescales. Using a longer time baseline, we explore cluster dissolution timescales through detailed modeling of the 2D cluster age and mass distribution for M31 in a separate work (M. Fouesneau, in preparation).

One piece of observational evidence from our study points to the possibility of cluster dissolution acting on timescales <300 Myr: the anomalously low Γ constraint for the 100–300 Myr age range in the OB30/31 analysis region (Region 2a) of $2.3^{+0.9}_{-0.8}$%. This value is significantly discrepant with the 10–100 Myr result of $5.8^{+1.1}_{-1.0}$% and the theoretical expectation of $\sim$5%. While this single data point does not justify changing our assumptions with regard to the effects of dissolution, it is interesting to hypothesize the possible reasons for this single outlier.

Before we consider possible reasons for this discrepancy, there are two other notable pieces of evidence to keep in mind. First, this region lies at the leading end of a continuous string of star forming regions that make up the northeastern portion of the 10 kpc star

forming ring. Second, the SFH results show that there is significant star formation activity throughout the full 10–300 Myr age range within and in close proximity in this spatial region. Possible hypotheses to explain the low Γ value in the OB30/31 region for the 100–300 Myr time bin include:

Errors on cluster and total stellar mass constraints? — Completeness results derived for the field star photometry and for the cluster catalog show no significant difference between this region and other nearby 10 kpc ring regions. We find no reason to doubt the validity of the observational results for this particular region.

Cluster destruction via tidal disruption by dense gas? — The presence of high intensity star formation in an epoch that immediately follows a previous episode of star formation raises the possibility that the local gas environment was more chaotic than typically found within M31. It is possible that clusters formed during the earlier epoch were dispersed at a rate that outpaced other regions due to a "cruel cradle"-like mechanism.

External Interactions — Multiple studies have explored galaxy interactions with M32 and M33 (Block et al. 2006; McConnachie et al. 2009; Dierickx et al. 2014) and accretion events (producing the giant southern stream; Fardal et al. 2008) in M31's recent history. While still under investigation, the consequences of external perturbations on the M31 disk could provide a viable, although exotic, explanation.

Stochastic Variation — Finally, while the $\Sigma_{\rm SFR}$ value is high, we cannot rule out the possibility that the unexpectedly low Γ measurement is simply due to a random statistical fluctuation of sampling from the cluster mass function.

Overall, the OB30/31 region provides impetus for us to explore how cluster dissolution would imprint its signal onto the Γ results presented here, but we find no compelling evidence showing that cluster dissolution has any significant effect on our 10–100 Myr results.

4.6 Summary and Future Directions

We close this work with a summary of the major contributions of this study, followed by a brief discussion about the broader implications of Γ constraints and modeling and future directions for observational and theoretical progress. The results of our observational Γ work in M31 are summarized here:

1. We combine high quality cluster and field star formation history constraints from the PHAT survey, include detailed cluster and stellar completeness information in our calculations, and utilize a probabilistic modeling approach to provide the most detailed analysis to-date of the fraction of stellar mass born in long-lived star clusters.

2. We make spatially resolved measurements of Γ across the disk of M31 and find values between 3-6%. Our study significantly extends the range of environments for which observations of long-lived cluster formation efficiency have been obtained.

3. We find excellent agreement between our Γ observations and theoretical predictions from Kruijssen (2012). The lone exception is the gas-poor inner disk region. Here the physical conditions depend on an unusually dominant mid-plane pressure contribution from the stellar disk and provide a future challenge for model improvements.

4. We employ knowledge about how the star formation law ($\Sigma_{\rm SFR} \propto {\Sigma_{\rm gas}}^N$) changes in the transition from a H_2 to HI-dominated ISM to refine predictions for spatially-resolved Γ observations as a function of $\Sigma_{\rm SFR}$. The new relation flattens at low $\Sigma_{\rm SFR}$, in agreement with observations.

5. We apply our newly-derived $\tau_{\rm dep}$-dependent star formation relation to model Γ observations in starburst environments. We propose an observational test to determine whether the theoretical Γ model predictions using a $\tau_{\rm dep}$=100 Myr star formation relation hold for starburst systems, or whether it is necessary to incorporate radiative feedback into the model for these systems.

Measurements of the fraction of stellar mass that is formed in long-lived star clusters as a function of star forming environment provide useful constraints towards understanding star formation behavior. Following the interpretation of Kruijssen (2012), these star clusters trace the stellar populations that are formed in environments where total star formation efficiencies (integrated over the lifetime of a star forming region, as opposed to per free-fall time) are high enough to produce stellar structures that survive gas expulsion during the

transition out of an initial embedded phase. Particularly when these Γ measurements are combined with a characterization of the natal ISM, these observations paint an interesting picture connecting stellar feedback processes, formation efficiencies, and characteristics of the resulting stellar products.

We have only scratched the surface when it comes to using clusters and the spatial structure of newly formed stars to constrain star formation physics. As pointed out in the review by Krumholz (2014), the theoretical model for Γ from Kruijssen (2012) can only predict the overall percentage of cluster mass locked up in long-lived clusters; it currently lacks the sophistication necessary to predict the mass function of these emergent clusters. Work by Hopkins (2013) makes headway in predicting the spatial clustering of stars, therefore making predictions for the shape of the cluster mass function, but it also falls short of a complete treatment of cluster formation that accounts for the influence of stellar feedback on cluster outcomes. In concurrence with Krumholz (2014), we conclude that a theoretical understanding of long-lived cluster formation would benefit from the combination of theories that not only predict the overall fraction of bound mass, but also the distribution of that mass into the discrete systems we observe. Once such a model is derived, we hope to use it as a tool to constrain stellar feedback and other star formation processes.

We are only beginning to utilize the full potential of Γ-based star formation studies. As we discussed in Section 4.5.3, follow-up observations to characterize the star forming ISM in starburst systems would allow the differentiation between feedback mechanisms within the Kruijssen (2012) Γ model framework. Also, the growing number of well-constrained Γ results span a wide variety of star forming environments and cluster formation activity. However, as we saw in our study, sometimes it is exceptional regions like the inner disk of M31 that contribute significantly toward testing theoretical models. Nevertheless, there is plenty of rewarding observational work still to be done.

4.7 Supplementary Information

4.7.1 Calculating Average Surface Densities

Previous Γ studies typically adopted a simple approach for deriving SFR and gas surface densities ($\Sigma_{\rm SFR}$ and $\Sigma_{\rm gas}$) that used a single galaxy-wide aperture and measured area-averaged quantities. There are a number of weaknesses in this approach. First, the subjective definition of an outer boundary directly affects derived surface density values. Adopting uniform definitions and procedures can serve to reduce these biases and uncertainties (e.g., see discussion in Section 3.1.2 in Adamo et al. 2011), but defining a outer limit for an inherently continuous distribution is difficult. Second, area-averaged quantities assume a uniform intrinsic distribution, whereas star formation is inherently clumpy and irregular forming structures such as bars, arms, and rings.

The excellent spatially resolution available for all relevant M31 datasets allows us to compute surface densities using a deprojected 0.5 kpc^2 measurement kernel (with deprojected radius of ∼100 arcsec). We calculate mass-weighted averages to account for filling factor variations in the gas and SFR distributions, and explore how mass-weighted averages compare to area-weighted metrics used in previous Γ studies.

We use the outer disk region to illustrate the difference between mass and area-weighted $\Sigma_{\rm SFR}$ measurements. In the left panel of Figure 4.12, we compare the distribution of 0.5 kpc^2 smoothed, unweighted (thus, area-weighted) $\Sigma_{\rm SFR}$ measurements with the distribution of mass-weighted measurements. The thick vertical lines denote the area-weighted and mass-weighted mean values; we report mass-weighted mean values as our primary $\Sigma_{\rm SFR}$ metric. While the ∼0.4 dex relative difference in $\Sigma_{\rm SFR}$ for the outer disk region is the biggest weighting-dependent difference among the seven M31 analysis regions (due to the relatively high contrasts between ring/arm and interarm/outskirts environments), a similar offset exists for all regions; we visualize these offsets in the right panel of Figure 4.12.

It is also important to acknowledge that each of the M31 analysis regions contains a range of $\Sigma_{\rm SFR}$ values. We compute the interquartile range (from the 25th to 75th weighted percentile; thick line segments in right panel of Figure 4.12) of the $\Sigma_{\rm SFR}$ distribution for each analysis region, finding values from 0.2–0.4 dex. While this is an unsurprising consequence

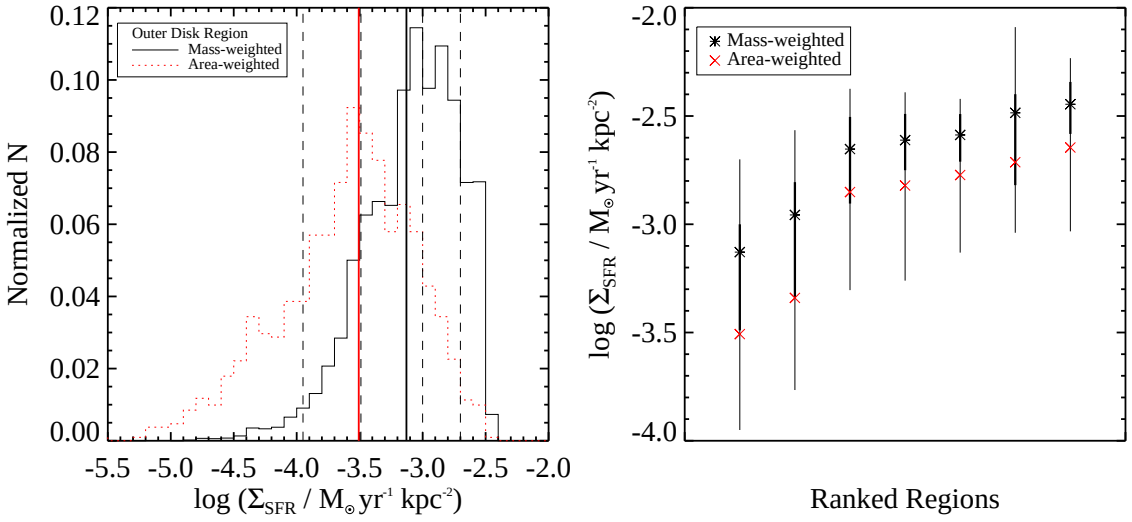

Figure 4.12 Left: We compare the $\Sigma_{\rm SFR}$ distributions derived with mass-weighting (black solid histogram) and without (area-weighted; red dotted histogram) for the outer disk analysis region. Thick vertical lines denote weighted (black) and unweighted (red) mean $\Sigma_{\rm SFR}$ values, and vertical dashed lines denote 5th, 25th, 75th, and 95th percentiles of the weighted distribution. Right: For each analysis region (ordered according to increasing $\Sigma_{\rm SFR}$) we plot the mass-weighted mean $\Sigma_{\rm SFR}$ values (black asterisk) and its associated 25th–75th percentile (thick black line) and 5th–95th percentile (thin black line) ranges. We compare these distributions to the area-weighted mean values (red x), show the systematic discrepancy between these estimates.

of the clumpy, varying nature of star formation, it is important to keep in mind that characteristic mean $\Sigma_{\rm SFR}$ values represent differences between broad underlying distributions of star formation intensities.

In addition to calculating $\Sigma_{\rm SFR}$, we also use a mass-weighted methodology to calculate robust measurements of $\Sigma_{\rm HI}$, $\Sigma_{\rm H_2}$, and $\Sigma_{\rm gas}$. M31's gas phase is dominated by a neutral HI component, which has shown to have a low sub-kpc to kpc clumping factor (Leroy et al. 2013a). Therefore, measuring $\Sigma_{\rm gas}$ using a 0.5 kpc^2 kernel provides accurate characterizations of intrinsic, HI-dominated total gas densities. This is not the case, however, when considering molecular gas and $\Sigma_{\rm H_2}$ alone. High-resolution (20 pc) molecular gas observations obtained using CARMA (A. Schruba, in preparation) reveal molecular gas structures on the scale of 10–100 pc. Therefore, one should take care when interpreting $\Sigma_{\rm H_2}$ values

calculated in this work, as these densities are likely to be significantly diluted.

4.7.2 Γ *Results from the Literature*

As discussed in Section 4.5.2, we assemble a compilation of Γ results from the literature to place our M31 results in a broad context. We present these literature results in Table 4.4. We are not the first to pursue this task (e.g., see the recent compilation in Appendix B Adamo et al. 2015), however it is important to make clear the choices we have made in assembling (and in some cases, transforming) this set of results. In particular, we make an effort to highlight where our choices differ from others.

We sought to utilize the highest quality results for individual galaxies when compiling this dataset. In the case of M83 (NGC5236), we prefer the recent results from Adamo et al. (2015) due to its near-complete coverage of the galaxy, surpassing previous constraints from Silva-Villa & Larsen (2011) and Silva-Villa et al. (2013). For the LMC, we utilize the recent Γ result from Baumgardt et al. (2013) of 15%. This work makes use of updated cluster constraints from a compilation of sources for clusters with $\log(\mathrm{Mass}/M_{\odot}) > 3.7$. It also updates an far-IR luminosity-based integrated SFR estimate from Larsen & Richtler (2000) with a CMD-based total star formation constraints from Harris & Zaritsky (2009). Inferred masses in clusters and total stars and the derived Γ increased significantly with respect to Goddard et al. (2010) ($\Gamma = 5 \pm 0.5\%$), but agrees with the 10–20% derived by Maschberger & Kroupa (2011) who use the same recent star formation history constraints. We note that Baumgardt et al. (2013) assume a power law with an index of -2.3 versus the traditional -2 for their cluster mass function extrapolation down to 100 $M_{\odot}$. An extrapolation using an index of -2 would give a result that was a factor of $\sim$0.7 smaller. Also note that Baumgardt et al. (2013) provides no accompanying Σ_{SFR} value; we adopt the area normalization (79 kpc^2) used previously by Goddard et al. (2010) to normalize the SFR (0.29 $M_{\odot}$ yr^{-1}).

In contrast to eliminating duplicate Γ observations made on a common galaxy-wide scale, spatially resolved Γ determinations provide unique constraints we do not want to ignore. We tabulate individual spatially resolved Γ constraints of the nuclear region of M83 from Goddard et al. (2010), as well as separate disk and nuclear measurements of NGC2997 from

Table 4.4. Γ Results from the Literature

Galaxy	Σ_{SFR} ($M_\odot$ yr^{-1} kpc^{-2})	Γ (%)	Reference
NGC1569	0.03	13.9 ± 0.8	Goddard et al. 2010
NGC3256	0.62	$22.9^{+7.3}_{-9.8}$	Goddard et al. 2010
NGC6946	0.0046	$12.5^{+1.8}_{-2.5}$	Goddard et al. 2010
SMC	0.001	$4.2^{+0.2}_{-0.3}$	Goddard et al. 2010
Milky Way	0.012	$7.0^{+7}_{-3.0}$	Goddard et al. 2010
ESO338	1.55	50.0 ± 10.0	Adamo et al. 2011
Haro 11	2.16	50.0^{+13}_{-15}	Adamo et al. 2011
ESO185-IG13	0.52	26.0 ± 5.0	Adamo et al. 2011
MRK930	0.59	25.0 ± 10.0	Adamo et al. 2011
SBS0335-052E	0.95	49.0 ± 15.0	Adamo et al. 2011
NGC45	0.00101	17.3	Silva-Villa & Larsen 2011
NGC1313	0.011	9.0	Silva-Villa & Larsen 2011
NGC4395	0.00466	2.6	Silva-Villa & Larsen 2011
NGC7793	0.00643	9.8	Silva-Villa & Larsen 2011
NGC4449	0.04	9.0	Annibali et al. 2011
LMC	0.00366	15.0	Baumgardt et al. 2013
NGC2997	0.0094	10.0 ± 2.6	Ryon et al. 2014
IC10	0.03	4.2	Lim & Lee 2015
M83 (0.45–4.5 kpc)[a]	0.019	12.5 ± 1.4	Adamo et al. 2015
ANGST Dwarfs (<10 Myr)	3e-5–1e-2	5.0	Cook et al. 2012
ANGST Dwarfs (<100 Myr)	3e-5–1e-2	1.65	Cook et al. 2012
M83 (Nuclear)	0.54	$26.7^{+5.3}_{-4.0}$	Goddard et al. 2010
NGC2997 (Disk)	0.0049	7.0 ± 2.0	Ryon et al. 2014
NGC2997 (Nuclear)	0.164	12.0 ± 4.0	Ryon et al. 2014
M83 (0.45–2.3 kpc)[a]	0.013	26.5 ± 4.0	Adamo et al. 2015
M83 (2.3–3.2 kpc)[a]	0.028	19.2 ± 2.6	Adamo et al. 2015
M83 (3.2–3.9 kpc)[a]	0.022	9.8 ± 1.6	Adamo et al. 2015
M83 (3.9–4.5 kpc)[a]	0.014	8.0 ± 1.5	Adamo et al. 2015

[a]We utilize the 10–50 Myr Γ results from Adamo et al. (2015).

Ryon et al. (2014), but these constraints do not appear in Figure 4.8, naturally, due to their sub-galaxy scale. However, we decide to omit these results from our presentation of spatially resolved results in Figure 4.10 due to a lack of available ISM constraints in the case of Ryon et al. (2014), and due to the <10 Myr age limitation of the Goddard et al. (2010) result.

From the Cook et al. (2012) dwarf galaxy work, we opt to use their "binned" Γ results. For each of the age ranges considered (4–10 Myr and 4–100 Myr), the authors combine the set of observed galaxies with $-4.5 < \log \Sigma_{\rm SFR} < -2.0$ into a single meta-galaxy. This calculation serves to alleviate the problem of small numbers of clusters per individual galaxy (leading to large Γ uncertainties), and to fold in galaxies that independently can only provide upper limit constraint on Γ.

We note that Γ result for NGC4449 from Annibali et al. (2011), quoted for ages <10 Myr, depends completely on the inclusion or exclusion of the massive nuclear super star cluster; this single system hosts >70% of the cluster mass considered in the Γ calculation. Similar to the behavior seen in the Cook et al. (2012) results, this galaxy further demonstrates that the stochastic nature of star formation in dwarf galaxies can lead to large variations in the derived result. Further, the Γ calculation in this work uses a mass function extrapolation assuming a power law form with -2 slope, down to a lower mass limit of 1000 $M_{\odot}$. A correction factor of $\sim$1.4 could be applied to bring the data in line with the standard $M_{\rm min}$ =100 $M_{\odot}$ assumption, but we opt to tabulate and plot the work's original values.

We utilize the "P1" results from Silva-Villa & Larsen (2010, 2011) that do not include dissolution modeling, as opposed to their mass independent destruction (MID) or mass dependent destruction (MDD) constraints. These results, which were also used by Cook et al. (2012), provides a better match to the model-independent, empirical approach of the other studies with which we compare. In addition, we note that these results were calculated using a mass function extrapolation assuming a Schechter function with $M_c = 2 \times 10^5$ $M_{\odot}$down to a lower mass limit of 10 $M_{\odot}$, which differs from the canonical value of 100 $M_{\odot}$ used in other Γ studies. Similar to the case of NGC4449 discussed previously, a correction factor of $\sim$0.8 could be applied to bring the data in line with the standard $M_{\rm min}$ assumption, but we opt to tabulate and plot the work's original values. In addition, this work utilizes

an age-dependent observational completeness limit for mass function extrapolation, and includes a scaling factor applied to the observed CFR to account for coverage differences between data used for cluster fitting versus that used for total star formation fitting of the field populations.

Finally, we would like to highlight two cases where galaxy-integrated Γ constraints deviate strongly from the observed: IC10 and NGC45. Both of these galaxies are dwarf systems with relatively low integrated SFR, which imply large associated uncertainties. Unfortunately, neither of these results were accompanied by reported uncertainties (although, uncertainty estimates for NGC45 were made available for MID and MDD based results by Silva-Villa & Larsen 2011). Finally, we wish to highlight that Silva-Villa et al. (2013) mentions the possibility that a number of ancient massive globular cluster systems were assigned integrated light-based ages that erroneously placed them in the 10–100 Myr range used to estimate Γ. This case serves as an example that, particularly in the case of small numbers of clusters, errors stemming from a variety of sources (many of which are not accounted for in uncertainty calculations) can contribute to the large scatter in reported Γ results.

4.7.3 Representing Observational Constraints on a Spatially Resolved Star Formation Law

Here we report the parameter space limits we adopt in our characterization of the Bigiel et al. (2008) dataset using a series of three two-component broken power laws, as plotted in Figure 4.9 and discussed in Section 4.5.3. The median relation is defined to agree with the most recent star formation law ($\Sigma_{\rm SFR} \propto {\Sigma_{\rm gas}}^N$) results from Leroy et al. (2013b) in the molecular-dominated high density regime ($N = 1.0$), and track the transition to HI-dominated star formation environments using a steeper slope ($N = 3.3$) for $\Sigma_{\rm gas} \lesssim 10\ M_\odot$ pc^{-2}. We define this median relation as:

$$\frac{\Sigma_{\rm SFR}}{M_\odot\ {\rm yr}^{-1}\ {\rm kpc}^{-2}} = \begin{cases} 1 \times 10^{-3.4} \left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right), & \text{if } 1.0 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) < 2.0 \\ 1 \times 10^{-5.93} \left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right)^{3.3}, & \text{if } 0.3 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) \leq 1.0. \end{cases} \tag{4.12}$$

We bracket the median relation with an upper and lower envelope that are chosen to reproduce the spread in $\tau_{\rm dep}$ observed by Bigiel et al. (2008). We define the upper envelope

as:

$$\frac{\Sigma_{\rm SFR}}{M_\odot\ {\rm yr}^{-1}\ {\rm kpc}^{-2}} = \begin{cases} 1\times10^{-3.1}\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right), & \text{if } 1.0 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) < 2.0 \\ 1\times10^{-5.055}\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right)^{3.3}, & \text{if } 0.3 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) \leq 1.0. \end{cases} \tag{4.13}$$

The lower envelope is defined as:

$$\frac{\Sigma_{\rm SFR}}{M_\odot\ {\rm yr}^{-1}\ {\rm kpc}^{-2}} = \begin{cases} 1\times10^{-3.7}\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right), & \text{if } 1.0 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) < 2.0 \\ 1\times10^{-6.69}\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right)^{3.3}, & \text{if } 0.3 \leq \log\left(\frac{\Sigma_{\rm gas}}{M_\odot {\rm pc}^{-2}}\right) \leq 1.0. \end{cases} \tag{4.14}$$

The median, upper envelope, and lower envelope relations have inflection points ($\log \Sigma_{\rm gas}$, $\log \Sigma_{\rm SFR}$) at (1.1, -2.3), (0.85, -2.25), and (1.3, -2.4), respectively. We define these relations over the range of $\Sigma_{\rm gas}$ parameter space spanned by the observations: $0.3 \leq \log(\Sigma_{\rm gas}/M_\odot {\rm pc}^{-2}) \leq 2.0$. Outside this range of total gas densities, observations tentatively point to qualitatively different behavior. For gas densities <2 $M_\odot$ pc^{-2}, Bigiel et al. (2010) presents evidence for a flattening of the star formation law slope that hints at an asymptotic $\tau_{\rm dep}$ value of $\sim10^{11}$ yr. For gas densities >100 $M_\odot$ pc^{-2}, a starburst mode of star formation likely prevails (e.g., see Daddi et al. 2010), characterized by $\tau_{\rm dep}$ on the order of $\sim10^7$–10^8 yr and a slope of $N\sim$1.3–1.4. Therefore, extrapolation of this relation beyond our adopted limits is not advised.

Chapter 5

THE HIGH-MASS TRUNCATION OF THE CLUSTER MASS FUNCTION AND ITS CORRELATION WITH STAR FORMATION RATE DENSITY

We find that the mass function of young star clusters observed by the Panchromatic Hubble Andromeda Treasury (PHAT) survey in M31 is truncated for masses $> 10^4$ $M_\odot$. Adopting a Schechter function parameterization, we find a statistically significant exponential truncation with a characteristic mass of $M_c = 8.5^{+2.8}_{-1.8} \times 10^3$ $M_\odot$ for a sample of 840 clusters with ages between 10–300 Myr. The clear detection of a truncation in the cluster mass function confirms the work of several previous studies, but the M_c value we measure for PHAT is more than an order of magnitude lower than any previous measurement. By combining our M31 results with previous work, we find a strong dependence of the cluster truncation mass on the star formation rate surface density, with $M_c \propto {\Sigma_{\rm SFR}}^{\sim 1.3}$. We also explore the possibility that the M_c–$\Sigma_{\rm SFR}$ relation derived for young clusters might also apply to globular cluster systems. Under this assumption, globular cluster mass functions could be used to constrain the star formation rate density of their galaxy hosts in the early universe.

5.1 Introduction

Numerous studies in the astronomical literature have shown that the mass function of young star clusters is well characterized by a power law ($M \propto M^\alpha$) with an index of α=-2.0 ± 0.3 over a wide range of masses (e.g., Zhang & Fall 1999; Gieles et al. 2006a; Portegies Zwart et al. 2010; Fall & Chandar 2012). There is on-going debate, however, whether the behavior of the cluster mass function at its high-mass end is best described by a pure power law distribution (e.g., Chandar et al. 2010b; Whitmore et al. 2010), or a Schechter (1976) function form (e.g., Gieles 2009; Larsen 2009; Bastian et al. 2012a) that includes an exponential, high-mass truncation ($M \propto M^\alpha \exp(-M/M_c)$, where M_c is the

characteristic truncation mass).

Multiple studies have previously presented evidence in support of an exponential mass function truncation, both in the form of direct mass function fits (e.g., Gieles 2009; Adamo et al. 2015) and indirect modeling of the few most massive and most luminous clusters (e.g., Larsen 2009; Bastian et al. 2012a). Together, these studies find evidence for truncations with M_c of $\sim 10^5$ $M_\odot$ in normal star forming galaxies, and larger values ($\sim 10^6$ $M_\odot$) for the interacting, starburst Antennae galaxies. Due to the indirect nature of some of these analyses and the relatively small predicted differences between the two mass function regimes, a definitive consensus on the behavior of the high-mass end of the cluster mass function has not yet emerged.

The study we present here uses data from the *Hubble Space Telescope* (HST) obtained by the Panchromatic *Hubble* Andromeda Treasury survey (PHAT; Dalcanton et al. 2012) to conduct an unprecedented study of the star cluster mass function in the neighboring Local Group galaxy M31. The high spatial resolution imaging provided by HST allows a detailed study of its stellar cluster population, resolving each cluster into a collection of individual stellar members. This level of data quality allows us to derive individual cluster ages and masses by fitting color-magnitude diagrams (CMDs) of individually-resolved member stars, providing stronger constraints than those obtained through multi-band SED fitting for young clusters, and avoids uncertainties caused by stochastic variations in the integrated light of low mass clusters (see e.g., Fouesneau & Lançon 2010; Krumholz et al. 2015). Finally, the uniform imaging from the PHAT survey combined with detailed cluster cataloging and associated completeness estimation provide an unparalleled, well-characterized cluster sample. The catalog's 50% completeness limit for young clusters is $\sim$500–1000 $M_\odot$, providing access to low mass clusters that are undetectable in most extragalactic surveys.

Our uniform, high-quality cluster analysis of a large, well-characterized cluster population allows us to make a direct measurement of the cluster mass distribution through maximum likelihood fitting. This cluster mass function determination for the PHAT survey region in M31 is particularly valuable because it provides a constraint at the low intensity end of the galactic star formation rate density spectrum. Previous observations have focused on galaxies with moderate star formation activity (e.g., M51 and M83 Gieles 2009; Adamo

et al. 2015), as well as high intensity starburst galaxy mergers (e.g., the Antennae Zhang & Fall 1999; Whitmore et al. 2010). Our M31 measurement significantly extends the range of star formation environments analyzed and gives us strong leverage when investigating environmentally-dependent trends.

We begin by introducing the PHAT cluster sample and CMD fitting in Section 5.2. Next, we introduce our probabilistic cluster mass function fitting technique in Section 5.3, and present results in Section 5.4. We compare our M31 results to Schechter mass function measurements for other young cluster systems and discuss the systematic variation of high-mass truncation masses with star formation rate intensity in Section 5.5.1. In Section 5.5.2, we consider the implications that a M_c–$\Sigma_{\rm SFR}$ relation may have on the interpretation of old globular cluster systems. Finally, we close with a summary in Section 5.6.

5.2 Data

We draw our cluster sample from the Andromeda Project (AP) cluster catalog (Johnson et al. 2015b). This catalog provides identifications of 2753 star clusters in the PHAT survey data that span a wide age and mass range. The AP catalog was constructed through visual identification of clusters in optical (F475W, F814W) images by volunteer citizen scientists, facilitated through a website hosted by the Zooniverse organization. The final sample of clusters was selected according to a candidate's weighted frequency of identification, where each image was examined by >80 AP volunteers. We adopt a catalog threshold that maximizes completeness and minimizes the number of contaminates with respect to the expert-derived PHAT Year 1 cluster catalog (Johnson et al. 2012) and its initial 25% survey coverage.

Cluster catalog completeness determinations are provided by a suite of 3000 synthetic cluster tests, which allow us to characterize catalog selection as a function of cluster properties and its variation across the survey footprint. Please see Section 2.2 in Johnson et al. (2015b) for detailed properties of the artificial cluster sample. Each synthetic cluster was injected into an AP search image and subsequently identified and analyzed in the same way as the genuine clusters. We compute survey-averaged 50% completeness limits, taking into account the relative distribution of clusters and spatially-varying completeness through a

star formation and stellar surface density weighted combination of synthetic cluster results. We find a 50% mass completeness of 740 $M_{\odot}$ for the 10–100 Myr age bin and 1080 $M_{\odot}$ for the 100–300 Myr age bin, and parameterize the full completeness as a function of mass using a logistic function.

Individual cluster members were among the 117 million resolved stars measured as part of the PHAT survey. The completeness limits of these stellar catalogs allow the detection of main sequence stars down to $\sim$3 $M_{\odot}$. Please refer to Dalcanton et al. (2012) and Williams et al. (2014) for full details on the survey's crowded field stellar photometry analysis. We extract optical CMDs for each cluster, and obtain constraints on cluster parameters through CMD fitting. We use the MATCH software package to perform maximum-likelihood CMD analysis following techniques described in Dolphin (2002). For cluster fitting, we adopt a M31 distance modulus of 24.47 (785 kpc; McConnachie et al. 2005), a binary fraction of 0.35, a Kroupa (2001) IMF for masses from 0.15 to 120 $M_{\odot}$, and stellar models from the Padova group (Marigo et al. 2008) that include updated low-mass asymptotic giant branch tracks (Girardi et al. 2010). We employ a restrictive prior on [M/H] (from -0.2 to 0.1) to constrain solutions to $\sim Z_{\odot}$ in an effort to match gas phase metallicity observations within M31 (e.g., Zurita & Bresolin 2012). We publish a full catalog of cluster parameter determinations, demonstrate that these results provide reliable results using synthetic cluster tests, and compare the CMD-based fits to those derived from integrated light fitting in L. Beerman et al. (in preparation).

The CMD fitting yields 1249 clusters with best fit ages between 10–300 Myr and masses between 300–20,000 $M_{\odot}$, where the median age uncertainty is 0.2 dex and the median mass uncertainty is 0.04 dex. We plot the derived mass distribution for the cluster sample in the left panel of Figure 5.1. We adopt a 10 Myr lower limit due to the uncertain and subjective nature of cluster identification at younger ages. Fortunately, Gieles & Portegies Zwart (2011) demonstrate that differentiating between long-lived clusters and rapidly expanding, unbound associations becomes well-defined for ages >10 Myr, so we adopt this lower age bound at little expense in terms of integrated star formation and number of clusters. The upper age bound of 300 Myr is based on the limit where CMD fitting becomes dramatically less precise when the MS turnoff drops below our completeness limit.

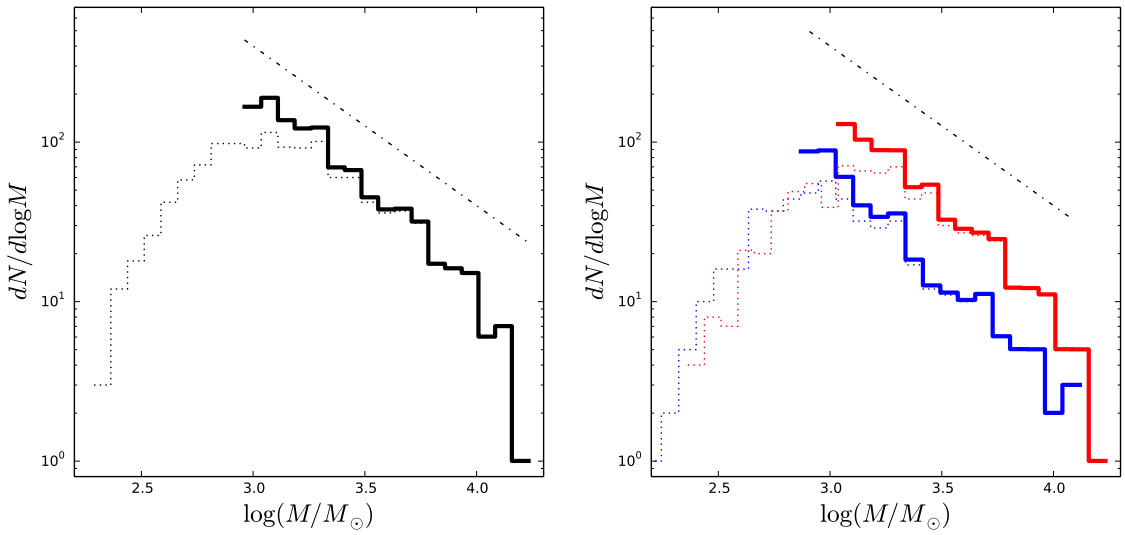

Figure 5.1 Observed mass distributions for PHAT star cluster sample, showing raw observed counts (dotted lines) as well as completeness corrected distributions (thick solid lines). The full 10-300 Myr cluster sample (black) in the left panel shows the same overall shape ($M \propto M^{-2}$; dash-dotted line) as the younger 10-100 Myr population subset (blue) and the older 100-300 Myr population (red) shown in the right panel. We only plot and fit the portion of the distribution that lies above sample's 50% completeness limit (thick solid lines), which occurs at around 10^3 $M_\odot$. The downturn in the raw cluster counts (dotted lines) at low masses is due to catalog completeness.

We observe that the age distribution of clusters is consistent with a near-constant formation history and little or no cluster destruction, in agreement with initial PHAT cluster results presented by Fouesneau et al. (2014). Full modeling of cluster dissolution timescales and its mass dependence is pursued in a separate work (M. Fouesneau, in preparation). While we expect little or no difference in the shape of the mass function with age due to cluster destruction, we will explicitly test for age-dependence of our mass function fitting results by separately analyzing 10–100 Myr and 100–300 Myr age bins. However, Figure 5.1 shows that the mass distributions of the total sample as well as the two separate age bins all show similar power law shapes above the completeness limit at 10^3 $M_\odot$.

5.3 Analysis

We derive mass function constraints using a probabilistic maximum likelihood fitting. Here, we follow a similar approach as Weisz et al. (2013) in their probabilistic modeling of the stellar mass function, but instead for the star cluster mass function. The likelihood for the mass of an observed cluster, M, is expressed as

$$p_{\mathrm{MF_o}}(M|\theta, \mathrm{obs}) \equiv \frac{1}{Z}\, p_{\mathrm{MF}}(M|\theta)\; p(\mathrm{obs}|M), \tag{5.1}$$

where $p_{\mathrm{MF}}(M|\theta)$ represents the cluster mass distribution function, p(obs|M) represents the observational completeness function, and Z represents the normalization required for $p_{\mathrm{MF}_0}(M|\theta, \mathrm{obs})$ to properly integrate to 1. This normalization is given as

$$Z = \int p_{\mathrm{MF}}(M|\theta)\; p(\mathrm{obs}|M)\; dM. \tag{5.2}$$

We adopt a Schechter (1976) functional form for the cluster mass distribution, whose shape is controlled by two parameters, $\theta = \{\alpha, M_c\}$; α is the low mass power law index and M_c is the characteristic mass that defines the exponential high mass truncation. This distribution follows the form

$$p_{\mathrm{MF}}(M|\theta) \propto M^{\alpha} \exp(-M/M_c). \tag{5.3}$$

Please note that the Schechter function simplifies to a simple power law function ($M \propto M^{\alpha}$) in the limit that $M_c \to \infty$. We model the cluster completeness function using a logistic function, parameterized by M_{lim} and a_{lim}, the 50% mass completeness limit and the curve's

maximum slope, respectively. To ensure that we are not too sensitive to the completeness corrections, we impose a step function lower limit to this completeness function (as well as to the sample of cluster masses) restricting our fitting to the mass range that lies above the 50% completeness limit, resulting in the expression:

$$p(\mathrm{obs}|M) = \begin{cases} \left(1 + \exp\left[\frac{-a_{\mathrm{lim}}(M - M_{\mathrm{lim}})}{M_\odot}\right]\right)^{-1}, & M > M_{\mathrm{lim}} \\ 0, & \text{otherwise.} \end{cases} \tag{5.4}$$

We use Bayes' theorem to derive the posterior probability distribution function for the Schechter function fit of our set of N observed cluster masses, $\{M_i\}$, expressed as

$$p(\theta|\{M_i\}, \mathrm{obs}) \propto p_{\mathrm{MF_o}}(\{M_i\}|\theta, \mathrm{obs})\ p(\theta). \tag{5.5}$$

The probability of mass function parameters θ is expressed in terms of the likelihood for a set of observed masses, calculated by taking the product of the individual cluster mass likelihoods:

$$p_{\mathrm{MF_o}}(\{M_i\}|\theta, \mathrm{obs}) = \prod_{i=1}^{N} \frac{1}{Z}\ M_i^{\alpha} \exp(-M_i/M_c)\ p(\mathrm{obs}|M_i), \tag{5.6}$$

where the normalization term becomes

$$Z = \int_{M_{\mathrm{lim}}}^{\infty} M^{\alpha} \exp(-M/M_c)\ p(\mathrm{obs}|M)\ dM. \tag{5.7}$$

We integrate this normalization term numerically during the course of fitting. The term $p(\theta)$ in Eq. 5.5 represents the priors on the Schechter parameters. We adopt uniform top-hat probability distributions that generously cover the range of possible α (-3 to -1) and M_c ($\log M_c/M_\odot$from 3 to 8) parameter values.

We use a Markov Chain Monte Carlo (MCMC) technique to sample the posterior probability distributions of the Schechter mass function parameters. In particular, we use the `emcee`[1] Python package (Foreman-Mackey et al. 2013) and its implementation of an affine invariant ensemble sampler from Goodman & Weare (2010). For our fitting, we use 500 walkers, each producing 600 step chains, of which we discard the first 100 burn-in steps. We report the median value of the marginalized posterior probability distribution for M_c and α, $p(M_c|\{M_i\}, \mathrm{obs})$ and $p(\alpha|\{M_i\}, \mathrm{obs})$, accompanied by a 1σ confidence interval defined by the 16th to 84th percentile range of the marginalized posterior.

[1] `http://dan.iel.fm/emcee/`

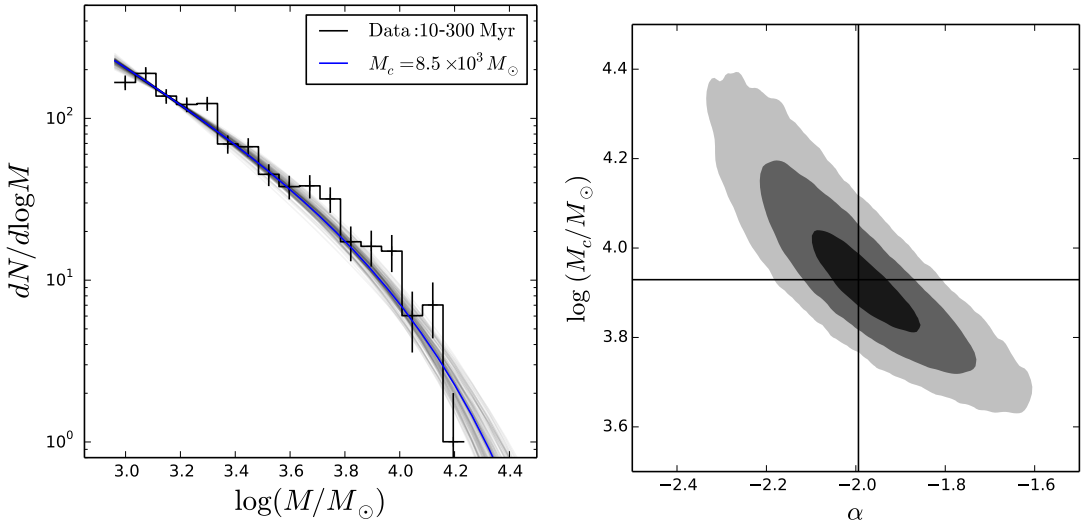

Figure 5.2 Schechter function fitting results for PHAT cluster sample. Left: Histogram shows binned cluster mass results (black). We plot 100 samples from the posterior probability distribution function to show the variance in Schechter fits (gray) best fit model (blue). Right: Two-dimensional posterior constraints on α and M_c, where contours represent 1, 2, and 3σ confidence intervals.

5.4 Results

The fitting results for the 10–300 Myr PHAT young cluster sample are presented in Figure 5.2, derived for 840 clusters that lie above the age-dependent 50% mass completeness limit. In the left panel, we plot Schechter function realizations using pairs of M_c and α values drawn from the posterior PDF, overlaid on a binned version of the observed cluster mass distribution. We emphasize that the maximum likelihood fitting is performed in an unbinned manner, but we use a binned histogram to help visualize and compare the fitting results. We plot the 2D posterior PDF for M_c and α in the right panel of Figure 5.2, showing the covariance in the values of the fitted parameters.

From the marginalized 1D PDFs we find $M_c = 8.5^{+2.8}_{-1.8} \times 10^3\ M_\odot$ ($\log M_c/M_\odot = 3.93^{+0.13}_{-0.10}$) and $\alpha = -1.99 \pm 0.12$. This characteristic Schechter truncation mass is the lowest value ever obtained for a star cluster population, falling an order of magnitude below the nominal $2 \times 10^5\ M_\odot$ value derived for a sample of star forming galaxies by Larsen (2009). The index

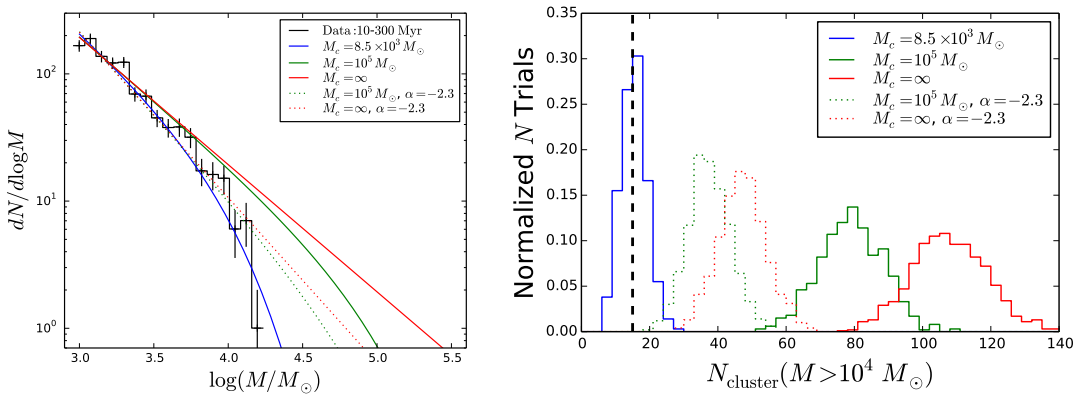

Figure 5.3 A comparison of mass function fits for a variety of M_c and α values. Solid lines represent α=−2.0 and dashed lines represent α=−2.3. Left: A comparison of mass function shapes to the observed 10-300 Myr cluster mass distribution. Model normalizations reflect best fit values based on least squares fitting to the binned observations between $\log(M/M_\odot)$ of 2.9–3.5. Right: Predictions for the number of clusters with $M > 10^4\ M_\odot$ based on the normalizations shown in the left panel. The observed value (N=15) is represented by the vertical dashed line. These panels demonstrate that the observed PHAT cluster mass function is incompatible with a pure power law mass function, or a Schechter function with M_c larger than $\sim 3 \times 10^4\ M_\odot$.

of the low mass slope, however, agrees perfectly with the canonical value of −2, supporting the notion that the cluster mass function for the PHAT survey region in M31 is otherwise rather typical for lower mass clusters. The detection of a significant truncation of the cluster mass distribution in M31, yielding few clusters with masses $>10^4\ M_\odot$, definitively shows that there is a physical limit to the maximum mass of star clusters, and rules out the notion of a universal pure power law cluster mass function.

To demonstrate how significantly the PHAT M31 star cluster mass distribution departs from a pure power law mass function, or even Schechter distributions with larger values of M_c, it is illustrative to compare how the mass function shapes vary for masses $>10^4\ M_\odot$ with varying values of M_c. We compare the observed cluster mass distribution to a set of Schechter realizations (using α=−2.0 & −2.3) in the left panel of Figure 5.3, where each of the models were normalized according to a least-squares fit to the low mass portion of the mass distribution ($2.9 < \log M/M_\odot < 3.5$). While a M_c value of $\sim 10^4$ provides an acceptable fit to the distribution, models with M_c of $10^5\ M_\odot$ or ∞ (the pure power law

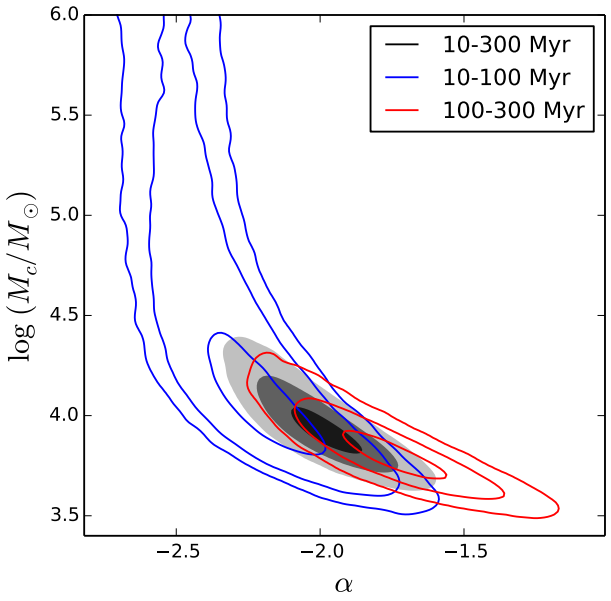

Figure 5.4 Two-dimensional posterior constraints on α and M_c for the 10-100 Myr (blue) and 100-300 Myr (red) cluster samples, overlaid on sample-wide (10-300 Myr) constraints (grayscale). Contours represent 1, 2, and 3σ confidence intervals.

model) predict large numbers of massive clusters that are not present in the PHAT dataset. Even with a steeper low mass slope of α=−2.3 instead of the canonical −2, the power law model provides a visibly poorer fit relative to the best fit Schechter function.

We can quantify this comparison using the number of high-mass clusters with masses $>10^4$ $M_\odot$ observed by PHAT and predicted for this set of models. For each combination of Schechter function parameters, we perform a set of 1000 simulated mass function draws to account for stochastic variations in the number of massive clusters. We plot the number of $>10^4$ $M_\odot$clusters resulting from these draws for each set of mass function parameters as histograms in the right panel of Figure 5.3. Only the model with M_c of 10^4 $M_\odot$ matches the observed value of fifteen $>10^4$ $M_\odot$ clusters, while the models with larger M_c all over-predict the number of massive clusters by factors of ∼2–8. This absence of large numbers of massive clusters cannot be easily explained by any observational bias (as these are the most easily detected clusters), and stand as clear evidence of a mass function truncation.

To test for age-dependence in our mass function fits, we divide the sample into two age bins: 10–100 Myr and 100–300 Myr. We obtain very similar results for the fits to these two age bins with respect to the total 10–300 Myr sample; we compare the 2D posterior PDFs for these three cases in Figure 5.4. We find no significant age dependence on our mass function results. Note that the M_c constraint from younger age bin alone is significantly weaker due to reduced number of clusters where only 324 clusters lie above the 50% mass completeness limit. This demonstrates that our large sample of clusters, obtained by integrating over wide areas and relatively long age ranges, was key to obtaining a robust result.

We note that a clear signature of mass dependent cluster dissolution is a flattening of the low mass slope of the cluster mass function with increasing age (Gieles 2009). The α constraints for the two separate age bins are statistically consistent with a single -2 power law slope at just over 1σ confidence. We conclude that there is no clear signature of cluster dissolution present, in agreement with previous PHAT results (Fouesneau et al. 2014), pointing to longer dissolution timescales (>300–500 Myr). As previously stated, we pursue constraints on timescales and mass dependence of cluster dissolution in future work (M. Fouesneau, in preparation).

Previous studies of the M31 young cluster mass function did not detect a truncation mass of $\sim 10^4$ $M_\odot$. For example, Vansevičius et al. (2009) compare their sample (and the sample from Caldwell et al. 2009) with a Schechter function distribution and argue that their results are consistent with the Larsen (2009) M_c value of 2×10^5 $M_\odot$, but do not fit their own value of M_c.

There are a number of points to consider when comparing our PHAT results to the work of Vansevičius et al. (2009) and Caldwell et al. (2009). First, both of these works examine cluster ages from a broader age bin, including clusters with ages between 1–3 Gyr. We prefer to restrict our analysis to an age regime where we can obtain robust cluster fits from CMD fitting, in addition to the availability of detailed total star formation history constraints for this same time period (Lewis et al. 2015). Second, both Vansevičius et al. (2009) and Caldwell et al. (2009) derive masses using conversions based on fully-sampled mass functions. As mentioned in the introduction, this strategy can lead to significant mass discrepancies due to the stochastic contribution of luminous evolved members. Third, these two studies

were both significantly limited by the mass completeness of their catalogs; Vansevičius et al. (2009) and Caldwell et al. (2009) only reaching 50% completeness masses of $\log(M/M_{\odot})$=3.7 and 4.0, respectively. Without a full accounting of the cluster population at masses below the knee of the distribution, it is difficult to properly constrain the characteristic mass.

Finally, the potential exists that the cluster population surveyed by these previous works, which include clusters that lie on the southwest side of the M31 disk opposite that of the PHAT survey region, might truly represent a different star formation environment with higher intrinsic values of M_c. The southwest portion of M31 hosts the star forming complex NGC206 (Hunter et al. 1996) and vigorous star formation near the split in the 10 kpc star forming ring (Gordon et al. 2006), and is known to host a number of notable massive (10^4–10^5 $M_{\odot}$) clusters (e.g., VdB01; Perina et al. 2009b). Indeed, Elmegreen & Efremov (1997) points out that the southwestern portion of the M31 disk hosts a spiral arm segment (S4; also OB79–82 in parlance of van den Bergh 1964) with particularly high intensity star formation, highlighting this same region of interest. With this in mind, we note that our results apply directly to the PHAT survey region of M31, and that variations across the disk of M31 are possible. Further study of the active southwest portion of the M31 disk could provide an interesting counterpoint to the more moderate star formation survey by PHAT.

5.5 Discussion

5.5.1 Mass Function Truncations for Young Cluster Systems: Correlation with $\Sigma_{\rm SFR}$

In this section we examine whether the physical conditions of star formation in the PHAT survey region can explain the low value of M_c measured here relative to previous studies. We find that combining our result with those from the literature show a clear correlation between the mass truncation, M_c, and the star formation rate surface density, $\Sigma_{\rm SFR}$.

We complement our PHAT M_c result with determinations from the literature. We use the M_c fit for the Antennae calculated by Jordán et al. (2007) using the 2.5–6.3 Myr cluster mass distribution data from Zhang & Fall (1999), yielding $\log(M_c/M_{\odot})=6.3^{+0.7}_{-0.3}$ (but see Whitmore et al. 2014). We also use M51 results from Gieles (2009), and galaxy-wide M83

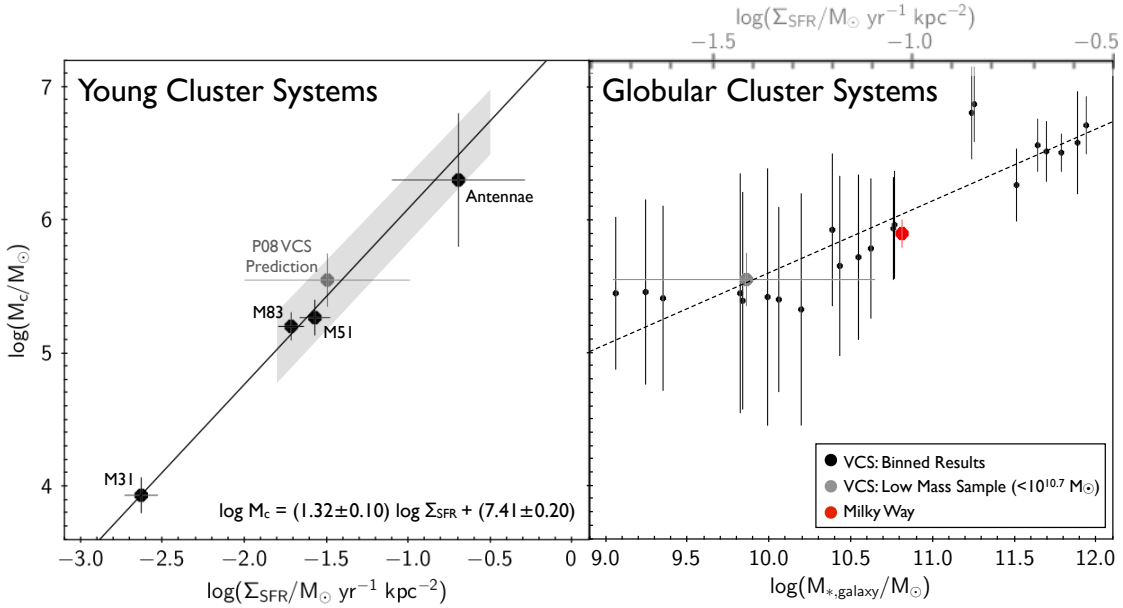

Figure 5.5 Left: Comparison of M_c fits for young cluster samples as a function of $\Sigma_{\rm SFR}$ for M31, M83 (Adamo et al. 2015), M51 (Gieles 2009), and Antennae (Zhang & Fall 1999; Jordán et al. 2007). Right: Comparison of M_c fits for old GC systems as a function total stellar mass of the host galaxy. Results come from the ACS Virgo Cluster Survey (VCS) and the Milky Way (Jordán et al. 2007). A combined result for a low mass sample of ellipticals and early type galaxies ($M_B > -19.5$; $M_* < 10.7\ M_\odot$) is plotted in gray in both panels. The placement in the right panel reflects the average M_c and M_* values derived for the subsample. We also place this binned result in the left panel, where its placement reflects the theoretical prediction for $\Sigma_{\rm SFR}$ in this low mass sample at the epoch of peak cluster formation ($z \sim 4.5$) based on results from the Millennium Simulation (Peng et al. 2008).

results from Adamo et al. (2015); both of these results are consistent with M_c=1–2×10^5 $M_\odot$, which agrees with the result from Larsen (2009) for a combined analysis of ∼20 nearby spiral galaxies (of which M51 and M83 were members).

Next, we assess characteristic, galaxy-averaged $\Sigma_{\rm SFR}$ values for each of these galaxies. We use values from L.C. Johnson et al. (in preparation) and Adamo et al. (2015) for M31 and M83, respectively. We calculate a new $\Sigma_{\rm SFR}$ value for M51 using the radial profile published in Leroy et al. (2008), and obtain a measurement for the Antennae using a galaxy-wide IR-based SFR (Brandl et al. 2009) and estimating our own area normalization from publicly available Spitzer imaging. Due to limitations in the availability of spatially-resolved SFR estimates, our adopted $\Sigma_{\rm SFR}$ value for the Antennae includes a large (0.4 dex) uncertainty representing these shortcomings.

A caveat regarding the $\Sigma_{\rm SFR}$ values of all four galaxies quoted here: these values are characteristic, mass-weighted average values that represent an intrinsic range of star forming environments. The uncertainties we associate with the M31, M51, and M83 $\Sigma_{\rm SFR}$ values are based on the standard deviation of kpc-scale regions within the galaxy and serves as a first-order attempt to capture the intrinsic variation of star forming environments with each of these galaxies. However, one must take care when interpreting how these galaxy-wide values, based on kpc-scale measurements, translate to cloud and complex properties on parsec scales.

We compile M_c and $\Sigma_{\rm SFR}$ measurements for young cluster systems in Table 5.1 and plot these results in the left panel of Figure 5.5. This plot shows a tight correlation between M_c and $\Sigma_{\rm SFR}$, spanning >2 orders of magnitude along each axis. This clear relation suggests a strong dependence of the cluster mass function truncation on the characteristics of the galactic star forming environment.

Although our dataset only includes four galaxy systems, we leverage the large dynamic range in $\Sigma_{\rm SFR}$ and quantify the relationship between M_c and $\Sigma_{\rm SFR}$. We fit the following linear relation in log M_c–log $\Sigma_{\rm SFR}$ parameter space:

$$\log M_c = (1.32 \pm 0.10) \times \log \Sigma_{\rm SFR} + (7.41 \pm 0.20). \tag{5.8}$$

It is interesting to observe that the proportionality measured here ($M_c \propto {\Sigma_{\rm SFR}}^{1.3\pm0.1}$) is sim-

Table 5.1. M_c and $\Sigma_{\rm SFR}$ Data

Galaxy	Region Name	log ($M_c/M_\odot$)	log ($\Sigma_{\rm SFR}/M_\odot$ yr^{-1} kpc^{-2})	References
M31	PHAT	$3.93^{+0.13}_{-0.10}$	-2.63 ± 0.10	This Work
M51	···	5.27 ± 0.13	-1.57 ± 0.10	Gieles 2009
M83	···	5.20 ± 0.10	-1.72 ± 0.08	Adamo et al. 2015
Antennae	···	6.30 ± 0.50	-0.7 ± 0.4	Jordán et al. 2007
Normal Galaxies	···	5.32 ± 0.10	···	Larsen 2009

ilar to the simple expectation for how mid-plane pressure, $P_{\rm mp}$, varies with $\Sigma_{\rm SFR}$ (Elmegreen 2009): $P_{\rm mp} \propto {\Sigma_{\rm gas}}^2$, and $\Sigma_{\rm gas} \propto {\Sigma_{\rm SFR}}^{1.4-1.5}$ from empirical measurements (Kennicutt 1998), resulting in $P_{\rm mp} \propto {\Sigma_{\rm SFR}}^{1.3-1.4}$. This dependence is weaker than the $M_{\rm max} \propto {\Sigma_{\rm SFR}}^2$ discussed in Billett et al. (2002), derived under the assumption of virial equilibrium of a progenitor cloud.

We emphasize that the connection between galactic disk mid-plane pressure and the observed dependence between M_c and $\Sigma_{\rm SFR}$ may be indirect. For instance, the physical conditions that lead to the formation of the most massive star clusters are likely reached through non-equilibrium processes that otherwise scale with the average mid-plane pressure. Cloud-cloud collisions are one example (Fukui et al. 2014), as are shock-based pressure enhancements due to spiral or other structure (Efremov 1985; Elmegreen & Efremov 1997), or large scale compressive shocks in the case of galaxy interactions (e.g., Johnson et al. 2015a). The physical processes that are directly responsible for the formation of the most massive star clusters are likely more complex than a simple dependence on increased average mid-plane pressure.

Finally, we note the possible connection between the variation of M_c and cluster formation efficiency (Γ). Both Γ and M_c have been shown to vary systematically with $\Sigma_{\rm SFR}$, and reducing M_c at a fixed mass function normalization does decrease the integrated total mass in star clusters. While the variation in M_c does tend to decrease Γ with decreasing

$\Sigma_{\rm SFR}$, this effect only explains a small fraction of the overall Γ variation observed. For example, decreasing M_c from 10^6 to 10^4 $M_\odot$ only accounts for a factor of $\sim$2 change in Γ, while galaxy observations and theoretical predictions show evidence for greater than an order of magnitude change over the same range of $\Sigma_{\rm SFR}$ (L.C. Johnson et al., in preparation; Kruijssen 2012).

5.5.2 Mass Function Truncations for Globular Cluster Systems: Similarity to Young Clusters?

Old globular cluster (GC) systems have a dramatically different mass function shape than the young cluster systems discussed in the previous section. Often parameterized using a Gaussian or log-normal form, the globular cluster mass function (GCMF) shows a clear turnover at a near-constant mass of 2×10^5 $M_\odot$ (e.g., Jordán et al. 2007; Villegas et al. 2010). While early theoretical work proposed that these ancient star clusters were formed as part of a peaked distribution through a process that yields clusters at a particular characteristic mass scale (e.g., Peebles & Dicke 1968; Fall & Rees 1985), recent work has pursued an approach where an initial power law mass distribution is evolved due to mass-dependent dynamical evolution and destruction (e.g., Gnedin & Ostriker 1997; Fall & Zhang 2001).

Analysis of GC systems in early-type Virgo and Fornax galaxy cluster members presented by Jordán et al. (2006, 2007) and Villegas et al. (2010) show clear evidence for variation in the width of the peaked GC luminosity function, and therefore the GCMF. Jordán et al. (2007) demonstrate that this increase in width of the mass distribution can be interpreted equivalently as an increase in dispersion ($\sigma_{\rm LN}$) for a traditional lognormal functional form, or as an increase in M_c for an evolved Schechter function, using a parameterization inspired by Fall & Zhang (2001).

We find the variation in M_c observed for GC systems particularly intriguing in light of the M_c–$\Sigma_{\rm SFR}$ correlation we derived for young cluster populations. Let us explore the hypothesis that GC formation in the early Universe followed the same correlation with star formation intensity as found for young clusters in the present day. Under this assumption, the M_c values derived for GC systems can be used to infer the characteristic $\Sigma_{\rm SFR}$ during

their epoch of formation.

We note that Jordán et al. (2006, 2007), Villegas et al. (2010), and Harris et al. (2014) report that derived M_c values (or equivalently, $\sigma_{\rm LN}$ values) vary systematically with host galaxy luminosity or mass. We show the relation between M_c and galaxy mass in the right panel of Figure 5.5 using results from the ACS Virgo Cluster Survey (VCS; Côté et al. 2004); we use binned M_c results for Virgo GC systems based on z-band luminosity function fitting from Jordán et al. (2007), and derive representative galaxy masses ($M_{*,\rm galaxy}$) using z-band luminosities from Ferrarese et al. (2006), $g-z$ colors and distances from Blakeslee et al. (2009), and a color-based stellar mass-to-light ratio from Into & Portinari (2013). We fit the following linear trend for $\log M_c$ as a function of $\log M_{*,\rm galaxy}$:

$$\log M_c = (0.50 \pm 0.11) M_{*,\rm galaxy} + (0.7 \pm 1.2). \tag{5.9}$$

We also include constraints for the Milky Way GC system to demonstrate that the GCMF behavior shown here is not a special feature of early type host galaxies in galaxy cluster environments. We use an M_c determination reported in Jordán et al. (2007) and a galaxy stellar mass from McMillan (2011); this data point is not included in the M_c–$M_{*,\rm galaxy}$ fit.

Following the hypothesis that the M_c variation observed in GC systems results from the correlation with star formation intensity that we observe for young clusters in Section 5.5.1, we combine relations for M_c in terms of $\Sigma_{\rm SFR}$ and $M_{*,\rm galaxy}$. As a result, we make predictions for the $\Sigma_{\rm SFR}$ at the epoch of GC formation as a function of present day galaxy mass for this sample of early-type Virgo galaxies. We add the predicted $\Sigma_{\rm SFR}$ scaling as a secondary axis in the right panel of Figure 5.5, and highlight the range of M_c–$\Sigma_{\rm SFR}$ parameter space relevant to these GC systems as a gray shaded region in the left panel. We predict that GCs formed in star forming environments with $\Sigma_{\rm SFR}$ values that span a range from those characteristic of active star forming galaxies like M83 and M51 at the low intensity end, up to extreme environments similar to and more intense than found in the Antennae.

We find theoretical support for the hypothesis of a universal relation between cluster mass function truncations and star formation intensity in a study of GC formation by Peng et al. (2008). They find that relatively low mass (10^9 $M_\odot$), early type galaxies in the Virgo galaxy cluster have $\log$ ($\Sigma_{\rm SFR}/M_\odot$ yr^{-1} kpc^{-2}) ~ -1.5 at the $z \sim 4.5$ peak

GC formation epoch, according to their analysis of Millennium simulation results (Springel et al. 2005). We observe that a comparable low-mass sample of VCS galaxies is observed to have log $(M_c/M_\odot) \sim 5.5$, and plot this averaged result in the right panel of Figure 5.5. We find excellent agreement between the $\Sigma_{\rm SFR}$ prediction from Peng et al. (2008) and the star formation rate intensity derived for the observed M_c of the low mass sample using our young cluster relation in Equation 5.8. This agreement represents theoretical support for the hypothesis that present day young clusters and ancient GCs follow the same M_c scaling relation as a function of star formation environment.

The notion of a universal dependence of massive cluster formation on the star formation intensity of its natal environment is very exciting. If our hypothesis is true, measurements of the upper end of the GCMF would allow observers to infer important details about the hierarchical build-up of galaxies and the physical conditions of star formation in the early universe. There are many aspects of the GC populations we have yet to fully understand (e.g., their specific frequencies, metallicity distributions, destruction and mass loss mechanisms, the origin of He and light element abundance variations among cluster members), but we are hopeful that M_c measurements can provide important insights that will better inform our overall understanding of star formation in the early Universe.

5.6 Summary

We find statistically significant evidence for a high-mass truncation of the star cluster mass function within the PHAT survey region in M31. Parameterized using a Schechter function, this exponential truncation has a characteristic mass of $M_c = 8.5^{+2.8}_{-1.8} \times 10^3\ M_\odot$. This M_c is the lowest value ever observed for a star cluster population, and provides definitive evidence for a physical upper mass limit for cluster formation within M31 that is not due to stochastic sampling of a pure power law mass distribution.

When we combine our M31 results with previous M_c results derived for young cluster systems from the literature, we identify for the first time a strong systematic correlation between the star cluster mass function truncation and star formation environment, as characterized by a host galaxy's $\Sigma_{\rm SFR}$. As $\Sigma_{\rm SFR}$ decreases, M_c also decreases according to a super-linear dependence on star formation rate surface density: $M_c \propto {\Sigma_{\rm SFR}}^{\sim 1.3}$.

Finally, we highlight that GC systems also show systematic variations in the high-mass truncations of their mass functions. We hypothesize that these mass function variations are the result of the same environmentally-dependent truncation relation that we observe for young cluster systems in the present day within nearby galaxies. If true, this enables the use of GC systems in present day galaxies to infer the star formation conditions during their formation epoch. While further study is required, we show that there is preliminary theoretical support from simulations of galaxy formation in support of this idea. This proposed commonality shared between ancient GCs and young clusters forming today could represent a long-sought link demonstrating that while star formation in the early Universe was generally more active and intense, cluster formation at any epoch follows the same universal trends.

Chapter 6

CONCLUSION

6.1 Summary of Contributions

My dissertation research significantly contributes to the study of star clusters, stellar populations, and star formation. Here, I summarize my work and outline my contributions to the field of astrophysics.

In Chapter 3 (Johnson et al. 2015b), I described the effort to construct a high quality sample of 2753 star clusters using imaging from the Panchromatic Hubble Andromeda Treasury survey (PHAT; Dalcanton et al. 2012). The catalog was assembled using 1.8 million image classifications contributed by thousands of volunteers as part of the Andromeda Project citizen science project. This effort built upon an initial, traditional visual search conducted by a group of professional astronomers presented in Chapter 2 (Johnson et al. 2012), and provided the efficiency necessary to systematically search $\sim$0.5 deg^2 (totaling $\sim$7 gigapixels) of imaging data.

The PHAT cluster catalog is a well-characterized census of Andromeda's star cluster population, providing one of the best galaxy-wide samples to date. This catalog increases the number of known clusters in the PHAT survey footprint by a factor of 6 with respect to previous catalogs in the literature, which shows the value of high spatial resolution imaging from PHAT. The PHAT cluster sample extends to low mass clusters ($\lesssim 10^3$ $M_\odot$) at young ages ($\lesssim$300 Myr) that go undetected in more distant galaxies (>1 Mpc). We conducted a thorough determination of catalog completeness, computing how detection varies as a function cluster properties (e.g., age, mass, and size) and galaxy properties (e.g., background stellar density and spatial distribution). Furthermore, each star cluster is resolved into individual stellar members, allowing detailed analysis of individual clusters through color-magnitude diagram (CMD) fitting, providing a richness of data that is only possible within the Local Group.

These catalogs provide a basic characterization of the M31 star cluster population, including sample-wide six-band integrated photometry, luminosity functions, and photometrically-determined sizes. These basic properties serve as input for additional characterization work, including age and mass determinations (Fouesneau et al. 2014, L. Beerman et al., in preparation).

A number of valuable investigations build out of the initial cluster discovery and characterization analysis. The first of these was an exploration of the cluster formation efficiency in Chapter 4: the fraction of young stars formed as members of long-lived star clusters (Γ). We find that 3–6% of young stars (10–100 Myr old) are born as cluster members, where this fraction varies systematically across the M31 disk according to star formation rate intensity (measured using star formation rate surface density, $\Sigma_{\rm SFR}$) and gas depletion time ($\tau_{\rm dep}$). These M31 measurements agree with the established trend of rising cluster formation efficiency with increasing star formation rate intensity, in which Γ increases from $\sim$1% in regions with low $\Sigma_{\rm SFR}$ to $\sim$50% in intense starburst environments. While this result adds to a growing collection of Γ studies in the literature (e.g., Goddard et al. 2010; Adamo et al. 2011; Cook et al. 2012; Adamo et al. 2015), the spatially resolved nature of the M31 analysis improves upon galaxy-integrated determinations, providing important leverage for physical interpretation of environmentally dependent cluster formation behavior.

We also compare our Γ measurements to theoretical predictions from Kruijssen (2012). We build on this initial work by deriving a new set of Γ relations as a function of $\Sigma_{\rm SFR}$ and gas depletion time ($\tau_{\rm dep}$), derived under a new set of assumptions concerning the underlying star formation law ($\Sigma_{\rm SFR}$ as a function of $\Sigma_{\rm gas}$) that are well-suited for use with spatially-resolved observations. Not only do these new relations help to understand Γ measurements in M31, but also help in proposing new tests of feedback prescriptions in starburst environments. Overall, we demonstrate good agreement between theoretical predictions and observations, providing further support for the underlying model of cluster formation in which clusters emerge from regions with high local star formation efficiency that occur in high density regions of a hierarchically-structured interstellar medium.

In addition to examining the dependence of cluster formation efficiency on galactic environment, we use the M31 cluster sample to investigate the star cluster mass function in

Chapter 5. We find that the mass function constructed for 840 young (10–300 Myr old) star clusters in the PHAT sample is truncated at its high mass end, and can be described using a Schechter function with a characteristic mass of $M_c = 8.5^{+2.8}_{-1.8} \times 10^3\ M_\odot$. When combined with other reported cluster mass function truncations, we show for the first time that the location of the exponential cutoff depends on star formation rate surface density, where $M_c \propto {\Sigma_{\rm SFR}}^{\sim 1.3}$. Additionally, we explore whether observed variations in the high-mass end of the globular cluster mass function could be connected to the M_c–$\Sigma_{\rm SFR}$ relation dependence we derive for young cluster systems, and present evidence supporting the hypothesis that cluster formation processes behave the same in the early universe as they do in the present day.

6.2 Related Studies

Beyond investigations of cluster formation efficiency and the cluster mass function in M31, the PHAT star cluster sample serves as the basis for other investigations of Andromeda's stellar populations. First, PHAT observations of M31 star clusters facilitate a systematic study of the high-mass stellar initial mass function (IMF). In their recent review of the literature, Bastian et al. (2010) concludes that that there is no clear evidence for significant variations in the IMF, however the current set of observational IMF constraints still allow for the possibility of systematic variations. A majority of current IMF studies are performed in single star clusters, one cluster at a time, leading to a large amount of heterogeneity in authors, analysis techniques, and data quality among the current set of published results. Concerns about systematic errors and underestimated uncertainties (see Weisz et al. 2013) severely limit the community's ability to interpret underlying IMF behavior.

Using observations of 85 young (>25 Myr) PHAT clusters, Weisz et al. (2015) systematically analyze the stellar IMF using a uniform data set, homogeneous data processing and preparation methods, and probabilistic fitting techniques. This consistent, robust analysis of a high quality data set provides the largest systematic study of the high-mass stellar IMF to date. Finding no evidence of systematic variations, this study concludes that the high-mass IMF appears universal and can be characterized by a power law slope of $-1.45^{+0.03}_{-0.06}$, which is slightly steeper than the canonical Kroupa (-1.30) or Salpeter (-1.35) power law

indices.

In a different application of the PHAT cluster catalog, Senchyna et al. (2015) combines age constraints for PHAT clusters with Cepheid catalogs published by Kodric et al. (2013) to study the period-age relation for Cepheid variable stars. The availability of a large sample of clusters is essential for a study of intrinsically rare objects. This study identifies 10 candidate Cepheid cluster members, significantly adding to 23 known Galactic candidates (Anderson et al. 2013), and contributes additional age constraints in support of future observational comparisons to theoretical pulsation and stellar evolution models.

These studies demonstrate that star clusters are valuable, multi-faceted astrophysical laboratories. The construction of a large, well-characterized sample of clusters from the PHAT survey allows countless new and exciting avenues of research. Examples of future applications include studies that use age-tagged cluster members to improve late stage stellar evolution models of asymptotic giant branch stars (L. Girardi et al., in preparation) and red/blue supergiants (L.C. Johnson et al., in preparation). However, the fact that all PHAT data is publicly available[1] permits, and encourages, a multitude of future applications.

6.3 Future Directions

The science results from my dissertation lead to an array of future directions. Additional high-quality, spatially-resolved measurements of cluster formation efficiency would allow us to explore new, higher intensity star forming environments and improve our understanding of the environmental dependencies of cluster formation. The Magellanic Clouds and M33 represent auspicious targets for extending the range of star formation rate intensity upward by nearly an order of magnitude, while still providing the ability to perform detailed analyses using individually resolved cluster member stars. Future HST programs and wide-area ground-based surveys (e.g., the Survey of the Magellanic Stellar History, SMASH; D. Nidever, in preparation) provide amazing opportunities for future observational studies of cluster formation.

The topic of cluster formation is also ripe for theoretical advancement. As discussed

[1] `https://archive.stsci.edu/prepds/phat/`

in Chapter 4, constructing a comprehensive framework for cluster formation that not only predicts the fraction of mass formed in long-lived star clusters, but also describes the mass function of the systems that are produced is now within reach (Krumholz 2014). These theoretical models are essential for understanding and interpreting observational differences revealed by the next generation of cluster and star formation studies.

Looking beyond the realm of studies that focus solely on star formation products (young cluster and field star populations), there is great potential in research that studies the link between properties of giant molecular clouds and the star forming ISM to the stellar populations they produce. While short timescales and the embedded nature of young stellar populations tend to complicate studies of early phases of star formation, vast improvements of southern hemisphere observatories (in the form of the DECam optical imager on the 4-m Blanco telescope at CTIO, and the ALMA sub-millimeter observatory) will enable detailed, revolutionary studies of star formation in the Magellanic Clouds.

Finally, the hypothesis that globular cluster mass function truncations could yield insights about the star formation environments of galaxy hosts in the early universe is extremely exciting. There are many avenues for future study in this line of research. For instance, can we find corroborating evidence that supports the notion of a universal correlation between star cluster mass functions and $\Sigma_{\rm SFR}$? How does the globular cluster mass function truncation differ for metal-rich and metal-poor subsamples of globular clusters? What do these differences (or similarities) tell us about the origin and relative differences among globular clusters subpopulations? How do the characteristic star formation rate intensities predicted by the young cluster relation with $\Sigma_{\rm SFR}$ compare to values inferred from the latest theoretical galaxy formation models? This open-ended topic suggest many avenues of future theoretical and observation study, and while still rather speculative, perhaps holds the greatest potential for wide-scale impact across a broad range of astronomical research.

6.4 Closing

My dissertation makes a significant contribution to the understanding of Andromeda's star cluster population and cluster formation on the whole. As is the case for most scientific

investigations, the work I have begun here is by no-means complete. For every question I was able to answer throughout the course of my research, a new intriguing puzzle rose to take its place. However, I take pride in participating in the scientific endeavor, and I hope that I never run out of cosmic questions to answer.

BIBLIOGRAPHY

Adamo, A., Kruijssen, D., Bastian, N., Silva-Villa, E., & Ryon, J. 2015, ArXiv e-prints

Adamo, A., Östlin, G., & Zackrisson, E. 2011, MNRAS, 417, 1904

Allison, R. J., Goodwin, S. P., Parker, R. J., Portegies Zwart, S. F., & de Grijs, R. 2010, MNRAS, 407, 1098

Anderson, R. I., Eyer, L., & Mowlavi, N. 2013, MNRAS, 434, 2238

Annibali, F., Tosi, M., Aloisi, A., & van der Marel, R. P. 2011, AJ, 142, 129

Barmby, P., & Huchra, J. P. 2001, AJ, 122, 2458

Barmby, P., Huchra, J. P., & Brodie, J. P. 2001, AJ, 121, 1482

Barmby, P., Huchra, J. P., Brodie, J. P., Forbes, D. A., Schroder, L. L., & Grillmair, C. J. 2000, AJ, 119, 727

Barmby, P., McLaughlin, D. E., Harris, W. E., Harris, G. L. H., & Forbes, D. A. 2007, AJ, 133, 2764

Barmby, P., et al. 2006, ApJ, 650, L45

Bastian, N. 2008, MNRAS, 390, 759

Bastian, N. 2013, in Astronomical Society of the Pacific Conference Series, Vol. 470, 370 Years of Astronomy in Utrecht, ed. G. Pugliese, A. de Koter, & M. Wijburg, 287

Bastian, N., Covey, K. R., & Meyer, M. R. 2010, ARA&A, 48, 339

Bastian, N., Konstantopoulos, I. S., Trancho, G., Weisz, D. R., Larsen, S. S., Fouesneau, M., Kaschinski, C. B., & Gieles, M. 2012a, A&A, 541, A25

Bastian, N., et al. 2011, MNRAS, 417, L6

—. 2012b, MNRAS, 419, 2606

Battistini, P., Bonoli, F., Braccesi, A., Federici, L., Fusi Pecci, F., Marano, B., & Borngen, F. 1987, A&AS, 67, 447

Battistini, P. L., Bonoli, F., Casavecchia, M., Ciotti, L., Federici, L., & Fusi-Pecci, F. 1993, A&A, 272, 77

Baumgardt, H., & Kroupa, P. 2007, MNRAS, 380, 1589

Baumgardt, H., Parmentier, G., Anders, P., & Grebel, E. K. 2013, MNRAS, 430, 676

Beerman, L. C., et al. 2012, ApJ, 760, 104

Bica, E., Bonatto, C., Dutra, C. M., & Santos, J. F. C. 2008, MNRAS, 389, 678

Bica, E., & Dutra, C. M. 2000, AJ, 119, 1214

Bica, E. L. D., Schmitt, H. R., Dutra, C. M., & Oliveira, H. L. 1999, AJ, 117, 238

Bigiel, F., Leroy, A., Walter, F., Blitz, L., Brinks, E., de Blok, W. J. G., & Madore, B. 2010, AJ, 140, 1194

Bigiel, F., Leroy, A., Walter, F., Brinks, E., de Blok, W. J. G., Madore, B., & Thornley, M. D. 2008, AJ, 136, 2846

Billett, O. H., Hunter, D. A., & Elmegreen, B. G. 2002, AJ, 123, 1454

Blakeslee, J. P., et al. 2009, ApJ, 694, 556

Block, D. L., et al. 2006, Nature, 443, 832

Boissier, S., et al. 2005, ApJ, 619, L83

Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207

Bonnell, I. A., Clark, P., & Bate, M. R. 2008, MNRAS, 389, 1556

Borissova, J., et al. 2011, A&A, 532, A131

Boutloukos, S. G., & Lamers, H. J. G. L. M. 2003, MNRAS, 338, 717

Brandl, B. R., et al. 2009, ApJ, 699, 1982

Braun, R., Thilker, D. A., Walterbos, R. A. M., & Corbelli, E. 2009, ApJ, 695, 937

Bressert, E., et al. 2010, MNRAS, 409, L54

Caldwell, N., Harding, P., Morrison, H., Rose, J. A., Schiavon, R., & Kriessler, J. 2009, AJ, 137, 94

Caldwell, N., Schiavon, R., Morrison, H., Rose, J. A., & Harding, P. 2011, AJ, 141, 61

Chandar, R., Fall, S. M., & Whitmore, B. C. 2010a, ApJ, 711, 1263

Chandar, R., Whitmore, B. C., & Fall, S. M. 2010b, ApJ, 713, 1343

Chandar, R., et al. 2010c, ApJ, 719, 966

Chiosi, C., Bertelli, G., Meylan, G., & Ortolani, S. 1989, A&A, 219, 167

Collins, M. L. M., et al. 2011, MNRAS, 413, 1548

Colucci, J. E., Bernstein, R. A., Cameron, S., McWilliam, A., & Cohen, J. G. 2009, ApJ, 704, 385

Cook, D. O., et al. 2012, ApJ, 751, 100

Corbelli, E., Lorenzoni, S., Walterbos, R., Braun, R., & Thilker, D. 2010, A&A, 511, A89

Côté, P., et al. 2004, ApJS, 153, 223

Crampton, D., Cowley, A. P., Schade, D., & Chayer, P. 1985, ApJ, 288, 494

Daddi, E., et al. 2010, ApJ, 714, L118

D'Agostino, R. B., & Stephens, M. A. 1986, Goodness-of-fit techniques, ed. D'Agostino, R. B. & Stephens, M. A. (New York: Dekker)

Dalcanton, J. J., et al. 2012, ApJS, 200, 18

Davies, B., Bastian, N., Gieles, M., Seth, A. C., Mengel, S., & Konstantopoulos, I. S. 2011, MNRAS, 411, 1386

Davies, B., de La Fuente, D., Najarro, F., Hinton, J. A., Trombley, C., Figer, D. F., & Puga, E. 2012, MNRAS, 419, 1860

Dias, W. S., Alessi, B. S., Moitinho, A., & Lépine, J. R. D. 2002, A&A, 389, 871

Dierickx, M., Blecha, L., & Loeb, A. 2014, ApJ, 788, L38

Dolphin, A. E. 2000, PASP, 112, 1383

—. 2002, MNRAS, 332, 91

—. 2012, ApJ, 751, 60

—. 2013, ApJ, 775, 76

Dorman, C. E., et al. 2015, ApJ, 803, 24

Dutra, C. M., Bica, E., Soares, J., & Barbuy, B. 2003, A&A, 400, 533

Efremov, Y. N. 1985, Soviet Astronomy Letters, 11, 69

Elmegreen, B. G. 2002, ApJ, 577, 206

—. 2008, ApJ, 672, 1006

—. 2009, Ap&SS, 324, 83

Elmegreen, B. G., & Efremov, Y. N. 1997, ApJ, 480, 235

Fall, S. M., & Chandar, R. 2012, ApJ, 752, 96

Fall, S. M., Chandar, R., & Whitmore, B. C. 2009, ApJ, 704, 453

Fall, S. M., & Rees, M. J. 1985, ApJ, 298, 18

Fall, S. M., & Zhang, Q. 2001, ApJ, 561, 751

Fan, Z., de Grijs, R., & Zhou, X. 2010, ApJ, 725, 200

Fardal, M. A., Babul, A., Guhathakurta, P., Gilbert, K. M., & Dodge, C. 2008, ApJ, 682, L33

Ferrarese, L., et al. 2006, ApJS, 164, 334

Figer, D. F., MacKenty, J. W., Robberto, M., Smith, K., Najarro, F., Kudritzki, R. P., & Herrero, A. 2006, ApJ, 643, 1166

Ford, G. P., et al. 2013, ApJ, 769, 55

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306

Fouesneau, M., & Lançon, A. 2010, A&A, 521, A22+

Fouesneau, M., Lançon, A., Chandar, R., & Whitmore, B. C. 2012, ApJ, 750, 60

Fouesneau, M., et al. 2014, ApJ, 786, 117

Friel, E. D. 1995, ARA&A, 33, 381

Frogel, J. A., Mould, J., & Blanco, V. M. 1990, ApJ, 352, 96

Fukui, Y., et al. 2014, ApJ, 780, 36

Fusi Pecci, F., Bellazzini, M., Buzzoni, A., De Simone, E., Federici, L., & Galleti, S. 2005, AJ, 130, 554

Gallazzi, A., Brinchmann, J., Charlot, S., & White, S. D. M. 2008, MNRAS, 383, 1439

Galleti, S., Federici, L., Bellazzini, M., Fusi Pecci, F., & Macrina, S. 2004, A&A, 416, 917

Gieles, M. 2009, MNRAS, 394, 2113

Gieles, M. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 423, Galaxy Wars: Stellar Populations and Star Formation in Interacting Galaxies, ed. B. Smith, J. Higdon, S. Higdon, & N. Bastian, 123

Gieles, M., & Bastian, N. 2008, A&A, 482, 165

Gieles, M., Larsen, S. S., Bastian, N., & Stein, I. T. 2006a, A&A, 450, 129

Gieles, M., Moeckel, N., & Clarke, C. J. 2012, MNRAS, 426, L11

Gieles, M., & Portegies Zwart, S. F. 2011, MNRAS, 410, L6

Gieles, M., Portegies Zwart, S. F., Baumgardt, H., Athanassoula, E., Lamers, H. J. G. L. M., Sipior, M., & Leenaarts, J. 2006b, MNRAS, 371, 793

Girardi, L., Rubele, S., & Kerber, L. 2009, MNRAS, 394, L74

Girardi, L., et al. 2008, PASP, 120, 583

—. 2010, ApJ, 724, 1030

Gnedin, O. Y., & Ostriker, J. P. 1997, ApJ, 474, 223

Goddard, Q. E., Bastian, N., & Kennicutt, R. C. 2010, MNRAS, 405, 857

Goodman, J., & Weare, J. 2010, Commun. Appl. Math. Comput. Sci., 5, 65

Goodwin, S. P., & Bastian, N. 2006, MNRAS, 373, 752

Gordon, K. D., et al. 2006, ApJ, 638, L87

Haas, M. R., Gieles, M., Scheepmaker, R. A., Larsen, S. S., & Lamers, H. J. G. L. M. 2008, A&A, 487, 937

Harris, J., & Zaritsky, D. 2009, AJ, 138, 1243

Harris, W. E. 1991, ARA&A, 29, 543

Harris, W. E., et al. 2014, ApJ, 797, 128

Hills, J. G. 1980, ApJ, 235, 986

Hodge, P., Krienke, O. K., Bianchi, L., Massey, P., & Olsen, K. 2010, PASP, 122, 745

Hodge, P. W. 1979, AJ, 84, 744

Hodge, P. W., Krienke, O. K., Bellazzini, M., Perina, S., Barmby, P., Cohen, J. G., Puzia, T. H., & Strader, J. 2009, AJ, 138, 770

Hodge, P. W., Mateo, M., Lee, M. G., & Geisler, D. 1987, PASP, 99, 173

Hopkins, P. F. 2013, MNRAS, 428, 1950

Hubble, E. 1932, ApJ, 76, 44

Hunter, D. A., Baum, W. A., O'Neil, Jr., E. J., & Lynds, R. 1996, ApJ, 468, 633

Hunter, D. A., Elmegreen, B. G., Dupuy, T. J., & Mortonson, M. 2003, AJ, 126, 1836

Huxor, A. P., Tanvir, N. R., Ferguson, A. M. N., Irwin, M. J., Ibata, R., Bridges, T., & Lewis, G. F. 2008, MNRAS, 385, 1989

Huxor, A. P., Tanvir, N. R., Irwin, M. J., Ibata, R., Collett, J. L., Ferguson, A. M. N., Bridges, T., & Lewis, G. F. 2005, MNRAS, 360, 1007

Huxor, A. P., et al. 2014, MNRAS, 442, 2165

Into, T., & Portinari, L. 2013, MNRAS, 430, 2715

Johnson, K. E., Leroy, A. K., Indebetouw, R., Brogan, C. L., Whitmore, B. C., Hibbard, J., Sheth, K., & Evans, A. S. 2015a, ApJ, 806, 35

Johnson, L. C., et al. 2012, ApJ, 752, 95

—. 2015b, ApJ, 802, 127

Jordán, A., et al. 2006, ApJ, 651, L25

—. 2007, ApJS, 171, 101

Kawamura, A., et al. 2009, ApJS, 184, 1

Kennicutt, Jr., R. C. 1998, ApJ, 498, 541

Kharchenko, N. V., Piskunov, A. E., Schilbach, E., Röser, S., & Scholz, R.-D. 2013, A&A, 558, A53

Kim, S. C., et al. 2007, AJ, 134, 706

King, I. 1962, AJ, 67, 471

Kodric, M., et al. 2013, AJ, 145, 106

Koekemoer, A. M., Fruchter, A. S., Hook, R. N., & Hack, W. 2002, in The 2002 HST Calibration Workshop : Hubble after the Installation of the ACS and the NICMOS Cooling System, ed. S. Arribas, A. Koekemoer, & B. Whitmore, 337

Krienke, O. K., & Hodge, P. W. 2007, PASP, 119, 7

—. 2008, PASP, 120, 1

—. 2013, PASP, 125, 636

Kroupa, P. 2001, MNRAS, 322, 231

Kruijssen, J. M. D. 2012, MNRAS, 426, 3008

Krumholz, M. R. 2014, ArXiv e-prints

Krumholz, M. R., Dekel, A., & McKee, C. F. 2012, ApJ, 745, 69

Krumholz, M. R., Fumagalli, M., da Silva, R. L., Rendahl, T., & Parra, J. 2015, ArXiv e-prints

Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250

Lada, C. J., & Lada, E. A. 2003, ARA&A, 41, 57

Lamers, H. J. G. L. M., Gieles, M., Bastian, N., Baumgardt, H., Kharchenko, N. V., & Portegies Zwart, S. 2005, A&A, 441, 117

Larsen, S. S. 2002, AJ, 124, 1393

—. 2009, A&A, 494, 539

Larsen, S. S., & Richtler, T. 2000, A&A, 354, 836

Leroy, A. K., Walter, F., Brinks, E., Bigiel, F., de Blok, W. J. G., Madore, B., & Thornley, M. D. 2008, AJ, 136, 2782

Leroy, A. K., et al. 2013a, ApJ, 769, L12

—. 2013b, AJ, 146, 19

Lewis, A. R., et al. 2015, ApJ, 805, 183

Lim, S., & Lee, M. G. 2015, ApJ, 804, 123

Lintott, C. J., et al. 2008, MNRAS, 389, 1179

Lundgren, A. A., Wiklind, T., Olofsson, H., & Rydbeck, G. 2004, A&A, 413, 505

Mackey, A. D., Broby Nielsen, P., Ferguson, A. M. N., & Richardson, J. C. 2008, ApJ, 681, L17

Mackey, A. D., et al. 2006, ApJ, 653, L105

—. 2010, ApJ, 717, L11

—. 2014, MNRAS, 445, L89

Maíz Apellániz, J., & Úbeda, L. 2005, ApJ, 629, 873

Marigo, P., Girardi, L., Bressan, A., Groenewegen, M. A. T., Silva, L., & Granato, G. L. 2008, A&A, 482, 883

Maschberger, T., & Kroupa, P. 2011, MNRAS, 411, 1495

Massey, P., Olsen, K. A. G., Hodge, P. W., Strong, S. B., Jacoby, G. H., Schlingman, W., & Smith, R. C. 2006, AJ, 131, 2478

Massey, P., Silva, D. R., Levesque, E. M., Plez, B., Olsen, K. A. G., Clayton, G. C., Meynet, G., & Maeder, A. 2009, ApJ, 703, 420

McConnachie, A. W., Irwin, M. J., Ferguson, A. M. N., Ibata, R. A., Lewis, G. F., & Tanvir, N. 2005, MNRAS, 356, 979

McConnachie, A. W., et al. 2009, Nature, 461, 66

McLaughlin, D. E., & van der Marel, R. P. 2005, ApJS, 161, 304

McMillan, P. J. 2011, MNRAS, 414, 2446

Meidt, S. E., et al. 2015, ArXiv e-prints

Mercer, E. P., et al. 2005, ApJ, 635, 560

Moeckel, N., Holland, C., Clarke, C. J., & Bonnell, I. A. 2012, MNRAS, 425, 450

Mora, M. D., Larsen, S. S., & Kissler-Patig, M. 2007, A&A, 464, 495

Nieten, C., Neininger, N., Guélin, M., Ungerechts, H., Lucas, R., Berkhuijsen, E. M., Beck, R., & Wielebinski, R. 2006, A&A, 453, 459

Peacock, M. B., Maccarone, T. J., Knigge, C., Kundu, A., Waters, C. Z., Zepf, S. E., & Zurek, D. R. 2010, MNRAS, 402, 803

Peebles, P. J. E., & Dicke, R. H. 1968, ApJ, 154, 891

Peng, E. W., et al. 2008, ApJ, 681, 197

Perina, S., Federici, L., Bellazzini, M., Cacciari, C., Fusi Pecci, F., & Galleti, S. 2009a, A&A, 507, 1375

Perina, S., Galleti, S., Fusi Pecci, F., Bellazzini, M., Federici, L., & Buzzoni, A. 2011, A&A, 531, A155

Perina, S., et al. 2009b, A&A, 494, 933

—. 2010, A&A, 511, A23+

Piskunov, A. E., Kharchenko, N. V., Schilbach, E., Röser, S., Scholz, R.-D., & Zinnecker, H. 2008, A&A, 487, 557

Popescu, B., Hanson, M. M., & Elmegreen, B. G. 2012, ApJ, 751, 122

Portegies Zwart, S. F., McMillan, S. L. W., & Gieles, M. 2010, ARA&A, 48, 431

Rich, R. M., Corsi, C. E., Cacciari, C., Federici, L., Fusi Pecci, F., Djorgovski, S. G., & Freedman, W. L. 2005, AJ, 129, 2670

Ryon, J. E., et al. 2014, AJ, 148, 33

San Roman, I., Sarajedini, A., & Aparicio, A. 2010, ApJ, 720, 1674

Sanders, N. E., Caldwell, N., McDowell, J., & Harding, P. 2012, ApJ, 758, 133

Sandstrom, K. M., et al. 2013, ApJ, 777, 5

Sargent, W. L. W., Kowal, C. T., Hartwick, F. D. A., & van den Bergh, S. 1977, AJ, 82, 947

Schechter, P. 1976, ApJ, 203, 297

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Schwamb, M. E., et al. 2012, ApJ, 754, 129

Senchyna, P., et al. 2015, ApJ, submitted

Silva-Villa, E., Adamo, A., & Bastian, N. 2013, MNRAS, 436, L69

Silva-Villa, E., & Larsen, S. S. 2010, A&A, 516, A10+

—. 2011, A&A, 529, A25+

Simpson, R. J., et al. 2012, MNRAS, 424, 2442

Skrutskie, M. F., et al. 2006, AJ, 131, 1163

Springel, V., et al. 2005, Nature, 435, 629

Strader, J., Caldwell, N., & Seth, A. C. 2011, AJ, 142, 8

Tamm, A., Tempel, E., Tenjes, P., Tihhonova, O., & Tuvikene, T. 2012, A&A, 546, A4

Thilker, D. A., et al. 2005, ApJ, 619, L67

van den Bergh, S. 1964, ApJS, 9, 65

van Dokkum, P. G. 2001, PASP, 113, 1420

Vansevičius, V., Kodaira, K., Narbutis, D., Stonkutė, R., Bridžius, A., Deveikis, V., & Semionov, D. 2009, ApJ, 703, 1872

Verley, S., Corbelli, E., Giovanardi, C., & Hunt, L. K. 2009, A&A, 493, 453

Vetešnik, M. 1962, Bulletin of the Astronomical Institutes of Czechoslovakia, 13, 180

Villegas, D., et al. 2010, ApJ, 717, 603

Weisz, D. R., et al. 2013, ApJ, 762, 123

—. 2015, ApJ, 806, 198

Whitmore, B. C., Chandar, R., Bowers, A. S., Larsen, S., Lindsay, K., Ansari, A., & Evans, J. 2014, AJ, 147, 78

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

Whitmore, B. C., et al. 2010, AJ, 140, 75

Whitney, B. A., et al. 2008, AJ, 136, 18

Willett, K. W., et al. 2013, MNRAS, 435, 2835

Williams, B. F., & Hodge, P. W. 2001a, ApJ, 559, 851

—. 2001b, ApJ, 548, 190

Williams, B. F., et al. 2014, ApJS, 215, 9

Zhang, Q., & Fall, S. M. 1999, ApJ, 527, L81

Zurita, A., & Bresolin, F. 2012, MNRAS, 427, 1463

VITA

L. Clifton (Cliff) Johnson was born and raised in Hannibal, Missouri. He attended Colby College in Waterville, Maine where he received a Bachelor of Arts degree in Physics and Mathematical Sciences in 2007. After working as a post-baccalaureate researcher and observer at the University of Wyoming in Laramie, Cliff moved to Seattle in 2008 to begin his graduate studies in the Astronomy department at the University of Washington. He completed his Masters of Science degree in 2009, and completed his doctorate in 2015.