

DISENTANGLING LUMINOSITY, MORPHOLOGY, STAR FORMATION,  
STELLAR MASS, AND ENVIRONMENT IN GALAXY EVOLUTION

by  
Daniel Christlein

---

A Dissertation Submitted to the Faculty of the  
DEPARTMENT OF ASTRONOMY  
In Partial Fulfillment of the Requirements  
For the Degree of  
DOCTOR OF PHILOSOPHY  
In the Graduate College  
THE UNIVERSITY OF ARIZONA

2004

UMI Number: 3145054

### INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

**UMI<sup>®</sup>**

---

UMI Microform 3145054

Copyright 2004 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company  
300 North Zeeb Road  
P.O. Box 1346  
Ann Arbor, MI 48106-1346

The University of Arizona ®  
Graduate College

As members of the Final Examination Committee, we certify that we have read the

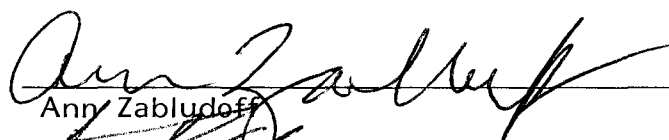
dissertation prepared by Daniel Christlein

entitled Disentangling Luminosity, Morphology, Star Formation,

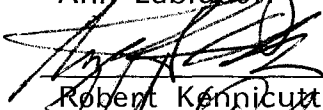
Stellar Mass, and Environment in Galaxy Evolution

and recommend that it be accepted as fulfilling the dissertation requirement for the

Degree of Doctor of Philosophy

  
Ann Zabludoff

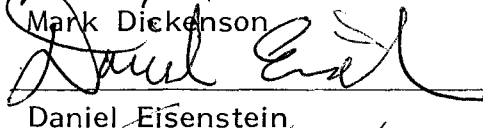
8/4/04  
date

  
Robert Kennicutt


8/4/04  
date

  
Mark Dickenson

8/4/04  
date

  
Daniel Eisenstein

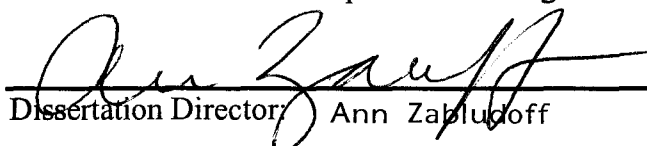
8/4/04  
date

  
William Tifft

8/5/04  
date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

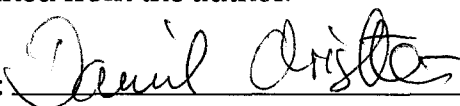
  
Dissertation Director: Ann Zabludoff

8/4/04  
date

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: \_\_\_\_\_

## ACKNOWLEDGMENTS

I would like to thank, first and foremost, my advisor and collaborator in the work presented here, Ann Zabludoff, for her guidance, support, persistence over many a manuscript revision, and her excellent dataset. Ann's work was a major motivation for me to come to Arizona, and I have never rued that decision. Further thanks go to Dennis Zaritsky, Dan Eisenstein, José Arenas, and Anthony Gonzalez for comments on, discussions about, and contributions to the work presented here, and especially Luc Simard for his GIM2D code and Dan McIntosh for his *U*-band photometry, both of which are important pillars upon which this work rests.

I also gratefully acknowledge financial support from a Graduate Assistantship and a Graduate College Fellowship from the University of Arizona, and from Ann's grants (NSF grant # AST-0206084 and NASA LTSA grant # NAG5-11108). Extra thanks to Michelle, Catalina, and Joy for fighting many bureaucratic battles and being persistently victorious.

Furthermore, I am particularly grateful to Bill Tifft and Chip Arp for reminding me that reality is not bound by theories and for providing a sense of balance and skepticism throughout my graduate school career. And special thanks are due to Kelly and Amy for their friendship and for helping me maintain vestiges of a life outside the office, Anna for teaching me more about life than I learned from five years of graduate school, and, above all, Wolfgang Amadé Mozart for bringing joy and serenity into even the loneliest nightly hour spent before the monitor.

## DEDICATION

This work is dedicated to the memory of Dr. Rainer Christlein (1940-1983).

## TABLE OF CONTENTS

LIST OF FIGURES . . . . .	10
LIST OF TABLES . . . . .	12
ABSTRACT . . . . .	13
CHAPTER 1 INTRODUCTION . . . . .	15
CHAPTER 2 GALAXY LUMINOSITY FUNCTIONS FROM DEEP SPECTROSCOPIC SAMPLES OF RICH CLUSTERS . . . . .	21
2.1 Chapter Summary . . . . .	21
2.2 Introduction . . . . .	22
2.3 The Data . . . . .	25
2.3.1 The Sample . . . . .	25
2.3.2 Spectroscopy . . . . .	26
2.3.3 Imaging . . . . .	29
2.3.4 Image processing . . . . .	29
2.3.5 Photometry . . . . .	31
2.3.6 The Galaxy Catalog . . . . .	34
2.4 Calculating the Luminosity Function . . . . .	36
2.4.1 The Sampling Fraction . . . . .	36
2.4.2 The Individual Cluster GLFs . . . . .	43
2.4.3 The Field and Composite Cluster GLFs . . . . .	44
2.5 Results and Discussion . . . . .	47
2.5.1 Individual Cluster GLFs . . . . .	47
2.5.1.1 Comparisons among Cluster GLFs . . . . .	48
2.5.1.2 Radial Sampling Bias . . . . .	53
2.5.1.3 Aperture Bias . . . . .	54
2.5.1.4 The Uniformity of Cluster GLFs . . . . .	56
2.5.2 Cluster Composite GLF . . . . .	57
2.5.3 Field GLFs . . . . .	60
2.5.4 Comparisons between Field and Clusters . . . . .	63
2.5.5 The Bright End . . . . .	71
2.6 Conclusions . . . . .	73
CHAPTER 3 THE U-BAND GALAXY LUMINOSITY FUNCTION OF NEARBY CLUSTERS . . . . .	75
3.1 Chapter Summary . . . . .	75
3.2 Introduction . . . . .	76

## TABLE OF CONTENTS — CONTINUED

3.3	The Data . . . . .	81
3.3.1	The Cluster Sample . . . . .	81
3.3.2	<i>R</i> -band Survey . . . . .	81
3.3.3	<i>U</i> -band survey . . . . .	82
3.4	Calculating the Luminosity Function . . . . .	84
3.4.1	A New Method . . . . .	84
3.4.2	The Sampling Fraction . . . . .	85
3.4.3	Consistency checks . . . . .	85
3.4.4	Coverage of $(M_R, M_U, \mu_R)$ Parameter Space . . . . .	87
3.5	Results and Discussion . . . . .	89
3.5.1	<i>U</i> -band GLFs for Individual Clusters . . . . .	89
3.5.2	Composite <i>U</i> -band GLFs . . . . .	92
3.5.3	Comparison to <i>R</i> -band GLFs . . . . .	95
3.5.4	Contribution from Clusters and Groups to Near-UV Back-ground . . . . .	100
3.6	Conclusions . . . . .	105
CHAPTER 4 CAN EARLY TYPE GALAXIES EVOLVE FROM FADING THE DISKS OF LATE TYPES? . . . . . 109		
4.1	Chapter Summary . . . . .	109
4.2	Introduction . . . . .	109
4.3	The Data . . . . .	112
4.4	Bulge-Disk Decomposition . . . . .	113
4.5	Calculating the Bulge and Disk Luminosity Functions . . . . .	116
4.6	Results and Discussion . . . . .	117
4.6.1	Bulge- and Disk Luminosity as a Function of Morphology . . . . .	117
4.6.2	Bulge- and Disk Luminosity as a Function of Environment . . . . .	128
4.7	Conclusions . . . . .	130
CHAPTER 5 DISENTANGLING MORPHOLOGY, STAR FORMATION, STELLAR MASS, AND ENVIRONMENT . . . . . 132		
5.1	Chapter Summary . . . . .	132
5.2	Introduction . . . . .	132
5.3	The Cluster Survey . . . . .	137
5.4	Galaxy Properties . . . . .	138
5.4.1	Star Formation Indices . . . . .	138
5.4.2	Morphologies . . . . .	139
5.4.3	Luminosities and Stellar Masses . . . . .	141
5.4.4	Mean Stellar Age . . . . .	144



## TABLE OF CONTENTS — CONTINUED

5.5	The Analysis . . . . .	145
5.5.1	Partial Correlation Coefficients . . . . .	145
5.5.2	Completeness-Corrected Correlation Coefficients . . . . .	148
5.5.3	The Discrete Maximum Likelihood Method . . . . .	150
5.5.4	The Completeness Function . . . . .	152
5.5.4.1	Spectroscopic Incompleteness . . . . .	152
5.5.4.2	Morphological Completeness . . . . .	154
5.5.4.3	Radial Incompleteness . . . . .	155
5.5.4.4	Summarial Completeness Function . . . . .	155
5.6	Results . . . . .	157
5.6.1	Morphology-Environment and Star Formation-Environment Relations . . . . .	157
5.6.2	Residual Star Formation Versus Environment . . . . .	159
5.6.2.1	Residual Correlation of Current Star Formation with Environment . . . . .	161
5.6.2.2	Residual Correlation of Recent Star Formation with Environment . . . . .	163
5.6.2.3	What is the effect of controlling individual vari- ables on the star formation gradient? . . . . .	165
5.6.2.4	Is the residual star formation gradient a major com- ponent of the total star formation gradient? . . . . .	168
5.6.3	Effect of Errors and Uncertainties on the Residual Correla- tions . . . . .	169
5.7	Conclusions . . . . .	174
CHAPTER 6 CONCLUSIONS . . . . .		177
APPENDIX A THE <i>R</i> -BAND PHOTOMETRIC DETECTION CATALOG . . . . .		183
APPENDIX B CORRECTION OF $f_s$ FOR SELECTION IN A DIFFERENT FILTER BAND . . . . .		189
APPENDIX C HIGHER-ORDER CORRECTIONS IN CALCULATING THE SAM- PLING FRACTION . . . . .		192
APPENDIX D THE DISCRETE MAXIMUM LIKELIHOOD METHOD . . . . .		194
APPENDIX E DETERMINATION OF THE SAMPLING FRACTION . . . . .		204
APPENDIX F HOW MUCH LIGHT FROM FAINTER GALAXIES? . . . . .		209

## TABLE OF CONTENTS — CONTINUED

APPENDIX G HOW MUCH LIGHT FROM LARGER RADII? . . . . .	215
REFERENCES . . . . .	217

## LIST OF FIGURES

2.1	Optical versus HI radial velocities . . . . .	28
2.2	Averaged Galaxy Sampling Fraction as a function of $(m_R; \mu_R)$ . . .	39
2.3	Cluster GLFs for All Spectral Types . . . . .	49
2.4	Cluster GLFs for Emission-Line Galaxies . . . . .	50
2.5	Cluster GLFs for Non-Emission-Line Galaxies . . . . .	51
2.6	Composite GLFs for Field and Clusters . . . . .	58
2.7	Error Contours for Schechter Fits . . . . .	61
2.8	Fraction of Emission-Line Galaxies as a Function of $M_R$ . . . . .	69
3.1	Comparison Between Rest-Frame and Observed Filter Bands . . . .	78
3.2	Comparison of $R$ -band GLFs Recovered with Different Methods . .	86
3.3	Galaxy Distribution in the $(M_U, M_R)$ Plane . . . . .	88
3.4	$U$ -band GLFs for A496, A754, A85 . . . . .	90
3.5	$U$ -band GLFs for All, Non-Emission-Line, and Emission-Line Galaxies . . . . .	93
3.6	Superposition of $U$ -band and $R$ -band GLFs . . . . .	97
3.7	Superposition of $U$ -band and $R$ -band GLFs for EL and NEL Galaxies	98
3.8	Integrated $U$ -band Luminosity as a Function of Cluster Mass . . . .	102
3.9	$U$ -band Luminosity Density from Groups and Clusters . . . . .	106
4.1	Distribution of Measured $B/T$ . . . . .	115
4.2	Disk Luminosity Functions . . . . .	120
4.3	Bulge Luminosity Functions . . . . .	121
4.4	$M_R^*(\alpha = -1)$ as a Function of $B/T$ . . . . .	122
4.5	$M_{bulge}^*(\alpha = -1)-x$ and $M_{disk}^*(\alpha = -1)-x$ Relations . . . . .	126
5.1	Completeness Functions of the Morphological Catalog . . . . .	156
5.2	Morphology-Environment Relation . . . . .	158
5.3	Star Formation-Environment Relation . . . . .	160

## LIST OF FIGURES — CONTINUED

5.4	Residual Star Formation-Environment Relation . . . . .	164
5.5	Residual Recent Star Formation-Environment Relation . . . . .	166
F.1	Integrated $U$ -band Luminosity as a Function of Magnitude Limit .	212
G.1	Integrated $U$ -band Luminosity as a Function of Sampling Radius .	216

## LIST OF TABLES

2.1	The Cluster Sample . . . . .	26
2.2	Galaxy Catalog (Example) . . . . .	35
2.3	Composite GLFs. GLF values given in decadic logarithms; number of galaxies given in parentheses. . . . .	59
2.4	Schechter parameters . . . . .	59
2.5	Field vs. Cluster comparisons. Probabilities of same parent distributions under $\chi^2$ and parametric comparisons. . . . .	64
2.6	Comparisons between individual clusters and field GLF, $m_R \leq 18$ . . . . .	66
3.1	The Cluster Sample . . . . .	81
3.2	<i>U</i> -band GLFs. . . . .	91
3.3	Composite <i>U</i> -band GLFs. . . . .	94
3.4	Integrated <i>U</i> -band Luminosities as a Function of Absolute Magnitude Limit . . . . .	103
4.1	The Cluster Sample . . . . .	113
4.2	Morphological Subsamples and Schechter Fits . . . . .	118
5.1	The Cluster Sample . . . . .	138
5.2	Spectroscopic Catalog (Example) . . . . .	140
5.3	Morphological Catalog (Example) . . . . .	142

## ABSTRACT

We present a study of the photometric and spectroscopic properties of galaxies in a sample of six nearby, rich galaxy clusters. We examine the variations of fundamental galaxy properties, such as luminosity, morphology, and star formation rates with environment, providing new constraints on the mechanisms that drive the evolution of galaxies. This study also introduces a new maximum likelihood algorithm to recover the true distribution function of galaxies from an incomplete sample. This algorithm is ideally suited for modern-day surveys that gather a large amount of information about each object.

The *R*-band luminosity function (GLF) shows no variation among clusters or between the field and clusters, with the exception of an enhancement of the luminous tip of the GLF in clusters. However, the GLF of quiescent galaxies steepens significantly between the field and clusters and is not universal in clusters either, suggesting that star formation properties may be more strongly correlated than the luminosity function with environment. The *U*-band GLF in clusters is slightly steeper than the *R*-band GLF, indicating that cluster galaxies are bluer at fainter magnitudes and that the GLF is thus weakly sensitive to star formation, dust, or metallicity effects. To constrain the mechanisms that shape the morphologies of cluster galaxies, we have calculated separate *R*-band luminosity functions for galaxy bulges and disks. Their distribution as a function of morphology and environment indicates that intermediate- and early-type galaxies can be generated from late-type galaxies by increasing the luminosity of the bulge, but not by fading the disks alone, favoring galaxy-galaxy interactions or mergers as the primary morphological transformation mechanism. Finally, we find a residual correlation of star formation with environment even after accounting for environmental

variations of morphology, stellar mass, and stellar age. Thus, the star formation gradient in clusters is not just another aspect of the morphology-density relation, and cannot be solely the result of initial conditions, but must partly be due to subsequent evolution through a mechanism (or mechanisms) sensitive to environment. These results thus constitute a true “smoking gun” pointing to the effect of environment on the later evolution of galaxies.

## CHAPTER 1

### INTRODUCTION

The first observations that the properties of galaxies are not the same throughout the Universe, but are correlated with their environment, were made early in the 20th century (Hubble & Humason, 1931). The morphologies, star formation properties, and possibly luminosities and/or stellar masses, of galaxies vary as a function of environment. This is reflected in the well-known morphology-environment relation (Dressler, 1980) (early-type galaxies, such as ellipticals and S0s, constitute a higher fraction of the galaxy population in dense environments than in sparse field environments), in the star formation-environment relation (Lewis et al., 2002; Gómez et al., 2003) (galaxies in denser environments have lower star formation rates), and in the debate about luminosity and mass segregation (Biviano, 2000) (whether the mass function is skewed towards more massive galaxies in denser environments).

It is clear that environment is connected to the properties of galaxies, through their initial formation and during their later evolution. Environmental transformation mechanisms may exist and could be active even today. However, the ways in which environment affects the lives of individual galaxies are poorly understood. This is in no small part due to the fact that, over the past decades, statistical investigations of galaxy properties have been limited in their scope and methodology.

For example, the question of whether there is luminosity or mass segregation with environment (Rood & Turnrose, 1968; Rood, 1969; Bahcall, 1973; Sarazin,



1980; Biviano et al., 1992) has not been resolved. Correlations between luminosity or stellar mass and environment provide constraints on the clustering and merging processes that drive the formation and evolution of structure in the universe in popular hierarchical clustering scenarios (Mo et al., 2004). Such a correlation could be detected as a variation of the luminosity function between environments such as the field and rich clusters of galaxies (Lugger, 1986; Colless, 1989; Gaidos, 1997; Valotto et al., 1997; Trentham, 1998; Paolillo et al., 2001; Christlein & Zabludoff, 2003; de Propris et al., 2003). However, many past determinations of the galaxy luminosity function in clusters have been unreliable because they used statistical background subtraction, rather than proper membership identification by redshift measurements (Christlein & Zabludoff, 2003; de Propris et al., 2003), to remove contamination of the sample by fore- and background galaxies. Insufficiently subtracted background contamination can bias a luminosity function estimate in an optically selected cluster sample (Valotto, Moore & Lambas, 2001) and may be responsible for the wide range of results quoted in the literature.

The question of luminosity segregation is inseparable from the question of which filter band the luminosity is measured in. Many past determinations of the luminosity function in clusters have used relatively blue filter bands, which are sensitive to star formation and dust attenuation, while red filter bands are more representative of the stellar mass. How reliable are luminosity function estimates in blue magnitude bands? Are they primarily sensitive to stellar mass or to star formation, dust, or metallicity? Finding an answer to these questions requires a systematic comparison of luminosity functions in different bands.

Likewise, although the morphology-environment relation (Dressler, 1980) is one of the best-established statistical properties of galaxies, its drivers are not well understood. Past statistical studies of the morphological properties of galax-

ies (Dressler, 1980; Solanes, Salvador-Solé & Sanromà, 1989; Benson, Frenk & Sharples, 2002) have often relied on small samples or on subjective morphological classifiers, such as Hubble types. With these methods, it is very difficult to separate real physical effects from biases introduced by the classification process. The systematic study of the morphological properties of galaxies has become feasible only through computational advances (Simard et al., 2002; Peng et al., 2002), which make it possible to determine quantitative morphological indices for large samples of galaxies. Even with this information available, further insight into the origin of the morphology-environment relation can only be gathered from simultaneously considering other variables, such as the luminosities of bulges and disks, in addition to morphology in order to understand how the physical properties of individual components of galaxies themselves vary along the morphological sequence and with environment.

In recent years, correlations of properties such as morphology (Dressler, 1980) and star formation (Balogh et al., 1998; Hashimoto et al., 1998; Lewis et al., 2002; Gómez et al., 2003) with environment have often been interpreted as the result of post-formation environmental mechanisms like mergers (Barnes, 1999; Bekki, 1998; Mihos & Hernquist, 1994), ram-pressure stripping (Gunn & Gott, 1972), or strangulation (Larson, Tinsley & Caldwell, 1980; Balogh, Navarro & Morris; Bekki, Couch & Shioya, 2002) acting on and shaping the fates of galaxies. However, these observed correlations do not provide unambiguous evidence for environmental processes acting at late epochs as they could also result from environment affecting the *formation* processes of galaxies themselves (Whitmore, 1990; Thuan et al., 1992). The observation that there is a gradient of galaxy properties with environment does not provide evidence that there are transformation processes specific to certain environments either. It is possible that any influences

on galaxy evolution simply produce more early-type galaxies in dense environments, and that all observed gradients — the morphology-environment relation, the star formation-environment relation, possibly luminosity and mass segregation — are simply a consequence of the same mechanisms. Whether that is the case, or whether environment-specific mechanisms exist that truly affect the late evolution of a galaxy, or whether all correlations between environment and galaxy properties are already seeded by initial conditions, are questions that can only be addressed by considering all these properties of galaxies at the same time.

One of the reasons for the surprising scarcity of unambiguous fundamental observational constraints on these questions is that past early galaxy surveys often contained little information on any individual galaxy beyond an apparent magnitude and, sometimes, a redshift. Furthermore, many mathematical tools that are in use today (Sandage, Tamman & Yahil, 1979; Efstathiou, Ellis & Peterson, 1988) were developed specifically to analyze simple distribution functions of galaxies over one or, at most, two variables (e.g., luminosity functions).

With the advent of large surveys containing a multitude of observables for each individual object (Shectman et al., 1996; Colless, 1999; York et al., 2000), the study of galaxy properties matures from simple studies of univariate distribution functions such as the luminosity function to true multi-variate distribution functions that analyze a multitude of parameters for any given galaxy at once, including luminosity, star formation properties, morphology, and environment. Such multi-variate studies have the potential to provide much better constraints on the mechanisms that control the lives and fates of galaxies.

In this thesis, we present an analysis of a survey of six nearby clusters (Christlein & Zabludoff, 2003). The dataset combines photometric and spectroscopic information, and thus makes it possible to not only examine the dependence of ba-

sis galaxy properties, such as luminosity, on environment, but also to separate variations of luminosity, morphology, stellar mass, mean stellar age, and star formation from each other, which provides new constraints on the mechanisms that control the evolution of galaxies. We also introduce a new maximum likelihood algorithm (Christlein, McIntosh & Zabludoff, 2004) for calculating completeness corrections in galaxy surveys. This algorithm, unlike previous maximum likelihood algorithms that were optimized to recover galaxy distribution functions only over one or two variables, has been designed specifically for use in multivariate problems.

The outline of this thesis is as follows: In Chapter 2, we address the basic question of the galaxy luminosity function and its variation among environments, specifically, between the field and clusters, and among clusters of different masses. We ask whether the *R*-band luminosity function of galaxies varies between clusters and the field, whether star-forming and quiescent galaxies are affected to different degrees by such variations, and whether the luminosity function is similar in all clusters or sensitive to environment even in the high-density regime.

In Chapter 3, we introduce our new statistical method, the discrete maximum likelihood (DML) method and apply it to calculating the galaxy luminosity function in the *U*-band. This allows us to compare luminosity function estimates for the same sample in two different filter bands and to constrain how sensitive luminosity functions are to effects such as star formation, dust, or metallicity. Thus, we can assess how reliable luminosity functions are as measures of the physical properties of galaxies.

In Chapter 4, we add information about the structural parameters of our sample galaxies into the analysis. This allows us to examine how the properties of

bulges and disks in our galaxies vary independently as a function of overall morphology and of environment. We address the question whether the morphological sequence of galaxies is generated primarily by variations of the bulge luminosity or the disk luminosity with environment. This provides a direct observational constraint on the mechanisms, such as ram-pressure stripping or tidal interactions between galaxies, that may be responsible for the morphological sequence. However, it does not resolve the question of whether these properties of galaxies are fixed during their initial formation or by such subsequent environmental transformations. Nor does it address whether other properties of galaxies, such as star formation, simply vary as a result of the same evolutionary mechanisms, or whether additional mechanisms are required to understand their variations with environment.

We address these last two questions in Chapter 5, where we simultaneously consider spectroscopic and photometric information to separate the star formation-environment relation from the environmental variations of morphology, stellar mass, and mean stellar age. This allows us to ask whether galaxies with similar morphologies, stellar masses, and star formation histories still have different star formation histories in different environments. We can thus test directly for any environmental influences on galaxy evolution. Furthermore, by fixing morphology, stellar mass, and stellar age, we remove any correlations that could have been seeded by initial conditions, and therefore isolate any signatures of late-term environmental transformation mechanisms.

In the Conclusions, we summarize what we have learned from our investigations and describe several outstanding issues that we plan to follow-up.

## CHAPTER 2

# GALAXY LUMINOSITY FUNCTIONS FROM DEEP SPECTROSCOPIC SAMPLES OF RICH CLUSTERS

### 2.1 Chapter Summary

Using a new spectroscopic sample and methods accounting for spectroscopic sampling fractions that vary in magnitude and surface brightness, we present R-band galaxy luminosity functions (GLFs) for six nearby galaxy clusters with redshifts  $4000 < cz < 20000$  km/s and velocity dispersions  $700 < \sigma < 1250$  km/s. In the case of the nearest cluster, Abell 1060, our sample extends to  $M_R = -14$  (7 magnitudes below  $M^*$ ), making this the deepest spectroscopic determination of the cluster GLF to date. Our methods also yield composite GLFs for cluster and field galaxies to  $M_R = -17$  ( $M^* + 4$ ), including the GLFs of subsamples of star forming and quiescent galaxies. The composite GLFs are consistent with Schechter functions ( $M_R^* = -21.14^{+0.17}_{-0.17}$ ,  $\alpha = -1.21^{+0.08}_{-0.07}$  for the clusters,  $M_R^* = -21.15^{+0.16}_{-0.16}$ ,  $\alpha = -1.28^{+0.12}_{-0.11}$  for the field). All six cluster samples are individually consistent with the composite GLF down to their respective absolute magnitude limits, but the GLF of the quiescent population in clusters is not universal. There are also significant variations in the GLF of quiescent galaxies between the field and clusters that can be described as a steepening of the faint end slope. The overall GLF in clusters is consistent with that of field galaxies, except for the most luminous tip, which is enhanced in clusters versus the field. The star formation properties of giant galaxies are more strongly correlated with the environment than those of fainter galaxies.

## 2.2 Introduction

The galaxy luminosity function (GLF) is one of the basic statistics of the properties of galaxies, and variations of the GLF as a function of environment provide important constraints on any attempt to model galaxy evolution. Despite recent progress on the determination of the field GLF (Blanton et al., 2001; Madgwick et al., 2001), less is known about the GLF in high density environments like rich clusters. In particular, there is controversy as to whether 1) the GLF in clusters is universal, 2) whether the GLF of clusters differs from the field, and, 3) which galaxy populations are most responsible for any differences between environments. These questions are unresolved for several reasons: 1) large volume- or magnitude-limited redshift surveys have far better statistics on galaxies in more common, less dense environments than clusters, 2) most past analyses have depended on statistical background subtraction, which is sensitive to cosmic variance (Valotto, Moore & Lambas, 2001), to constrain the faint end, 3) past work has often used B-band magnitudes, which are sensitive to dust and recent star formation, and 4) comparisons among clusters, or between clusters and the field, have been inhomogeneous, with data gleaned from different datasets. In this paper, we aim to address these problems with the first deep, spectroscopic determination of the R-band GLF for six clusters of galaxies and their associated fields.

Past studies examining the universality of the GLF in clusters have yielded conflicting results (Trentham, 1998; Smith, Driver & Phillipps, 1997; Driver, Couch & Phillipps, 1998). A definitive answer to this question requires not only a reliable determination of the mean GLF in clusters, but also strong constraints on the GLFs of individual clusters to at least 3-4 magnitudes below  $M^*$  (the characteristic magnitude at which the exponential cutoff at the bright end of the Schechter function (Schechter, 1976) begins to dominate over the power law describing the

faint end) in order to constrain their scatter around the mean. Semi-analytic models that attempt to reproduce the GLF in clusters (Springel et al., 2001) are so far limited to simulating relatively small numbers of systems and are thus affected by cosmic variance. An estimate of the observational variance of the cluster GLF will be of great relevance to judging the quality of such models.

It is known from other statistics, such as the morphology-density relation (Dressler, 1980), that the properties of galaxies in clusters are different from those in the field. Are these differences reflected in the shape of the GLF? This question has not been resolved unambiguously either. Zabludoff & Mulchaey (2000) find a consistency between poor group and rich cluster GLFs, and Muriel, Valotto & Lambas (1998) find group GLFs to be consistent with the field GLF. Christlein (2000) finds a systematic and continuous variation of the GLF faint end slope over a range of environments that mostly covers poor groups from the Las Campanas Redshift Survey (Schechter et al., 1996), but does not have sufficient data to extend the analysis to high-mass systems. To make a proper comparison from the highest to lowest density environments requires a field GLF that has been obtained from the same survey and processed using the same criteria (such as star/galaxy separation and surface brightness limits). Such an approach guarantees internal consistency and avoids the problems associated with comparing GLFs across different surveys (for example, see Blanton et al. (2001) for a discussion of the effect of different surface brightness cutoffs on GLF determinations).

A related question is whether it is possible to identify particular populations of galaxies (*e.g.*, early or late types) that individually show variations with the environment. This test would be a potentially even stronger constraint on galaxy evolution models than the total GLFs alone. Furthermore, we would like to know whether any differences between the field and cluster GLFs arise solely from the



morphology-density relation, from mixing differently populations with universal, type-specific GLFs (Bromley et al., 1998), or whether the type-specific GLFs themselves depend on environment.

In the present paper, we measure cluster GLFs for six nearby ( $cz < 20000$  km/s) clusters and their surrounding fields. Our study is based on a spectroscopic sample that includes 300-500 galaxies per cluster, thus making statistical background subtraction unnecessary. The samples extend to  $M_R \approx -18$  ( $M^* + 3$ ) for the highest-redshift cluster, A3266, and to  $M_R \approx -14$  ( $M^* + 7$ ) in the case of the lowest-redshift cluster, A1060. For the purposes of determining composite cluster and field GLFs and comparing them, we impose magnitude limits that typically restrict our analysis to  $M_R \leq -17$ . Of 1860 spectroscopically confirmed cluster members, 1563 are within the magnitude limits of this study.

The clusters in our sample span a range of velocity dispersions ( $700 < \sigma < 1250$  km/s), providing a significant baseline for studies of any variations in the shapes of the GLF with cluster environment. A set of 703 galaxies confirmed non-members from the cluster fields allow for a self-consistent comparison between the field and cluster GLFs.

Our study provides complementary results to an independent study, also based on a spectroscopic sample, by de Propris et al. (2003) of cluster GLFs in the  $b_J$ -band. Our R-band results are not strongly biased by the current or most recent star formation history of a galaxy or by its dust content. The R-band is more sensitive to the total stellar mass than bluer bands. At the same time, the choice of the R-band allows for deeper samplings of the GLF than studies in the infrared (de Propris et al., 1999; Kochanek et al., 2001), which provide an even better (though not perfect) representation of the total stellar mass (Bell & de Jong, 2001).

In §2, we first describe the six cluster samples. We then review how we obtained and reduced the data, including the procedures used to compile the detection catalog. We took particular care to optimize our photometry and star/galaxy separation. We describe the calculation of the cluster and field luminosity functions in §4, with special care given to the treatment of the sampling fraction,  $f_s$ . We present our results in §5, including the six cluster GLFs, split into subsamples by their spectral properties, as well as the composite cluster and field GLFs. Finally, we test the six cluster samples for consistency with the composite cluster GLF to determine if the latter can serve as a common parent distribution for the galaxy populations in our clusters, and we compare the field and cluster composite GLFs using several tests.

## 2.3 The Data

### 2.3.1 The Sample

Our dataset is a spectroscopic survey of galaxies in the fields of six low-redshift ( $cz \leq 20,000$  km/s) galaxy clusters. These clusters were selected based on 1) their visibility from Las Campanas, 2) the availability of some prior spectroscopic and X-ray data in the literature, 3) their redshifts, which allowed us to sample a large fraction of the virial radius with the  $1.5 \times 1.5$  degree field of the fiber spectrograph field, and 4) their range of velocity dispersions, which suggest a wide range of virial masses. The properties of these clusters (for  $H_0 = 100$  km s<sup>-1</sup> Mpc<sup>-1</sup>,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ , as applied throughout this paper) are given in Table 2.1. In this table,  $\Delta m$  is the distance modulus that we adopt,  $\sigma$  the internal velocity dispersion of the cluster, and  $r_{\text{sampling}}$  the spectroscopic sampling radius.

Table 2.1. The Cluster Sample

Cluster	N	$\bar{cz}$ [km/s]	$\Delta m$ [mag]	$cz$ range [km/s]	$\sigma$ [km/s]	$r_{sampling}$ [Mpc]
A1060	252	$3683 \pm 46$	32.85	2292 - 5723	$724 \pm 31$	0.48
A496	241	$9910 \pm 48$	35.03	7731 - 11728	$728 \pm 36$	1.24
A1631	340	$13844 \pm 39$	35.78	12179 - 15909	$708 \pm 28$	1.71
A754	415	$16369 \pm 47$	36.16	13362 - 18942	$953 \pm 40$	2.00
A85	280	$16607 \pm 60$	36.19	13423 - 19737	$993 \pm 53$	2.03
A3266	331	$17857 \pm 69$	36.35	14129 - 21460	$1255 \pm 58$	2.19

### 2.3.2 Spectroscopy

We selected targets for the spectroscopic sample by running the FOCAS software (Jarvis & Tyson, 1981) on Digitized Sky Survey (DSS) plates of the survey region to detect diffuse objects and obtain approximate photometry in the  $b_j$  band. We describe in §3.1 the effect of target selection in  $b_j$  on the determination of R-band GLFs. We prioritized these targets in order of increasing magnitude. We carried out the spectroscopic observations with the multifiber spectrograph Shectman et al. (1992) at the 2.5m DuPont telescope at the Las Campanas Observatory (LCO), targetting each field multiple times to ensure that no galaxies were lost to fiber crowding problems.

We extract, flat-field, wavelength-calibrate and sky-subtract (based on the flux normalization of the 5577 Å 5890 Å and 6300 Å night sky lines) each spectrum. The spectra have a resolution of  $\sim 5$ -6 Å, a pixel scale of  $\sim 3$  Å, and a wavelength range of 3500-6500 Å. The average signal-to-noise ratio (S/N) in the continuum around the  $H\beta$   $\lambda 4861$ ,  $H\gamma$   $\lambda 4340$ , and  $H\delta$   $\lambda 4102$  absorption lines is typically  $\sim 8$  (calculated by determining the ratio of the mean square deviation about the continuum at the absorption line, after excluding the absorption line and any

nearby sky lines). The fiber aperture is  $3.5''$ .

We determine the radial velocities using the cross-correlation routine XCSAO and the emission-line finding routine EMSAO in the RVSAO package in IRAF (Mink & Wyatt, 1995). The velocities in Tab. 2.2 are either emission-line velocities, absorption-line velocities, or a weighted average of the two (for a discussion of the cross-correlation templates and the spectral lines typically observed, see Shectman et al. (1996), §2.2; Lin (1995)). We compute velocity corrections to the heliocentric reference frame with the IRAF/HELIO program.

We estimate the velocity zero-point correction and external velocity error by comparing our velocities with H I velocities from NED. Fig. 2.1 shows the residuals for 61 galaxies (22 galaxies in A1060 from McMahon (1993) and 39 galaxies in the fields of poor groups, which were observed with the same instrumental setup (Zabludoff & Mulchaey, 1998)) as a function of our internal velocity error estimate. We use only those H I velocities with quoted errors of  $<30 \text{ km s}^{-1}$ . The mean residual of  $11 \text{ km s}^{-1}$  (*solid line*) is small compared with the rms deviation of the residuals ( $\sim 76 \text{ km s}^{-1}$ ) and is consistent with the mean residual of the 336 stars ( $52 \text{ km s}^{-1}$ ) that were serendipitously observed with the same instrument (*dashed line*). Therefore, we do not apply a zero-point correction to the velocities.

We adopt the rms deviation of the residuals ( $\sim 80 \text{ km s}^{-1}$ ), which is constant over the range of internal errors, as the true velocity error when the internal or NED error is smaller than  $80 \text{ km s}^{-1}$ . Otherwise, we list the internal or NED error. Our error estimates are consistent with the average external error estimate of  $70 \text{ km s}^{-1}$  for the Las Campanas Redshift Survey (Shectman et al., 1996) and with Zabludoff & Mulchaey (1998), which both employ the same fiber spectrograph setup.

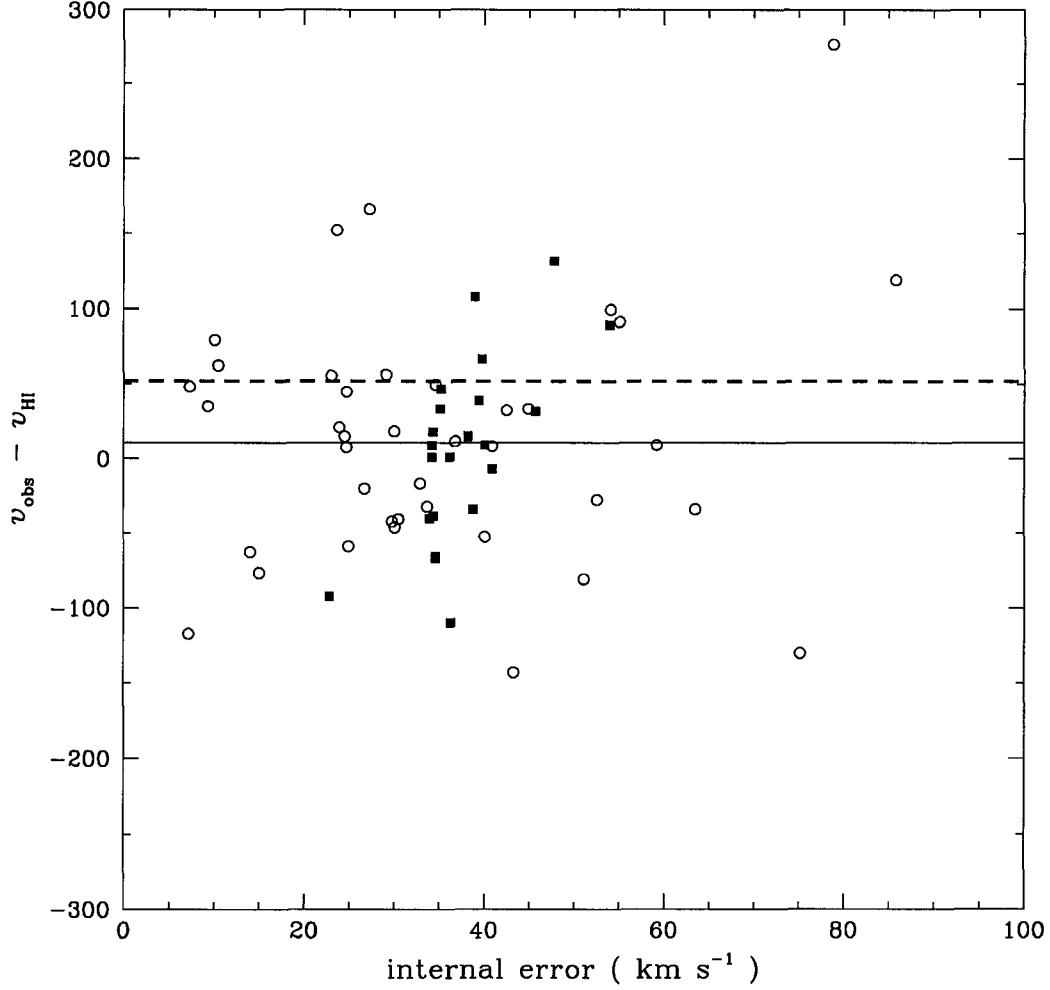


Figure 2.1: Residual radial velocities of 61 galaxies relative to H I redshifts from NED as a function of internal velocity error. The residuals are consistent with a Gaussian distribution with  $\sigma=76 \text{ km s}^{-1}$ . Filled boxes are galaxies in A1060, while open circles are galaxies from Zabludoff & Mulchaey (1998). The solid line shows the mean residual. The dashed line shows the mean residual of 336 stars observed with the same instrument, consistent with the distribution of galaxy residuals.

### 2.3.3 Imaging

We imaged each cluster in a mosaic of 25 tiles, which yielded a complete R-band photometric catalog of galaxies. The R-band photometry is both higher-resolution and less sensitive to recent star formation history than the  $b_J$ -band photometry obtained from DSS images. The R-band catalog defines the sample completeness of the spectroscopic survey. We discuss possible biases that may result from the fact that we used different magnitude bands for the target selection and the final photometric catalog in §3.1. We obtained the bulk of the R-band imaging data with the 40" Swope telescope and the TEK4 CCD at the Las Campanas Observatories. Some missing exposures were provided by Dennis Zaritsky with the same instrument and by Jose Arenas, using the SITe#3 CCD at the same telescope. Typical seeing for most exposures was 1-2 arcseconds, with a pixel scale of 0.696 arcsec/pixel, sufficiently small to allow a robust star/galaxy separation down to  $m_R \leq 19$ . We took most images under photometric conditions, but even the non-photometric images have a scatter in their photometric calibration of only a few hundredths of a magnitude.

The survey region around each cluster consists of a mosaic of  $5 \times 5$  tiles, covering an area of  $1.5 \times 1.5$  square degrees. The sampling radii in Mpc are also given in Tab. 2.1. The exposure time on each tile is  $2 \times 120$  s or, in a few cases, a single exposure of 240 s.

### 2.3.4 Image processing

We subject each frame to a bias subtraction and flatfielding using sky flats. Where both dome flats and sky flats are available, we use dome flats for the flatfielding and sky flats for a subsequent illumination correction. In a few cases where the standard flat fields are problematic, we create a flat field by combining a large number of object frames and rejecting high pixels to remove all objects. If more

than one exposure is available for a given tile, we combine them by addition.

We repair bad columns in the combined images using the IRAF fixpix task. We treat other artifacts, such as extended gradients emanating from a number of bad columns or pedestals in some regions of the chip, or scattered starlight from some of the brightest galactic foreground stars, using custom made algorithms that remove structures defined either by particular symmetries (such as circularly symmetric scattered light around stars or defects extending over entire columns) or by characteristically large scales of their brightness fluctuations.

From the resulting image, we then construct our catalog of detections and perform photometry. In order to identify and remove cosmic rays from the catalog, we also produce a second combined image by averaging the two exposures (where available) with the minmax rejection algorithm set to reject the highest value among the two frames at each pixel. This procedure produces a combined frame that is virtually free of cosmic rays; we use this frame only to identify spurious detections that are likely to be cosmic rays, but do not extract photometry from it.

We use the Source Extractor software (Bertin & Arnouts, 1996) to construct a catalog of detections from each combined object frame and to perform photometry. An object is detected if at least 5 pixels lie above the threshold of 23.61 mag arcsec<sup>-2</sup>, which corresponds to a deviation of an individual pixel of 1.5 - 2  $\sigma$  over the background on most tiles. Our surface brightness limit is comparable to that of Blanton et al. (2001), and the use of Source-Extractor total magnitudes avoids the biases that they find to be associated with the use of isophotal magnitudes with shallow surface brightness limits.

We use a combination of three different methods to separate stars and galaxies. The first is the ‘stellarity’ flag provided in the Source-Extractor output. This

classifier is based on a neural network algorithm. Our second method is based on the fact that galaxies and stars occupy two different, distinct regions in a plot of apparent magnitude versus surface brightness. For each frame, we visually determine a separatrix that optimizes the separation of stars and galaxies in this parameter space. The third method is visual inspection. We inspect every detection that is not unambiguously identified as a galaxy by the two automatic mechanisms, and classify it as stellar, galactic or uncertain. In cases of conflicts between the two automated methods, our visual classification tends to confirm the Source Extractor classification. Therefore, we count as galaxies any objects that we classify visually as galaxies, as well as those that we visually classify as uncertain, but that are classified as galaxies by Source Extractor.

Details of this procedure and of the construction of our detection catalog in general are provided in Appendix A.

We calibrate the coordinate transformation between image and equatorial coordinates (2000.0) for our detections using the Guide Star Catalog-I (Lasker et al., 1990) and, in few cases, the Guide Star Catalog-II (Morrison et al., 2001). For photometric calibration, we use standard star fields from Graham (1982) and Landolt (1992). The astrometric errors are on the order of the internal consistency of the GSC; we have not encountered any problems matching up detections with spectroscopic targets or with other detections of the same object in overlapping fields.

### 2.3.5 Photometry

For the total magnitude  $m_R$ , we adopt the photometric value provided by Source Extractor as  $m_{best}$ , except in the few instances described in Appendix A. This is an automatic estimate of the total magnitude of the object, determined either from an adaptive aperture magnitude or from a correction to the isophotal magnitude (see Bertin & Arnouts (1996)). For a few objects (not selected systematically), we



have verified the Source Extractor magnitudes using a separate program that performs photometry on individual objects and find the agreement to be very good (*i.e.*, typically accurate to  $<0.05$  mag) for reasonably bright, well-isolated bright objects. We also calculate a surface brightness,  $\mu_R$ , which is the mean surface brightness within our isophotal limit of  $23.61 \text{ mag arcsec}^{-2}$ .

After performing an initial photometric calibration of the magnitude zero point using the appropriate standard star fields for each night, we allow for a small photometric correction for each plate that minimizes the photometric discrepancies in the overlap regions between adjacent tiles (approximately  $4'$  wide).

To determine that correction, we describe the quality of photometric agreement among the 25 tiles by

$$\Phi = \sum_i \sum_j (\delta_{ij} + \Delta_i - \Delta_j)^2 N_{ij} \quad (2.1)$$

Here,  $\delta_{ij}$  is the systematic magnitude difference between stars and galaxies on tiles  $i$  and  $j$ . We must solve for  $\Delta_i$ , the magnitude correction of tile  $i$  (*i.e.*, the systematic offset that is to be subtracted from all magnitude measurements in tile  $i$ ).  $\Phi$  is thus basically a least-squares estimator, weighted by the number  $N_{ij}$  of matches found in the overlap regions. By requiring that  $\partial\Phi/\partial\Delta_i = 0$ , we find

$$\Delta_i = \frac{\sum_j N_{ij} (\Delta_j - \delta_{ij})}{\sum_j N_{ij}} \quad (2.2)$$

This equation is suitable for an iterative solution for the photometric correction  $\Delta_i$ . An additional constraint is that there should be no net magnitude offset over the entire 25 tiles (the results do not differ noticeably if that requirement is restricted to exposures taken on photometric nights).

Even on tiles imaged on different nights, the photometric corrections  $\Delta_i$  are of the order of a few hundredths of a magnitude at most (the *rms* correction is

$\sim 0.035$ ), indicating that the photometric calibration is stable within our typical magnitude errors and that atmospheric extinction effects are minor.

We estimate the random errors remaining in our photometry by comparing the total magnitudes of galaxies in the overlap regions, and by determining the magnitude intervals within which about 2/3 of the comparisons agree. These intervals are dependent on apparent magnitude, and we approximate them by the following empirical results:  $\Delta m = 0.02$  at  $m = 13$ ,  $\Delta m = 0.04$  at  $m = 17$ ,  $\Delta m = 0.08$  at  $m = 18.5$ ,  $\Delta m = 0.11$  at  $m = 20$ . When quoting magnitude errors for individual galaxies, we extrapolate the errors linearly between these data points. Note that these errors reflect only internal consistency.

In addition, we have compared our photometry to literature values in the NED database. For these comparisons, we have used total (preferred) or isophotoal magnitudes to surface brightness limits of 26 or 25 mag arcsec<sup>-2</sup> in the R-band. The magnitude discrepancies for individual galaxies typically scatter between +0 and +0.2 mag (our measurements typically yield the fainter values), although there are outliers at magnitude differences of about 0.4 mag. The distribution of magnitude differences, especially at fainter magnitudes, is non-random, presumably because of the nonhomogeneity of the literature sources.

A systematic offset in the photometry provided by Source-Extractor has been reported before (Daniel McIntosh, 2001, private communication). Bertin & Arnouts (1996) also quote a possible offset as large as 0.06 mag at  $m_R = 17$ . We do not make an attempt to correct for this, as it does not affect our luminosity functions, but this problem should be kept in mind when interpreting the magnitudes that we quote for bright galaxies.

We examine galaxies with large ( $>0.2$  mag) magnitude discrepancies with previously published values by remeasuring their photometry as described in Ap-

pendix A. There is no indication of systematic errors affecting individual galaxies in our sample, revealing no need to revise our photometry for these objects. The two most difficult objects are NGC 3309 and NGC 3311, which are located in the center of A1060 and have extended, overlapping envelopes. We model these galaxies individually using the ELLIPSE and BMODEL tasks in IRAF. The magnitudes that we quote for these two galaxies are based on this flux measurement, not on Source-Extractor.

To correct our magnitudes and surface brightnesses for absorption from foreground Galactic dust, we use the all-sky dust maps and the conversion factor for the CTIO R-band by Schlegel, Finkbeiner & Davis (1998).

### 2.3.6 The Galaxy Catalog

Tab. 2.2 presents a sample from our catalog, listing spectroscopically sampled galaxies with positions, redshifts, and  $m_R$  (uncorrected for galactic extinction). The coordinates listed are the target coordinates for the spectrographic fiber; in cases where the spectroscopic and photometric coordinates deviate by more than  $3''$  (*e.g.*, because of confusion with a nearby star), we list the latter coordinates as a comment. The comments also note whether the object has been deblended from another detection (“deblend”) and/or manually added to the catalog (“add”). The comment “mag!” indicates that the apparent magnitude has been changed from the default Source Extractor value.

From this spectroscopic sample, we select various subsamples. For each of the six cluster fields, the cluster members lie in the redshift ranges in Tab. 2.1. The members are further split into emission line (EL) and non-emission line (NEL) samples. EL galaxies are those with  $\lambda 3727$  [OII] doublet equivalent widths  $> 5\text{\AA}$ ; NEL galaxies are the rest. Thus, EL galaxies represent star forming or active galaxies, and NEL galaxies are relatively quiescent.

Table 2.2. Galaxy Catalog (Example)

ID	RA (J2000)	Dec (J2000)	$m_R$	$\Delta m_R$	$cz$	$\Delta cz$	NED ID	Comments
1060C_494[104]	10 36 42.70	-27 31 42.00	10.07	0.015	3857	80	NGC 3311	10:36:43.21 -27:31:32.09 mag! add
1060A_494[103]	10 36 35.69	-27 31 5.30	10.74	0.015	4071	80	NGC 3309	mag! add deblend
1060A_494[68]	10 37 2.53	-27 33 53.60	11.34	0.015	2761	80	NGC 3312	
1060A_494[91]	10 37 47.30	-27 4 52.00	11.46	0.015	2973	80	IC 2597	
1060A_494[33]	10 33 35.60	-27 27 17.20	11.54	0.015	3295	80	NGC 3285	
1060A_494[96]	10 36 22.31	-27 26 17.50	11.77	0.015	3537	80	NGC 3308	
1060B_494[22]	10 36 57.88	-28 10 38.80	12.24	0.015	2503	80	ESO 437- G 015	10:36:58.02 -28:10:35.58
1060B_494[69]	10 37 37.26	-27 35 38.50	12.36	0.015	3922	80	NGC 3316	
1060B_494[3]	10 36 12.04	-27 9 43.20	12.43	0.015	4002	80	NGC 3305	10:36:11.74 -27: 9:43.92
1060A_494[47]	10 33 30.14	-26 53 50.10	12.62	0.015	3535	80	ESO 501- G 013	
1060A_494[70]	10 37 12.76	-27 41 1.10	12.62	0.015	2795	80	NGC 3314	
1060A_494[78]	10 37 19.17	-27 11 30.50	12.75	0.015	3753	80	NGC 3315	
1060B_494[25]	10 36 50.43	-27 55 8.80	12.78	0.015	4854	80	ESO 437- G 011	
1060A_494[39]	10 34 36.75	-27 39 9.30	12.90	0.015	3150	80	NGC 3285B	10:34:36.97 -27:39:10.25
1060A_494[17]	10 36 24.72	-26 59 57.60	12.97	0.015	4115	80	ESO 501- G 035	
1060A_494[20]	10 36 53.99	-27 54 58.90	13.08	0.015	3625	80	ESO 437- G 013	10:36:53.98 -27:55: 2.16
1060A_494[57]	10 38 33.32	-27 44 12.40	13.12	0.016	4412	80	ESO 501- G 065	
1060A_494[64]	10 39 18.26	-26 50 23.50	13.16	0.016	3113	80	ESO 501- G 068	
1060A_494[6]	10 35 20.48	-27 21 42.90	13.19	0.016	4539	80	ESO 501- G 021	
1060A_494[75]	10 37 4.89	-27 23 59.30	13.28	0.016	2690	80	PGC 031515	deblend
1060A_494[24]	10 36 32.45	-28 3 48.90	13.29	0.016	4362	80	ESO 437- G 008	

Note. — [The complete version of this table is published in Christlein & Zabludoff (2003).]

We also create a composite sample of all cluster galaxies, as well as a composite sample of field galaxies. We define the field sample as all galaxies not explicitly included in any of the cluster samples. As the redshift space in the direction of several of these clusters reveals large-scale structures at different redshifts, the field sample actually represents a range of environments and is not restricted to truly isolated galaxies. A NED database search shows no other major clusters within 1 degree of our lines of sight, but there are several clusters at slightly larger projected distances, indicating that the field sample includes higher density regions as well.

We further complement our catalog with redshift data taken from NED (this adds 86 redshifts to the sample of 1563 cluster galaxies and 46 redshifts to the sample of 703 field galaxies). The inhomogeneity of the literature sources raises concerns that they may be biased towards cluster members, but the overall contributions to our sample are small enough not to constitute a problem. These additional redshifts are used only for the calculation and analysis of the overall GLFs, not for the EL and NEL subsamples, because [OII] equivalent widths are typically not available for the objects supplemented from NED.

## 2.4 Calculating the Luminosity Function

### 2.4.1 The Sampling Fraction

In any statistical investigation of galaxy properties, each galaxy has to be weighted by the inverse of the sampling fraction  $f_s$ , which is the fraction of photometrically detected galaxies obeying certain selection criteria (*e.g.*, cluster membership, luminosity) that have been spectroscopically sampled.

Due to the design of this survey (target selection by apparent magnitude, multiple spectroscopic exposures of each field), the sampling fraction is not depen-

dent on the position of an object on the sky or the proximity of other targets. The primary dependence of  $f_s$  is on apparent magnitude, and possibly on surface brightness. It is possible that systematic discrepancies exist in the star/galaxy classification between the initial target selection (made from lower resolution Digitized Sky Survey (DSS) plates) and the final photometric catalog (made from CCD images as described earlier), but we expect these to be correlated with the position of a galaxy detection in the  $(m_R; \mu_R)$  plane.

We therefore choose to calculate the sampling fraction by counting photometrically detected and spectroscopically sampled galaxies as a function of  $(m_R, \mu_R)$  in overlapping bins of a fixed size on a fine grid in the  $(m_R; \mu_R)$  plane. We thus obtain a (smoothed) estimate of the sampling fraction at every point in the plane where galaxies have been detected.

We inspect visually all objects not identified as stars by both algorithms described in Appendix A. We assign full statistical weight to visually confirmed galaxies, as well as to visually uncertain objects classified as galaxies by Source-Extractor. To be conservative, we also visually inspect objects identified as stellar by both algorithms in up to a third of the mosaic tiles per cluster. We reclassify a small fraction ( $\ll 1\%$ ) of these objects as galaxies and assign this fraction as a fractional statistical weight to the uninspected stellar objects with similar  $(m_R; \mu_R)$  on the remaining tiles. All other (*i.e.* non-stellar) detections that we do not confirm visually are presumed spurious and discarded.

The spectroscopic sample covers a region of  $(m; \mu)$  space in which galaxies are unambiguously identified, so that most of the detections used in the calculation of  $f_s$  are unambiguous galaxies. To assess the impact of the star/galaxy classification on our results, we consider two extremes: objects classified as galaxies by all three methods and objects classified as non-stellar by at least one method.

Between these two extremes, the faint end slope  $\alpha$  of the GLF changes only by  $\sim 0.05$ , and our default option yields results in the middle of this interval.

Fig. 2.2 shows the  $(m_R; \mu_R)$  plane and the average sampling fraction (including the redshifts added from NED) for all six fields. The distribution of all galaxy detections in our catalog is bounded by the light solid line <sup>1</sup>. The figure also shows the region of  $(m_R; \mu_R)$  represented by the spectroscopic sample (bold envelope). The standard magnitude limit for our composite GLFs is  $m_R = 18$ . As we have no information on the redshift distribution of galaxies outside the spectroscopically sampled region, those detections do not contribute to our GLFs. It is obvious from Fig. 2 that such regions are very small; the only population not sampled is one of faint high-surface brightness galaxies that appeared stellar on the lower-resolution DSS plate material and thus were not targetted. Most are fainter than our magnitude threshold of  $m_R = 18$ .

We also note a dependence of the sampling fraction on surface brightness at faint magnitudes. The sampling fraction drops strongly at low-surface brightness end of the distribution, but also exhibits a decline towards the high-surface brightness end. We have randomly looked up several objects in this undersampled region that are clearly galaxies on our R-band exposures; these objects appear faint and star-like on the DSS plates, with no discernible diffuse component. We conclude that the reason for the decline in the sampling fraction with increasing surface brightness lies in the initial star/galaxy separation during the target selection.

A few galaxies stand out in this distribution at unusually high surface brightness values. For these, we find no evidence of errors in the photometry or visual classification. One of these objects with spectroscopic data, 1631A-494[13], is very

---

<sup>1</sup>Note that the sampling in individual fields may vary from this averaged distribution; our analysis therefore treats the sampling fraction as a function of  $m_R$ ,  $\mu_R$ , and cluster field

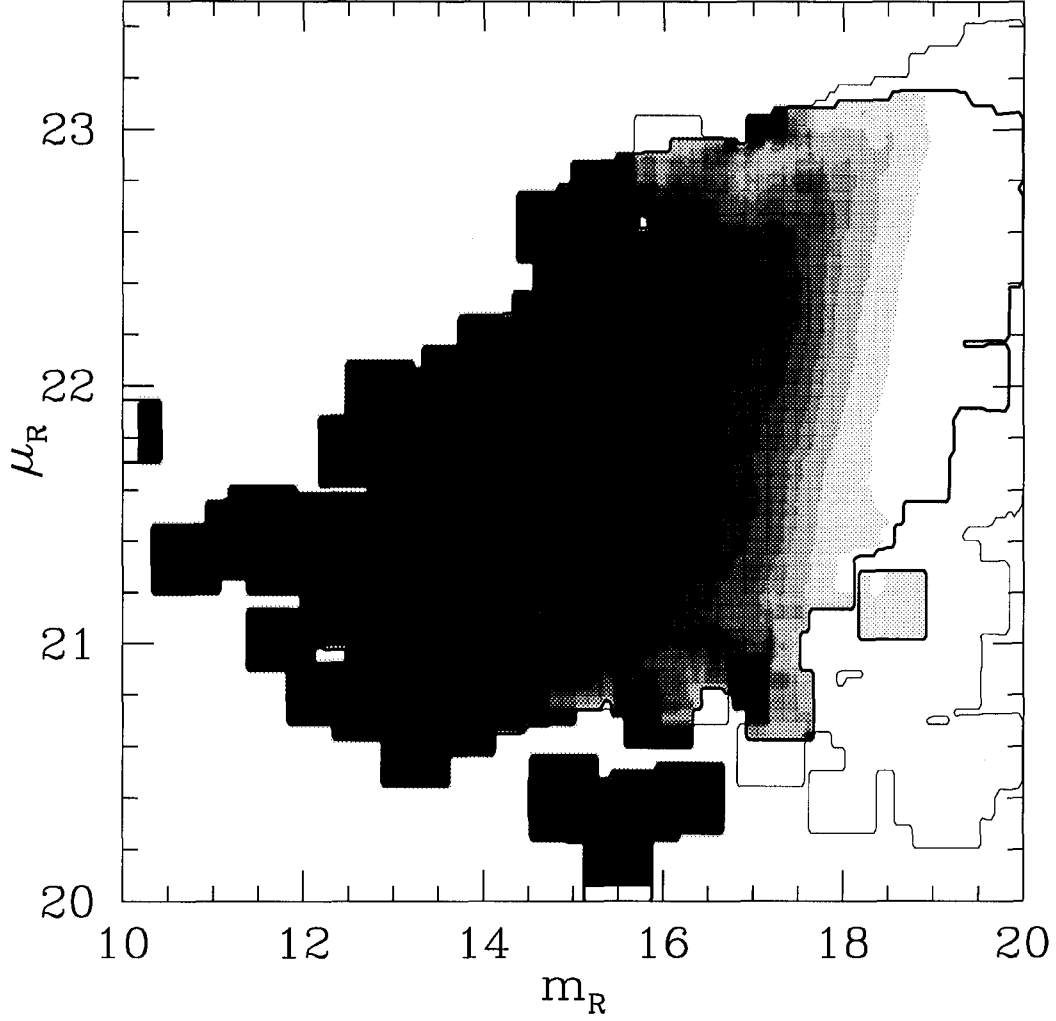


Figure 2.2: Averaged galaxy sampling fraction as a function of  $(m_R; \mu_R)$ . Greyscales indicate the spectroscopic sampling fraction, ranging from 0 to 1. The bold envelope denotes the spectroscopically sampled region (i.e., all  $(m_R; \mu_R)$  for which  $f_s > 0$ ); the light envelope denotes the regions containing photometric detections.



compact, but has a nonstellar profile. This galaxy is superimposed on the envelope of NGC 4756, with a redshift difference of more than 300 km/s. The large difference in surface brightness between the object and the envelope of NGC 4756 makes a confusion of the spectra unlikely. Another object, SERA\_1294[33] (APMUKS(BJ) B042429.42-611817.9), has a visible disk, albeit of very low surface brightness, and a bright, compact bulge. Neither of these objects has unusual activity in its spectrum.

The analysis of the GLF should not be based on detections from poorly sampled regions of  $(m_R; \mu_R)$  space, where possible systematic errors associated with the sampling fraction might become significant. This is particularly important for the composite GLFs, where systematic deviations in the GLF at faint apparent magnitudes could manifest themselves as systematic deviations of the GLF over a wide range of absolute magnitudes. We therefore introduce an apparent magnitude limit for our analysis. We also quantify below some of the major effects that could bias the sampling fraction at faint magnitudes.

For the calculation of the composite GLFs, we truncate our catalog at an apparent magnitude of  $m_R = 18$  for five of our clusters. At this magnitude, the fractional completeness has dropped to  $\sim 20\text{-}50\%$  for the individual clusters. (Over the entire magnitude range down to that limit, the sample is much more complete on average, reaching nearly 100% completeness for  $m_R < 16$ ). An exception is the field of A3266, where the sampling is shallower. Here we truncate at  $m_R = 17$  to achieve a similar level of completeness (although extending this limit to  $m_R = 18$  makes no significant difference in the GLF). We have also experimented with modeling the sampling fraction as a function of U-band magnitudes, which have been provided for three of our clusters by Daniel McIntosh. After correcting the galaxy counts by  $f_S(m_U; \mu_R)$ , we find that residual incompleteness remains

for galaxies with  $m_R > 18$ , indicating that color terms are important for fainter magnitudes. This provides an additional justification for cutting our sample at  $m_R = 18$ .

For the intercomparisons of our GLFs, we adopt similar magnitude limits for reasons of consistency. In particular, when comparing clusters to composite GLFs, we adopt a magnitude limit of  $m_R = 18$  for all six clusters. For A3266, this is fainter than the limit to which it contributes to the composite GLF, but the differences are insignificant both for the calculation of the GLF and for comparisons between the composite GLFs and A3266. We also choose a standard *absolute* magnitude limit for some of our analyses in order to provide a common baseline for all comparisons, rather than the maximum baseline permitted by the apparent magnitude limit. This absolute magnitude limit is  $M_R = -18.35$  ( $-18.5$  in binned distributions), which corresponds to the standard apparent magnitude limit of  $m_R = 18$  at the distance of the furthest cluster, A3266.

A potential source of bias is the fact that target selection for this survey is based on approximate  $b_J$  magnitudes, while our photometric catalog (and thus our calculation of the sampling fraction) uses R-band magnitudes. Cluster galaxies may have colors systematically different than field galaxies. A cluster galaxy of a given  $m_R$  would thus have a different probability of being selected as a spectroscopic target than a field galaxy with the same  $m_R$ . Our sampling fraction, which is based on total counts of galaxies, irrespective of their cluster membership, would then be biased. By assuming a mean color difference between field and cluster galaxies, it is possible to reconstruct the field- and cluster-specific sampling fractions. We use these to estimate the magnitude of this effect and find that it is negligible. Appendix B gives the details of our calculation.

Another potential bias is that our spectroscopic success rate may be correlated

with the spectral properties of the target galaxies and thus indirectly with cluster membership. It is not obvious if this bias would favor the sampling of galaxies in clusters (because of the prevalence of extreme early types with strong absorption features) or in the field (because of the prevalence of extreme late types with strong emission lines). However, the number of spectroscopic targets for which we could not obtain a redshift provides a constraint on the magnitude of any such uncertainty (presuming that the targetting itself is representative). The upper and lower limits on the sampling fraction due to these failed targets show that the impact on our results is minor.

Our procedure for determining  $f_S$  by counting discrete detections in  $(m_R; \mu_R)$  bins of finite size is also subject to several biases. We choose bin sizes to minimize these effects, although a compromise between good statistics and an accurate representation of the sampling fraction is necessary when a smooth, but non-analytic function such as  $f_S(m_R; \mu_R)$  is probed only at discrete points. We explain the details below.

A small bin size is likely to exclude photometric detections in regions with sparse spectroscopic sampling; photometric detections that do not fall into a bin with at least one spectroscopically confirmed galaxy are effectively discarded, while the sampling fraction for galaxies that fall in that bin may be overestimated. Of course, certain populations of galaxies in the  $(m_R; \mu_R)$  plane simply may not have been sampled. These detections should not be included in the calculation of the luminosity function anyway, as their redshift distribution is unknown. We must therefore choose a  $(m_R; \mu_R)$  bin size such that a spectroscopically sampled galaxy is representative of photometric detections in the same region.

A different problem lies in the fact that large bin sizes smooth over variations in the detection density and in the sampling fraction itself across the bin. Higher-

order variations in the density of photometric detections (or spectroscopic galaxies) mean that the integrated number of galaxies across the bin is not representative of the detection density at the bin center. We present a simple procedure to correct this effect to first order in  $m_R$  in Appendix C, but use it only to quantify the magnitude of the effect, which is negligible.

There may also be small-scale variations of the sampling fraction itself in the  $(m_R; \mu_R)$  plane; in particular, the extreme high and low surface brightness regions of the galaxy distribution are undersampled. A large bin size will smooth over these variations, overestimating the sampling fraction in the undersampled regions and underestimating it in the well-sampled regions. The effect on the luminosity function is dependent on how the population of the selected sample — field or cluster galaxies — is distributed in the  $(m_R; \mu_R)$  plane.

Our choice of the bin size for the calculation of the sampling fraction in  $(m_R; \mu_R)$  is motivated by the robustness of the GLF that it produces. The faint end slope of the field GLF is particularly sensitive to the second of the aforementioned effects, the variation of the detection density across the bin, and thus places constraints on the choice of the bin size in  $m$ . A bin size of  $\Delta m = 0.75$  is small enough to avoid these higher-order effects, while not running into the problem of excluding too many photometric detections in sparsely sampled regions. For this bin size,  $\Delta \mu = 0.25 \text{ mag arcsec}^{-2}$  yields robust results for the GLF.

#### 2.4.2 The Individual Cluster GLFs

We calculate the cluster GLFs using bins in  $M_R$  of width  $\Delta M_R = 0.5$  (adopted as a compromise between sufficiently high signal-to-noise in most bins and sufficient resolution of the shape of the GLF). We add up the number of galaxies in each bin, weighted by the inverse of their sampling fraction. This is a justifiable procedure for clusters, as all galaxies are at approximately the same distance and thus do

not require a volume correction.

The finite extent of the lowest redshift cluster in the sample, A1060, leads to a small additional uncertainty in the distance modulus and thus the absolute magnitude of each galaxy. For A1060 and  $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , this uncertainty is on the order of 0.07 mag, much smaller than our bin size.

### 2.4.3 The Field and Composite Cluster GLFs

For calculating the field GLF (and the composite cluster GLF), it is necessary to adopt a different approach, as the observation of galaxies over a wide range of distances requires a volume correction (more luminous galaxies can be seen over larger distances and are thus overrepresented by numbers in a magnitude-limited sample). Weighting by the inverse volume over which a galaxy would be visible would be unlikely to yield good results, as it assumes galaxies are distributed homogeneously in comoving space.

We use a stepwise maximum likelihood (SWML) estimator as described by Efstathiou, Ellis & Peterson (1988). The derivation below is modelled after that given in Lin et al. (1996), with modifications to account for the fact that, instead of fixed apparent magnitude limits, we have a variable sampling fraction as a function of  $(m; \mu)$ . Each galaxy  $i$  in the sample is characterized by an absolute magnitude  $M_i$  and an “absolute surface brightness”  $\mu_i$  (*i.e.*, before cosmological effects or Galactic extinction are applied). Also, each galaxy is associated with a particular redshift, Galactic extinction, and one of our six cluster fields. We parametrize these last three variables, which determine the sampling probability of a galaxy with a given absolute magnitude and surface brightness, by the vector  $F_i$ . The probability that a sampled galaxy with the properties  $F_i$  has the absolute magnitude  $M_i$  and surface brightness  $\mu_i$  is

$$p_i = p(M_i; \mu_i | F_i) = \Phi(M_i, \mu_i) * f_s(M_i; \mu_i; F_i) / \int \Phi(M; \mu) f_s(M; \mu; F_i) dM d\mu \quad (2.3)$$

Note that that  $f_s(M_i; \mu_i; F_i)$  is primarily a function of apparent magnitude and surface brightness and may also vary from one cluster field to another. Given  $F_i$ , the conversion between absolute and apparent variables is unambiguous. The integral is over all of  $(M; \mu)$  space.

The logarithmic likelihood function for a sample consisting of  $N$  galaxies is then

$$\ln \mathcal{L} = \sum_{i=1}^N [\ln(\Phi(M_i, \mu_i) * f_s(M_i; \mu_i; F_i)) - \ln \int \Phi(M; \mu) f_s(M; \mu; F_i) dM d\mu] \quad (2.4)$$

We then discretize  $\Phi$  in  $(M; \mu)$  bins, writing  $\Phi_{kl}$  for  $\Phi(M_k; \mu_l)$ :

$$\ln \mathcal{L} = \sum_i^N \sum_{k,l} \ln(\Phi_{kl} * f_s(M_k; \mu_l; F_i)) * W(M_i; \mu_i; M_k; \mu_l) - \quad (2.5)$$

$$\sum_i^N \ln \left[ \sum_{k,l} \Phi_{kl} f_s(M_k; \mu_l; F_i) * H(M_k; \mu_l; F_i) * \Delta M \Delta \mu \right] \quad (2.6)$$

where  $W$  is 1 if galaxy  $i$  falls into bin  $(k; l)$  and 0 otherwise.  $H$  contains fractional corrections to the bin widths to account for any overlap of the bin with the defined limiting magnitudes of the catalog (particularly the imposed apparent magnitude cutoff at  $m_R = 18$ ). We take the derivative by  $\Phi_{mn}$ :

$$\frac{\partial \ln \mathcal{L}}{\partial \Phi_{mn}} = \sum_i \Phi_{mn} W(M_i; \mu_i; M_m; \mu_n) - \sum_i \frac{f_s(M_m; \mu_n; F_i) H(M_m; \mu_n; F_i) \Delta M \Delta \mu}{\sum_{k,l} \Phi_{kl} f_s(M_k; \mu_l; F_i) H(M_k; \mu_l; F_i) \Delta M \Delta \mu} \quad (2.7)$$

Now we set this expression to zero and solve for  $\Phi_{mn}$  to obtain a prescription for an iterative solution:

$$\Phi_{mn} = \sum_i W(M_i; \mu_i; M_m; \mu_n) / \sum_i \frac{f_s(M_m; \mu_n; F_i) H(M_m; \mu_n; F_i) \Delta M \Delta \mu}{\sum_{k,l} \Phi_{kl} f_s(M_k; \mu_l; F_i) H(M_k; \mu_l; F_i) \Delta M \Delta \mu} \quad (2.8)$$

This maximum likelihood estimator converges fairly quickly. It has the advantage of being unbiased by large-scale structure inhomogeneities, as the sums implicitly trace the redshift distribution of the sample. Its application to clusters has the additional advantage of extracting information even from empty bins, which would be ignored by simply averaging individual GLFs.

When the sampling fraction for a given bin in  $(M_R; \mu_R)$  is referenced by the algorithm, we calculate an average sampling fraction over the area of that bin. In some cases, the algorithm may attempt to reference the sampling fraction at coordinates in the  $(m; \mu)$  plane where no galaxies have been sampled (*e.g.*, to determine the hypothetical visibility of a high-redshift giant galaxy if it were located in a low-redshift field). Therefore, we extrapolate sampling fractions in those unsampled bins prior to calculating the GLF with an iterative algorithm similar to the Liebmann method for solving the Poisson equation on a discrete grid. The effect of this approach on our GLFs, as compared to setting such undefined sampling fractions to zero, is smaller than our quoted uncertainties by more than an order of magnitude.

We do not apply a maximum likelihood estimator (Sandage, Tamman & Yahil, 1979) to calculate parametric luminosity functions (such as Schechter (1976) functions), as there is no standard analytical expression that would allow us to model the galaxy distribution in luminosity and surface brightness simultaneously (but see Cross & Driver (2002) for a proposal for such a two-dimensional analogue to the Schechter function). Instead, we fit Schechter functions to the results of the stepwise maximum likelihood estimator by  $\chi^2$  minimization (after they have been integrated over surface brightness) and estimate the errors in this procedure using the prescription by Avni (1976).

The maximum likelihood method leaves the normalization of the composite

GLFs undetermined, so that they are actually luminosity distributions. Nevertheless, we refer to them as GLFs throughout for reasons of simplicity.

To test the SWML algorithm, we apply it to individual clusters and compare the results to GLFs obtained from the direct binning method described in §3.2. Despite the slightly different treatment of the sampling fraction (for the direct binning method,  $f_s$  is calculated in bins centered on the individual detections, rather than on a regular grid), the GLFs obtained from these two methods are indistinguishable. Furthermore, the field GLFs from the six individual fields are all consistent with the composite field GLF, as would be expected from a sample of similar environments, indicating that our algorithm reproduces the same field GLF even in fields with different redshift distributions and sampling fractions.

## 2.5 Results and Discussion

### 2.5.1 Individual Cluster GLFs

The individual luminosity functions for each of the six clusters are shown in Fig. 2.3 for all spectral types and in Figs. 2.4 and 2.5 for the emission line (EL) and non-emission line (NEL) subsamples, respectively. The error bars denote the Poisson errors of the spectroscopically sampled galaxies, modified by the sampling fraction. The thin solid lines correspond to the uncertainties in each GLF — overall, EL, NEL — due to the failed spectroscopic targets. For example, in the case of the overall GLF, the upper limit assumes that all failed spectroscopic targets are cluster members, and the lower limit that none of them are. We note that these uncertainties are small (typically within the Poisson errors), so that any correlation of the spectroscopic success rate with cluster membership or spectral properties cannot bias our results significantly. The dashed lines represent the most pessimistic scenarios where *all* unsampled galaxies are cluster members (upper



limit) or not (lower limit).

#### 2.5.1.1 Comparisons among Cluster GLFs

We use two different comparison tests to examine whether the GLFs in our six clusters agree to within the statistical uncertainties. First, we compare the individual cluster GLFs to each other using a  $\chi^2$  test. For each of the overall, EL and NEL samples, this test yields 15 independent probabilities. We then analyze the results of these 15 individual comparisons statistically to determine whether they are consistent with the hypothesis that differences between the six clusters are only random, and not systematic.

While this procedure will tell us whether our data is consistent with a universal cluster GLF or not, it does not yield a quantitative estimate of the “universality” of the GLF in clusters (*i.e.*, the fraction of all clusters for which the average GLF is representative). A constraint on this universality is useful in estimating how representative the composite GLFs that we present in §4.2 are for clusters in general. Therefore, we perform a second comparison test, in which we compare the individual clusters to composite GLFs. This test shows how many, and which, clusters in our sample are consistent with the composites, and this in turn allows us to constrain the universality of the composite GLFs. The details of both tests are below.

We turn first to the cluster-to-cluster comparisons. In order to be consistent with our treatment of the composite GLFs, which are presented below, we adopt the same magnitude limits as discussed in §3.1. We impose an apparent magnitude cutoff of  $m_R = 18$  and compare the clusters only over absolute magnitude ranges brighter than this limit. Down to this apparent magnitude, we consider our sampling fractions reliable. For each of the EL, NEL and overall samples, we perform 15 cluster-to-cluster comparisons. If no systematic differences exist be-

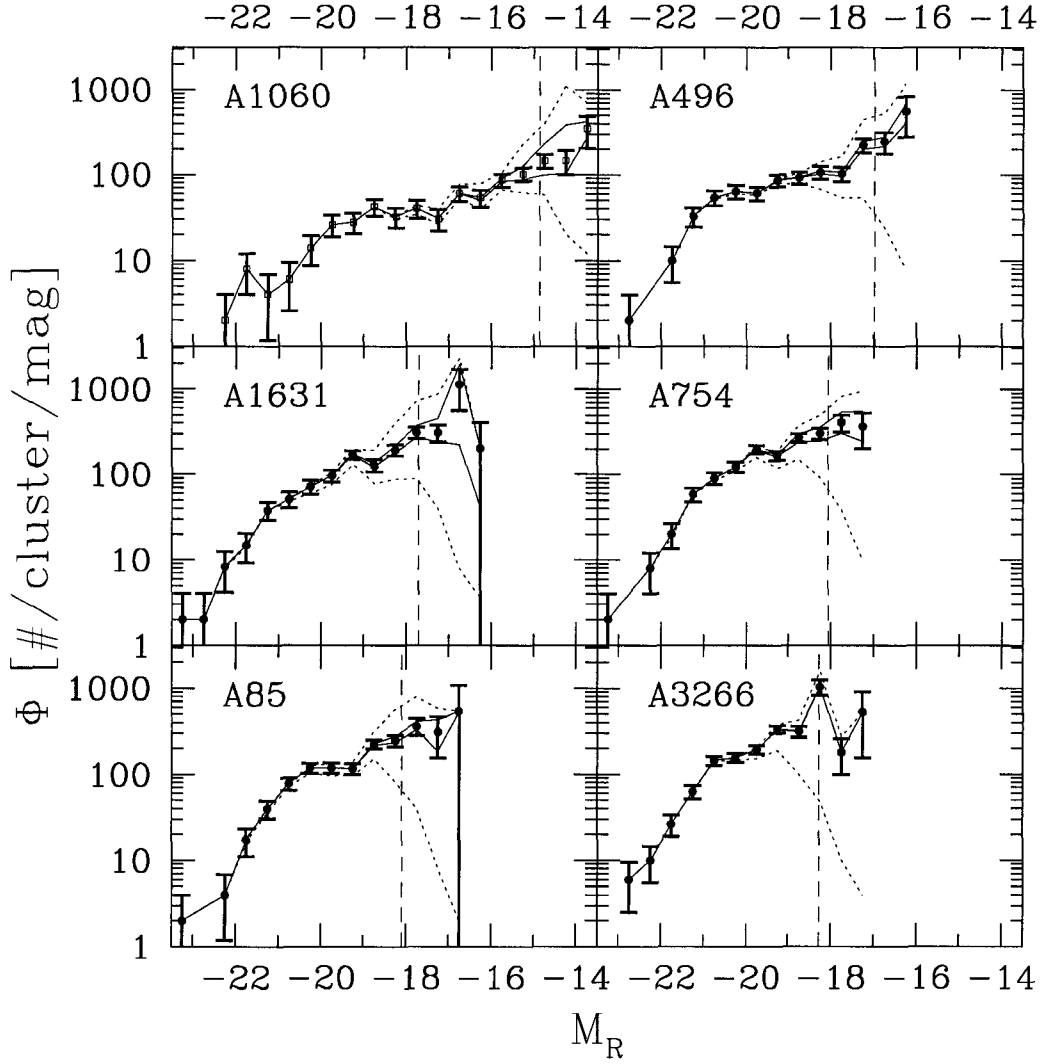


Figure 2.3: Cluster GLFs for all spectral types. Clusters are arranged left to right and top to bottom in order of increasing redshift. Error bars are  $1\sigma$ . Solid lines give upper and lower limits on the GLF, assuming that all spectroscopic targets for which we could not obtain redshifts are either members (upper limit) or non-members (lower limit). Dashed lines give number of spectroscopically confirmed sample members (lower limit) and total number of detections that are not confirmed non-members (upper limit). For orientation, vertical lines indicate the absolute magnitude corresponding to our standard analysis threshold of  $m_R = 18$ .

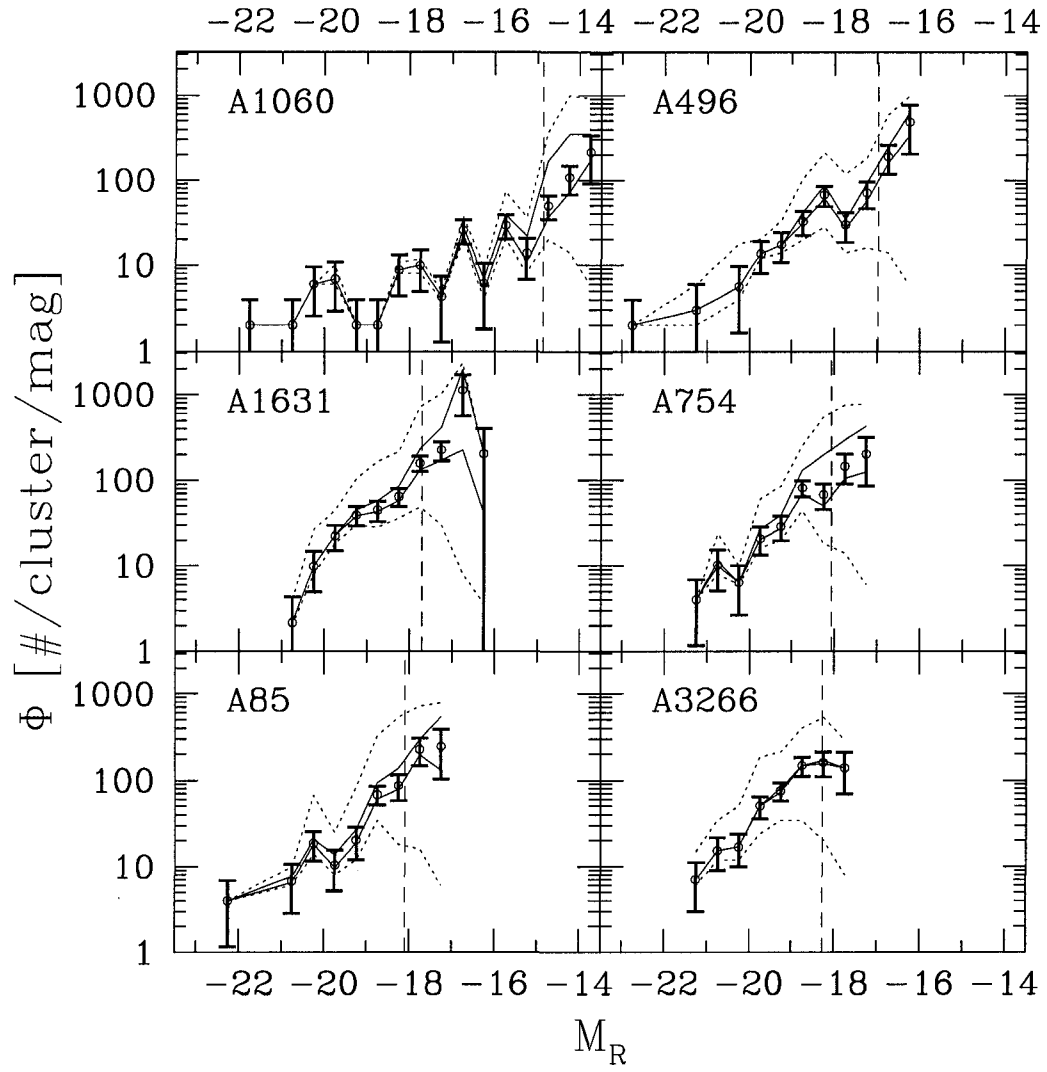


Figure 2.4: Same as Fig. 3 for EL galaxies only.

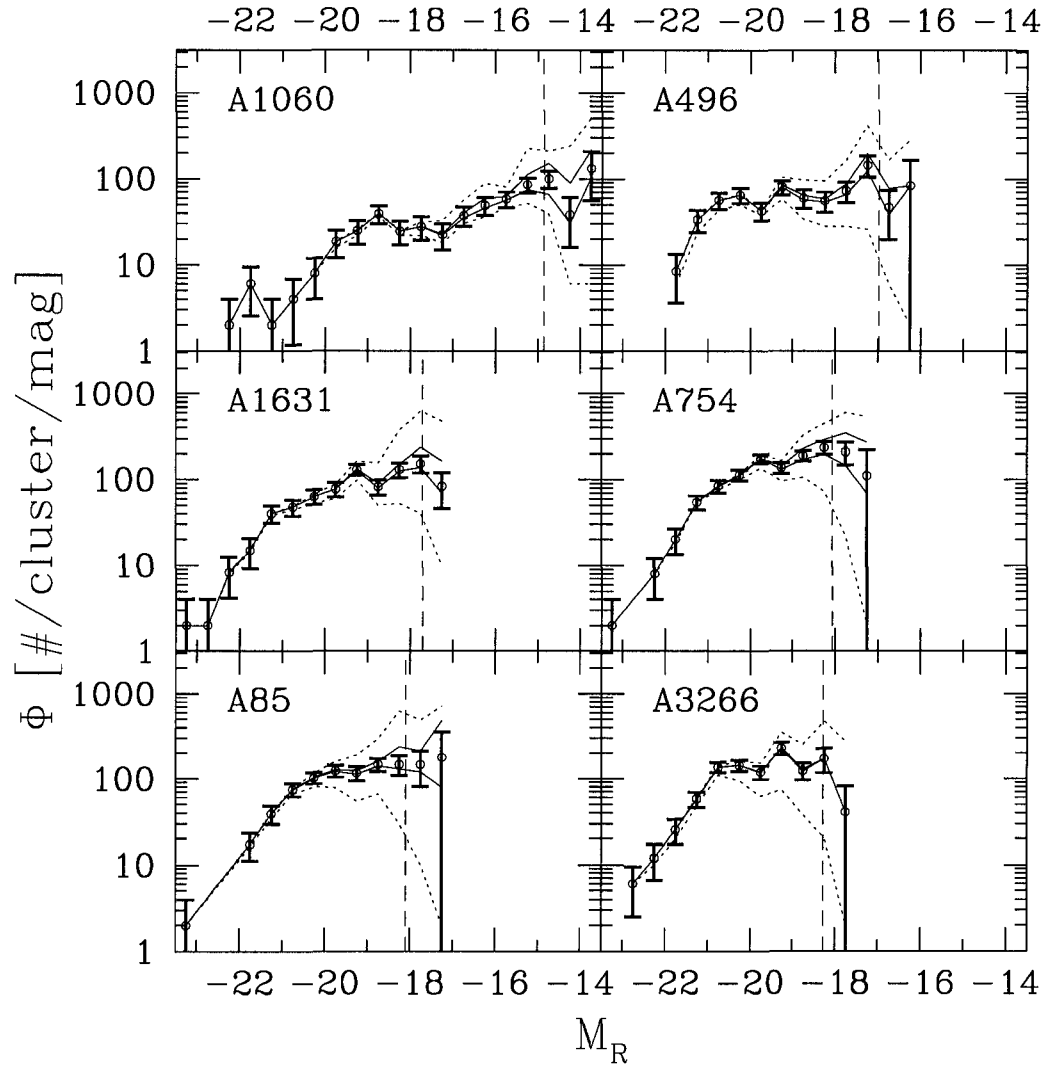


Figure 2.5: Same as Fig. 3 for NEL galaxies only.

tween clusters, we would expect about one in 20 comparisons to be inconsistent at the  $2\sigma$  level if the variance among clusters obeys Gaussian statistics.

Among the overall GLFs, only those of A1060 and A496 ( $M_R \leq -17.5$ ) are inconsistent. There are no significant inconsistencies between the EL GLFs. In the NEL sample, we find inconsistencies between A1060 and A496 ( $M_R \leq -17.5$ ), A496 and A754 NEL ( $M_R \leq -18.5$ ), and A754 and A3266 NEL ( $M_R \leq -18.5$ ).

To perform the same analysis over a common absolute magnitude range, we adopt a absolute magnitude limit of  $M_R = -18.35$  (effectively  $-18.5$  due to the bin boundaries), approximately the limit that corresponds to  $m_R = 18$  for the most distant cluster. All individual clusters are well sampled to and beyond this magnitude limit. While the results of some of the individual cluster-to-cluster comparisons differ from above, the general picture presented by this test is the same as above.

The results for the NEL GLFs suggest systematic differences among the six clusters. With a Monte Carlo test, we have verified that the probabilities are lower than expected if the six clusters had been drawn from a common parent distribution. We find no evidence for systematic differences of the overall or EL GLFs among our six clusters.

We have shown the non-universality of the NEL GLF, but have not yet constrained the universality of the GLF in clusters quantitatively. Therefore, we now turn to the second test, the comparison between individual clusters and composite GLFs, to examine to what extent these composites – overall, EL, and NEL – can serve as common parent distributions to the six clusters. To ensure that the distribution are statistically independent, we compare each individual cluster to a composite calculated from the five remaining clusters. We calculate the composites as described in §3.4. We convolve each composite GLF (in its original

form as a bivariate, luminosity-surface brightness distribution) with the Galactic extinction and sampling fraction applicable for the individual cluster. Thus, we predict how many galaxies should have been sampled in that cluster as a function of absolute magnitude if they had been drawn from the assumed parent distribution. This process is essentially the reverse of the SWML algorithm that we use to calculate the composite GLFs. We then project the predicted distribution onto the  $M_R$  axis and use a KS test and a  $\chi^2$  test to compare the distributions. In our case, the KS test is more sensitive to systematic discrepancies between the distributions, and therefore we base our discussion on its results. Again, we impose our common absolute magnitude limit of  $M_R = -18.35$  on all cluster-vs.-composite comparisons; this is the absolute magnitude corresponding to our standard apparent magnitude cutoff of  $m_R = 18$  for the most distant cluster, A3266.

Most clusters are consistent with the composites formed from the five remaining clusters. The exceptions are the EL and NEL populations of A1060. The KS-test probabilities are  $p = 0.04$  with 52 galaxies for the NEL GLF, and  $p = 0.003$  with 10 galaxies for the EL GLF. As this cluster is the lowest redshift cluster in our sample, this result raises the question of whether the fixed angular sampling radius of the fiber spectrograph has introduced an inhomogeneity into the sample by truncating the clusters at different physical radii.

#### 2.5.1.2 Radial Sampling Bias

To address the possibility of radial sampling bias, we determine new composite GLFs, for which we truncate the higher-redshift clusters to the same fraction of the virial radius sampled by the lower- $z$  clusters. In virialized systems, the virial radius scales with  $\sigma$  (Girardi et al., 1998). We therefore scale the angular sampling radii of the low-redshift clusters by  $\sigma/D_A$ , where  $D_A$  is the angular diameter distance, to find the correct angular truncation radius for the other clus-

ters. The composite GLFs truncated to the A1060 and A496 sampling radii are indistinguishable from the default composite GLFs under a  $\chi^2$  test. (This is not surprising for two reasons: Not only are the samples dominated by galaxies in the central regions of the clusters, but the scaling of the virial radius with  $\sigma$  and the angular scale with  $D_A$ , by coincidence, mostly cancel each other, making the default angular sampling radii of most clusters comparable.)

We now repeat the cluster-vs.-composite analysis from §4.1.1 with these truncated GLFs to find out whether the discrepancies between the composites and the EL and NEL populations in A1060 persist. The A1060 EL and A496 NEL GLFs are now inconsistent with the composites, but the A1060 NEL GLF is consistent. Therefore, we cannot rule out that the small sampling radius is at least partly responsible for the inconsistency in the A1060 NEL GLF that we observed with the untruncated samples. On average, however, we still find five out of six clusters to be consistent within  $2\sigma$  with the composite for the EL and NEL samples, and six out of six for the overall samples. We will discuss below what constraints this places on the universality of the shape of the GLF in clusters.

### 2.5.1.3 Aperture Bias

Another potential source of inhomogeneity in the sample is aperture bias. Due to the finite angular radius of the  $3''$  fibers, the spectra only sample light within a limited physical radius around the center of a given galaxy. The spectrum may thus not be representative of the galaxy as a whole, and, in particular, the [OII] equivalent width may be biased low. This effect is obviously of greater concern at lower redshift and for galaxies with large bulges. Therefore, we have to consider whether aperture bias may have enhanced the bright end of the NEL GLF in the most nearby clusters by causing misclassifications of EL galaxies as NEL galaxies.

We have found no strong indications that aperture bias is responsible for the

discrepancies observed in the A1060 EL and NEL and A496 NEL GLFs. Although the deviation responsible for the disagreement of A496 with the composite is indeed an excess of galaxies at the bright end and a deficit around  $M_R \approx -18.5$  (compared to a composite normalized to minimize  $\chi^2$ ), the same systematic deviation is qualitatively observed in the A496 overall GLF, which cannot be affected by aperture bias. In addition, the discrepancy between A1060 and the composite is not in the same sense as that between A496 and the truncated composite, even though we would expect a stronger effect for this, more nearby, cluster.

We examine how many galaxies would need to be misclassified in A496 to explain its deviant GLF by aperture bias. Because of the larger physical size of the most luminous galaxies, aperture bias would affect the bright end of the NEL GLF most for a given cluster. The observed deviation would require  $\sim 40\%$  of the presumed NEL galaxies around  $M_R \approx -20.5$  to be EL galaxies. Given the small number of identified EL galaxies, this would require  $\sim 85\%$  of the EL population to have been misclassified as NEL galaxies. For a sample very similar to this, Zaritsky, Zabludoff & Willick (1995) estimate that at most 20% of spiral galaxies with  $cz \leq 15000$  km/s might be affected by aperture bias resulting in a misclassification of their emission line properties. Therefore, judging primarily from the case of A496, aperture bias is unlikely to be the cause of the observed discrepancies between individual clusters and the composites, or between individual cluster NEL GLFs. Aperture bias cannot explain the deviation of the A1060 NEL GLF from the composites, as the composite overpredicts, rather than underpredicts, the number of bright NEL galaxies, and underpredicts the number of faint galaxies.



#### 2.5.1.4 The Uniformity of Cluster GLFs

We have demonstrated in §2.5.1.1 that  $\chi^2$  comparisons among individual clusters rule out a universal shape of the NEL GLF. The comparisons among individual clusters and the composite GLFs in the previous sections yield additional constraints on the degree of uniformity of GLFs in clusters.

For our comparisons between individual clusters and the composite GLFs formed from the five remaining clusters, we find all six clusters to be consistent with the composite within  $2\sigma$  in the case of the overall GLFs, and five out of six consistent within  $2\sigma$  when comparing EL or NEL populations. This allows us to place constraints on the “universality” of our composite GLFs, *i.e.*, on the fraction of all clusters (with selection criteria similar to those in our sample) that are consistent with our composite GLFs within  $2\sigma$ . To establish the universality of our composite GLF, that fraction would have to be shown to be at least 0.95 (*i.e.*, on average only one cluster out of 20 should show a discrepancy at the  $2\sigma$  level), which is impossible to prove with a sample of just six clusters. However, simple binomial statistics show that it is unlikely ( $p < 0.05$ ) to draw five consistent clusters out of a sample of six unless at least  $\sim 40\%$  of all clusters are in agreement with our composite GLFs. We therefore conclude that our EL and NEL composite GLFs are representative of at least  $\sim 40\%$  ( $2\sigma$  lower bound) of all clusters. Drawing six consistent clusters out of a sample of six, as in the case of the overall GLF, is unlikely unless at least  $\sim 60\%$  of all clusters are consistent with the composite.

We therefore adopt this fraction of  $\sim 60\%$  for the overall GLF and  $\sim 40\%$  for the EL/NEL GLFs as lower bounds on the fraction of all clusters for which our composites are representative. While this argument does not prove the existence of a universal GLF, it indicates a degree of uniformity that is interesting given the relatively wide range of velocity dispersions spanned by these six clusters.

Larger cluster samples, analyzed in a similar way, would provide tighter bounds on the universality of our composite GLF, provided that each cluster is sampled deeply enough to provide significant constraints on its consistency with the composite. It is unlikely, however, that the limit would be revised downward from  $\sim 60\%$ . The sampling in each of our six clusters is already almost complete over the magnitude range considered here. Furthermore, the uncertainties associated with the composite GLF, which are already much smaller than those associated with the individual clusters, are not propagated through the KS test, so that the lower bound that we derive from this test is a conservative one.

It is possible that differences among cluster GLFs exist at the very faint end, which has been sampled by us only in the nearest clusters. Given that the comparisons between NEL GLFs yield the strongest indications for non-random discrepancies (stronger than for the better-sampled overall GLFs), searches for systematic differences among clusters are more likely to succeed if NEL galaxies are considered separately.

### 2.5.2 Cluster Composite GLF

The right hand panels in Fig. 2.6 show the composite cluster GLFs for the complete (top row), EL (middle row) and NEL (bottom row) samples, based on all six cluster fields. The numerical values of these GLFs are given in Tab. 2.3.

We fit Schechter functions to these GLFs down to a limiting magnitude of  $-17$  (the magnitude at which the composite GLF is based on more than one cluster). The best fit Schechter parameters are given in the individual panels of Fig. 2.6 and are also listed in Tab. 2.4. All fits are consistent with the discrete GLFs within  $2\sigma$  over the specified magnitude range.

Fig. 2.7 shows the 1- and  $2\sigma$  error contours for the fits, as determined from the method of Avni (1976), with the  $\Delta\chi^2$  values taken to be the  $\chi^2$  values for two

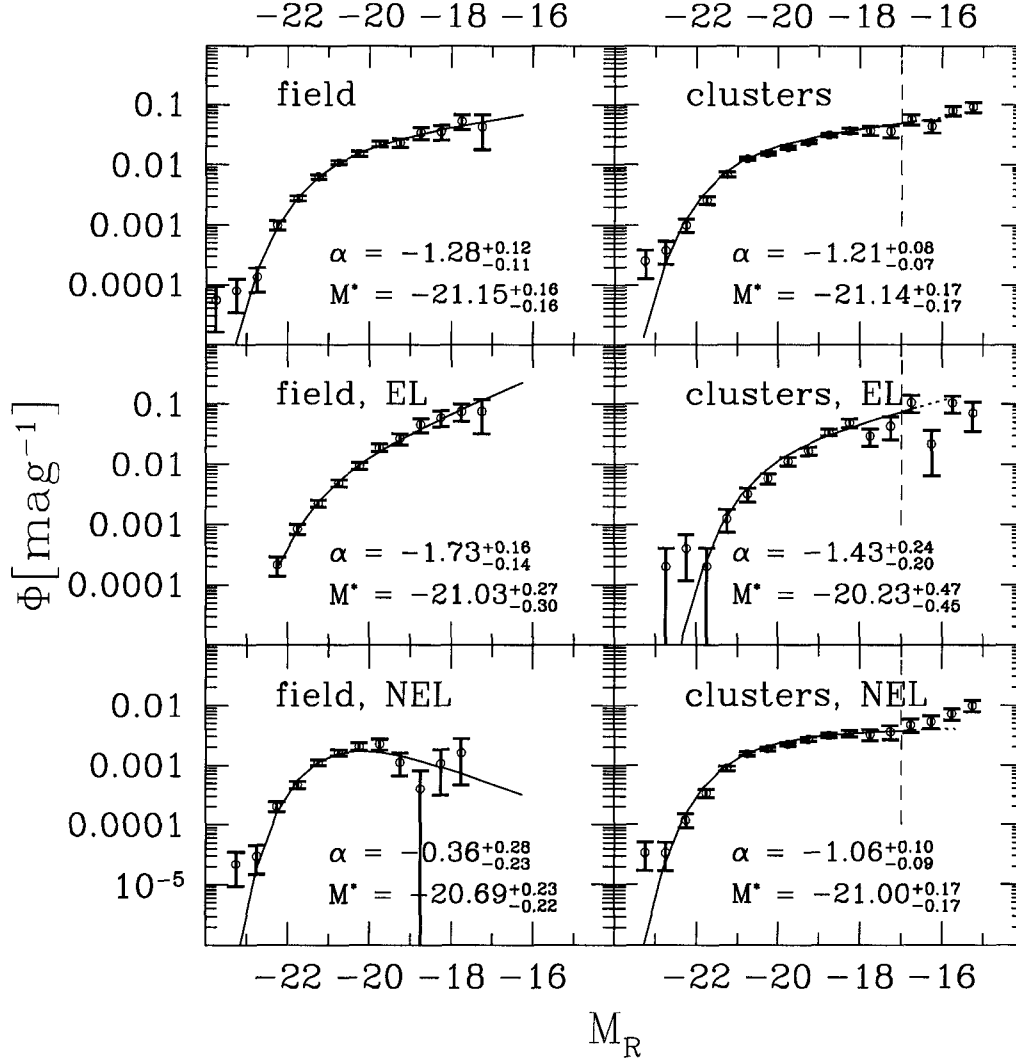


Figure 2.6: Composite GLFs for field and cluster subsamples and different spectral types. Schechter fits down to a magnitude limit of  $M_R = -17$  mag are also shown. Normalization is arbitrary. Vertical lines indicate limit beyond which only A1060 contributes ( $M_R \approx -17$  for our adopted apparent magnitude cutoff of  $m_R = 18$ ).

Table 2.3. Composite GLFs. GLF values given in decadic logarithms; number of galaxies given in parentheses.

$M_R$	field, all	field, EL	field, NEL	clusters, all	clusters, EL	clusters, NEL
-23.75	-4.256(2)	...	...	...	...	...
-23.25	-4.100(3)	...	-4.658(3)	-3.595(4)	...	-3.467(4)
-22.75	-3.872(5)	...	-4.525(4)	-3.418(6)	-3.698(1)	-3.469(4)
-22.25	-3.000(36)	-3.668(8)	-3.692(26)	-2.991(16)	-3.395(2)	-2.920(14)
-21.75	-2.555(91)	-3.071(28)	-3.325(54)	-2.582(41)	-3.693(1)	-2.472(39)
-21.25	-2.198(168)	-2.649(55)	-2.957(100)	-2.154(110)	-2.905(6)	-2.056(100)
-20.75	-1.964(158)	-2.311(62)	-2.789(77)	-1.898(198)	-2.495(15)	-1.811(172)
-20.25	-1.820(120)	-2.030(57)	-2.686(52)	-1.805(238)	-2.234(25)	-1.727(192)
-19.75	-1.653(82)	-1.724(48)	-2.648(24)	-1.704(279)	-1.952(42)	-1.652(195)
-19.25	-1.631(33)	-1.572(22)	-2.950(6)	-1.621(230)	-1.781(41)	-1.568(166)
-18.75	-1.468(20)	-1.338(15)	-3.396(1)	-1.501(243)	-1.468(64)	-1.509(146)
-18.25	-1.449(13)	-1.226(11)	-2.970(2)	-1.431(118)	-1.310(41)	-1.477(66)
-17.75	-1.278(13)	-1.124(10)	-2.795(2)	-1.426(41)	-1.527(10)	-1.492(24)
-17.25	-1.373(3)	-1.120(3)	...	-1.433(21)	-1.358(6)	-1.443(14)
-16.75	...	...	...	-1.242(26)	-0.975(10)	-1.335(15)
-16.25	...	...	...	-1.355(19)	-1.663(2)	-1.275(17)
-15.75	...	...	...	-1.100(33)	-0.979(10)	-1.145(23)
-15.25	...	...	...	-1.032(26)	-1.143(4)	-1.005(22)

Table 2.4. Schechter parameters

Sample	$\alpha$	$M_R^*$	$p_{chi^2}$
field, all	$-1.28^{+0.12}_{-0.11}$	$-21.15^{+0.16}_{-0.16}$	0.84
field, EL	$-1.73^{+0.16}_{-0.14}$	$-21.03^{+0.27}_{-0.30}$	0.94
field, NEL	$-0.36^{+0.28}_{-0.23}$	$-20.69^{+0.23}_{-0.21}$	0.10
clusters, all	$-1.21^{+0.08}_{-0.07}$	$-21.14^{+0.17}_{-0.17}$	0.19
clusters, EL	$-1.43^{+0.24}_{-0.20}$	$-20.23^{+0.47}_{-0.45}$	0.10
clusters, NEL	$-1.06^{+0.10}_{-0.09}$	$-21.00^{+0.17}_{-0.17}$	0.56

degrees of freedom ( $\alpha$  and  $M^*$ ) and probabilities of 0.315 and 0.05.

The failed spectroscopic targets introduce an additional uncertainty in the faint end slope of  $\Delta\alpha \approx \pm 0.03$  and  $\Delta M^* \approx \pm 0.01$ . The effect of selecting galaxies in a different magnitude band is on the order of  $\Delta\alpha \approx +0.02$ . The effect of higher-order variations of the sampling fraction is on the order  $\Delta\alpha \approx +0.01$ . As these corrections are all distinctly smaller than our statistical uncertainties and the sense of the largest of these is unclear, we neglect them.

We have tested whether these composite GLFs represent all of our individual cluster GLFs well by comparing them using the same procedure as in §4.1, *i.e.*, convolving them with the sampling function and Galactic extinction for any given cluster, and comparing the predicted distribution of sample galaxies to the observed distribution using a KS and  $\chi^2$  test. As our purpose is only to verify that our composites represent all of the individual cluster GLFs, we carry the comparisons as far as justifiable in each case, *i.e.*, to a magnitude limit of  $m_R = 18$  (but not fainter than  $M_R = -17$ , beyond which only A1060 contributes to the composite).

We find that the overall, EL and NEL composites represent the respective galaxy distributions in all six individual clusters well (*i.e.*, no comparison — either by a  $\chi^2$  or KS test — shows a discrepancy at the level of  $2\sigma$  or more). Therefore, we conclude that our composite GLFs are good representations of all six clusters in our sample and of  $\geq 60\%$  (as estimated in §4.1.4;  $\geq 40\%$  for the EL and NEL GLFs) of all clusters obeying similar selection criteria.

### 2.5.3 Field GLFs

The sample contains 1527 galaxies not associated with the six clusters, 749 of them within our redshift and magnitude limits. This enables us to calculate GLFs for field galaxies in the same way as our cluster composite GLFs, and to compare them in a self-consistent way.

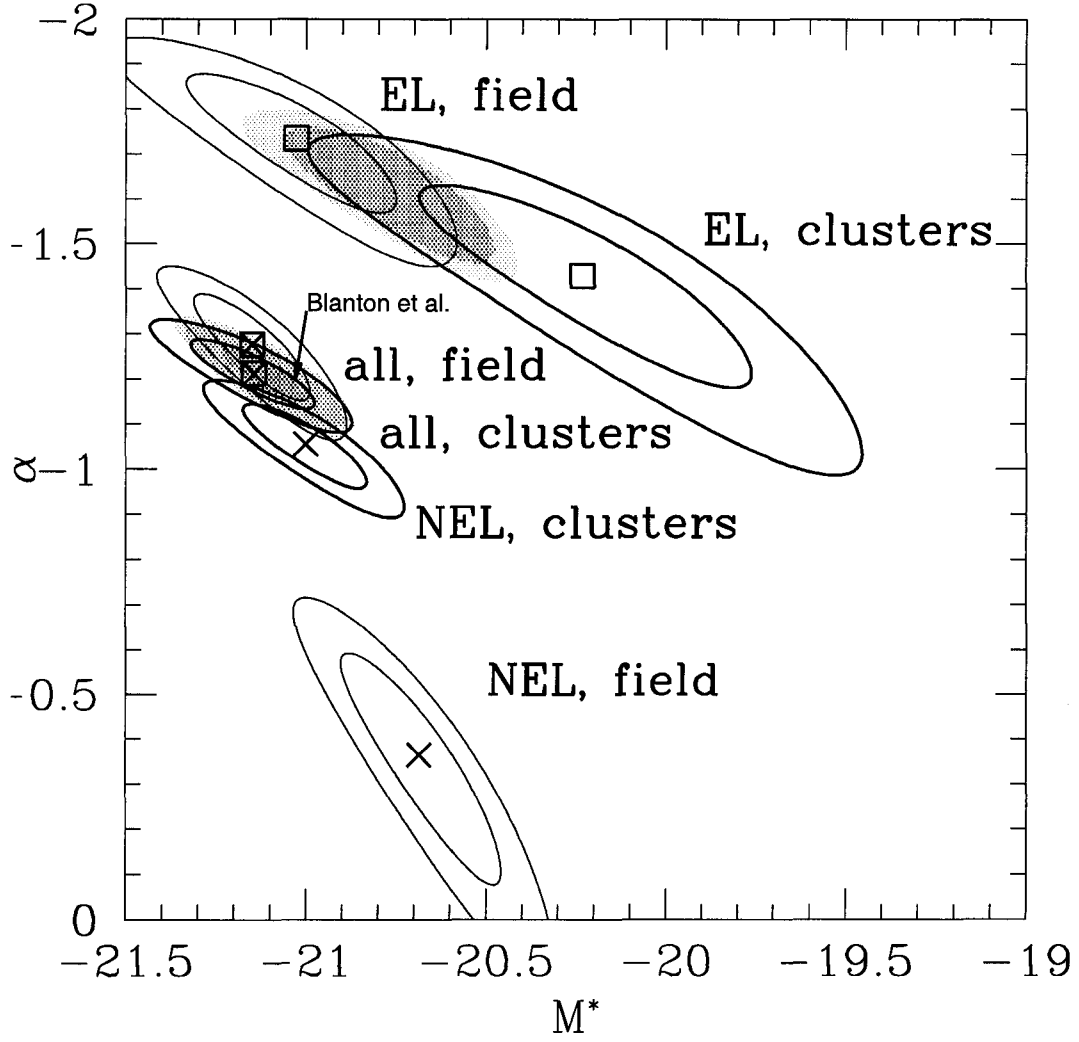


Figure 2.7:  $1\sigma$  and  $2\sigma$  error contours in Schechter parameter space for our Schechter fits to the combined field and cluster GLFs down to  $M_R = -17$  mag. The arrow shows the field GLF from Blanton et al. (2001) with a transformation  $M_R - M_{r^*} = -0.2$ .

We apply the SWML method with an apparent magnitude cutoff of  $m_R = 18$  ( $m_R = 17$  for the A3266 field) to calculate the field GLF. We impose redshift limits of  $6000 < cz < 50000$  km/s. The lower limit is motivated by the fact that there is a low- $z$  ( $cz \approx 4000$  km/s) feature in the field of A1631 (apparently associated with NGC4756) that would dominate the faint end of the field GLF if included. The distribution of these galaxies in a plot of [OII] EW versus  $M_R$  is uncharacteristic of field galaxies in the five other fields and at higher redshifts; the high fraction of galaxies with very low [OII] EW resembles a cluster population rather than a field population. While the field sample is supposed to cover a representative range of environments, including higher-density ones, the volume of space sampled at such low redshifts is not large enough to guarantee a representative sample of the field population if this feature is included.

The composite field GLFs are presented in the left hand column of Fig. 2.6 and in Tab. 2.4. As for the cluster composite GLFs, we fit Schechter functions for  $M_R \leq -17$ . The 1- and  $2\sigma$  error contours are given in Fig. 2.7.

Our results for the field GLF are in agreement with the values given by Blanton et al. (2001) for the Digital Sky Survey (SDSS) of  $\alpha = -1.20$  and  $M_{r^*} = -20.83$  (assuming  $M_{r^*} - M_R \approx 0.2$ , based on Fukugita et al. (1996)). If we adopt a reasonable color of  $(B - R) \approx 1.0$ , our field GLF is also in agreement with the determination of Madgwick et al. (2001) from the 2dFGRS.

As in the case of the cluster GLFs, the various biases afflicting the calculation of the sampling fraction do not affect the composite field GLF significantly. The failed spectroscopic targets introduce worst-case uncertainties of  $\pm 0.03$  in  $\alpha$  and  $M^*$  in the field GLF. Systematic effects in  $\alpha$  from the color selection bias and higher order variations in  $m_R$  are on the order of  $\Delta\alpha \leq 0.01$  each.

### 2.5.4 Comparisons between Field and Clusters

We use three different methods to compare our composite cluster and field GLFs to each other. The first is a simple  $\chi^2$  comparison of the SWML solutions for the overall, EL and NEL samples. We compare them to a limiting absolute magnitude of  $M_R = -17$ , the magnitude to which the cluster GLF is determined from more than one cluster and to which the field GLFs contain data. Tab. 2.5 summarizes the results of these comparisons. The overall and EL GLFs are consistent between the field and clusters, but the field NEL GLF is clearly distinct from the cluster NEL GLF at more than  $3\sigma$ .

To test for systematic differences in the shape of the GLF to which the  $\chi^2$  test may be insensitive, we also compare the composite GLFs in terms of their Schechter fits. Two GLFs are obviously inconsistent if the  $2\sigma$  error contours of their Schechter fits are disjoint. If that is not the case, a more differentiated evaluation is required. If the error contours overlap, any set of parameters within this overlap region would be individually consistent with both GLFs. However, the requirement for the correct Schechter values of both GLFs to lie simultaneously within this overlap region imposes an additional constraint on the likelihood of such a fit; even a point that is marginally consistent with both GLFs individually may thus not necessarily qualify as a likely simultaneous fit to both.

To avoid the difficulties of having to calculate the exact probability density for both fits in Schechter space, we approximate this probability by calculating the joint  $\chi^2$  probability of two different GLFs having been drawn from the same Schechter function. The joint probability is simply the product of the two individual probabilities, renormalized so that the best fit probability for each individual fit is 1. We apply this renormalization because, for the purposes of this comparison, we are not interested in the quality of the fits, but merely in using them as



Table 2.5. Field vs. Cluster comparisons. Probabilities of same parent distributions under  $\chi^2$  and parametric comparisons.

Comparison	$\chi^2$ -test	parametric
Field vs. Clusters	0.77	0.83
Field EL vs. Clusters EL	0.87	0.31
Field NEL vs. Clusters NEL	<0.0005	<0.0005

tools to characterize the shapes of the GLFs. (This idea is similar to the procedure for determining the error contours, which are always drawn relative to the minimum  $\chi^2$ , regardless of the quality of the best fit itself.) If a point exists in  $(\alpha; M_R^*)$  space for which the joint probability is high enough so as not to exclude a simultaneous fit (*i.e.*,  $>0.05$ ), and if that point lies within the error contours of both individual Schechter fits, we consider the two GLFs consistent. Deriving the joint probability by multiplying the two individual probabilities implies that the two realizations of the GLF are independent; therefore, we cannot use this procedure to compare, for example, the NEL composite to the overall composite GLF.

The shaded regions in Fig. 2.7 indicate regions of the  $(\alpha; M^*)$  plane where the Schechter function is in simultaneous agreement with two different GLFs, and Tab. 2.5 lists the results of this test. The parametric comparisons confirm the result from the  $\chi^2$  test; the field and cluster overall composite GLFs are consistent with each other, as are the EL composite GLFs. *However, the NEL GLF differs between the field and clusters under both tests.* Inspection of Fig. 2.7 reveals that it is steeper in clusters than in the field.

To confirm this important conclusion, and maintain consistency with the pro-

cedure used earlier to compare individual clusters to the composite GLFs, we apply a third test: can the field GLFs serve as parent to the six individual clusters? The procedure is the same as for the cluster-composite comparisons in §4.1, except that we now adopt the field GLFs (EL, NEL and overall) as the hypothetical parent distributions for the six cluster samples. As we did for the comparison between the cluster composite GLF and the individual clusters, we impose a magnitude limit of  $M_R = -17$  and  $m_R = 18$  (except for the NEL GLF, where we do not have any data for  $M_R \approx -17.25$  and therefore restrict the comparison to  $M_R \leq -17.5$ ). Table 2.6 shows the results of this test.

The overall field composite GLF is consistent with the cluster populations in all six cases, confirming our earlier finding that the field and cluster GLFs are consistent in shape. The distribution of probabilities is consistent with a normal distribution at  $p = 0.74$ . This conclusion also holds for the EL GLFs ( $p = 0.1$ ). However, the field NEL GLF is clearly inconsistent with the NEL populations in all six clusters.

Therefore, our conclusion, supported by all three tests, is that the NEL GLF differs between the field and clusters and is steeper in high-density environments. We cannot detect a systematic discrepancy between the field and cluster EL GLFs (but caution that the constraint is weaker because of the smaller number of EL galaxies).

Before we consider possible implications of this, we have to consider the potential role of aperture bias in introducing an inhomogeneity between the field and cluster NEL samples. The cluster sample, on average, is at lower redshift than the field sample, and is thus more susceptible to aperture bias, which might artificially increase the number of galaxies classified as NEL and raise the faint end slope of the NEL GLF. However, if we limit the field NEL GLF to  $cz \leq 20000$

Table 2.6. Comparisons between individual clusters and field GLF,  $m_R \leq 18$ 

Cluster	$p_{KS}$	$p_{\chi^2}$	# of galaxies
overall GLF			
A1060	0.11	0.82	101
A496	0.08	0.51	237
A1631	0.60	0.62	272
A754	0.81	0.94	367
A85	0.28	0.93	301
A3266	0.41	0.53	391
EL GLF			
A1060	0.06	0.71	20
A496	0.88	0.94	49
A1631	0.20	0.79	61
A754	0.11	0.78	52
A85	0.14	0.25	43
A3266	0.56	0.96	62
NEL GLF			
A1060	<b>0.000</b>	<b>0.000</b>	79
A496	<b>0.000</b>	<b>0.026</b>	159
A1631	<b>0.000</b>	<b>0.014</b>	210
A754	<b>0.000</b>	<b>0.000</b>	310
A85	<b>0.000</b>	<b>0.001</b>	201
A3266	<b>0.000</b>	<b>0.047</b>	232

km/s, the faint end slope does not steepen. If we impose the same lower redshift limit of  $cz \geq 6000$  km/s on the cluster NEL GLF as on the field GLF,  $\alpha$  does not become significantly shallower. Furthermore, we can perform a simple plausibility check: The difference between the field and cluster NEL GLF within the magnitude range  $-20 < M_R < -18$  amounts to  $> 60\%$  of the integral of the cluster NEL GLF in this range. This ratio implies that over 350 cluster EL galaxies would have to have been misclassified as NEL galaxies to explain the discrepancy. The error rate in identifying EL galaxies would then have to be  $> 90\%$ . For the reasons given in §4.1.3, this large an error is unlikely.

The sense of the variation of the NEL GLF is a steepening of the faint end slope from  $\alpha = -0.36$  in the field to  $\alpha = -1.06$  in clusters. This observation agrees with Christlein (2000), who also found a steepening of the NEL GLF faint end slope (measured to  $M_R = -17.5$ ) in increasingly denser environments. In contrast to Christlein (2000), our overall GLF is not significantly different between low- and high-density environments<sup>2</sup>. It is interesting that the overall GLF shows no significant difference between the field and clusters, despite the differences between the NEL GLFs and the well established morphology-density relation (Dressler, 1980). In our data, the effects of the morphology-density relation, in which early-type galaxies with their intrinsically shallower GLF are more abundant in denser environments, and the steepening of the faint end slope of the NEL galaxies (which are mostly early-type galaxies) cancel each other within our margins of uncertainty.

Given that the overall GLFs in the field and clusters are similar, the difference between the galaxy populations of these two environments is revealed by the dif-

---

<sup>2</sup>This discrepancy may result from the different survey parameters of the Las Campanas Redshift Survey, particularly its surface brightness cutoff, which has been demonstrated by Blanton et al. (2001) to produce an inaccurate field GLF.

ference between the fraction of EL (or, equally, NEL) galaxies with environment (Fig. 2.8). To construct this figure, we normalize our GLFs to appropriate units (galaxies per comoving  $\text{Mpc}^3$  for the field; galaxies per cluster for the clusters) and calculate the fraction of EL galaxies as a function of  $M_R$ . We display this fraction both as the ratio of the EL GLF to the sum of the EL and NEL GLF, and as the ratio of the EL GLF to the overall GLF. Because the EL, NEL and overall GLFs are calculated and normalized independently of each other, the overall GLF is not necessarily exactly identical to the sum of the EL and NEL GLF, but the two methods agree well. We calculate the error bars using a simple Monte Carlo algorithm. We take the observed EL fraction in each bin to be the parent distribution. For 1000 trials, we draw samples of galaxies equal to the total number of observed galaxies in that bin and determine the scatter in the distribution of the simulated EL fractions recovered from these trials. We then use this scatter as our error bars in each bin.

The upper panel of Fig. 2.8 shows the EL fraction  $f_{EL}$  in the field and clusters. The solid lines show the results calculated from the parametric fits, circles give the NEL fraction as calculated from the sum of the EL and NEL GLF, and triangles give the ratio of the NEL GLF to the overall GLF. All three methods agree. As we move from the field to the clusters, the fraction of EL galaxies decreases significantly over almost all magnitudes (with the exception of the bright end, where the EL fraction is almost zero in both environments). The most drastic differences in the EL fraction are at fainter absolute magnitudes. Around  $M_R = -18$ , the field population is almost entirely dominated by EL galaxies, while in our cluster sample, the fraction of EL galaxies has dropped to about one third on average<sup>3</sup>.

---

<sup>3</sup>The absolute numbers, of course, are likely to be dependent on the sampling radius, but our aim is to demonstrate qualitative variations, and with the exception of A1060, the sampling in our clusters is rather homogeneous.

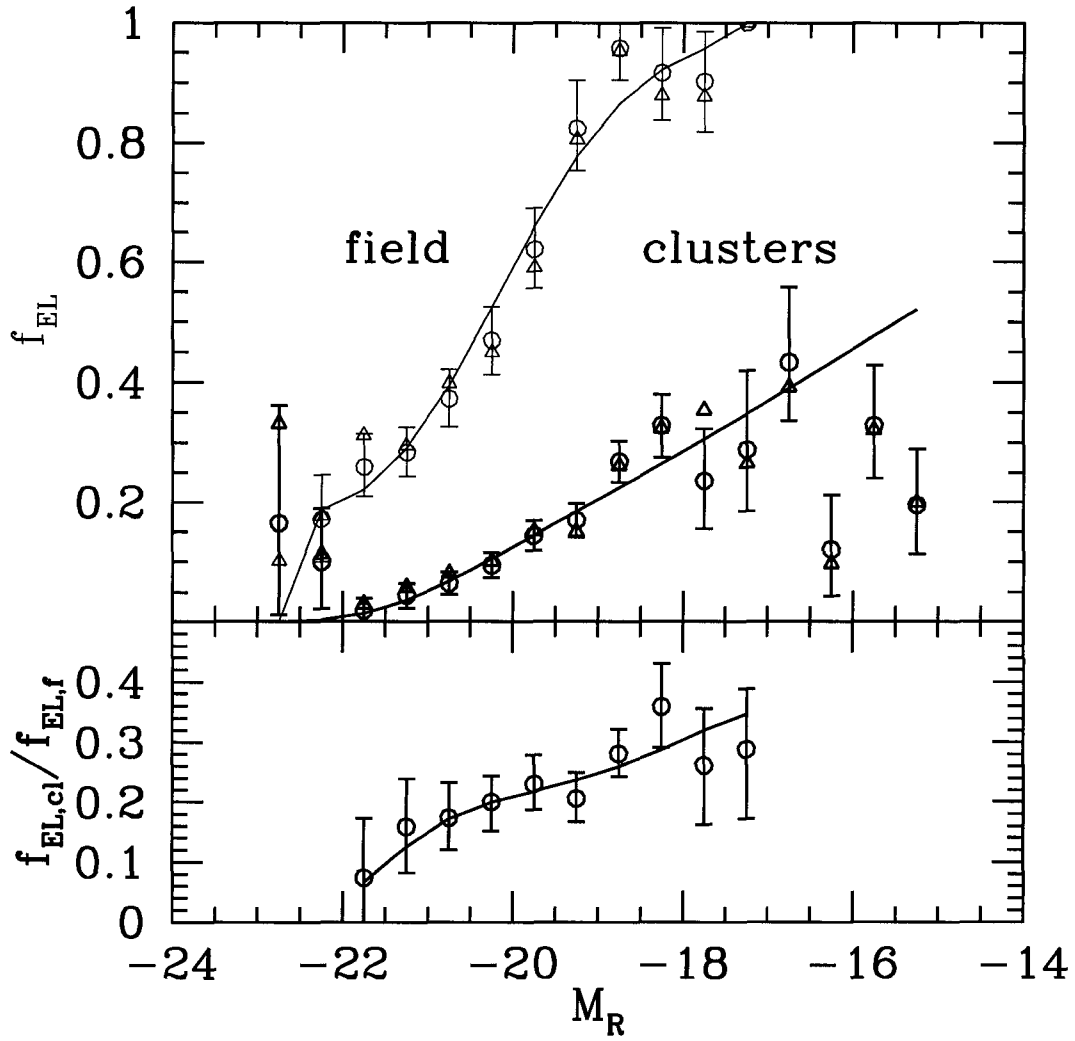


Figure 2.8: Upper Panel: Fraction of EL galaxies as a function of  $M_R$ , derived from the cluster and field composite GLFs. Data points quote fraction relative to sum of EL and NEL GLFs; triangles indicate fraction relative to total GLF. Solid lines are based on Schechter fits. Error bars indicate approximate  $1\sigma$  uncertainties, based on Monte Carlo estimate. Bottom panel: Ratio  $f_{EL}^{clusters} / f_{EL}^{field}$ .

While the change in the EL fraction,  $f_{EL}^{field} - f_{EL}^{clusters}$ , is larger for fainter magnitudes,  $f_{EL}$  itself is also larger at the faint end. It is thus more instructive to plot the ratio  $f_{EL}^{cluster}/f_{EL}^{field}$  versus  $M_R$  to learn whether the properties of dwarf or of giant galaxies are more strongly correlated with their environment. The bottom panel of Fig. 2.8 shows the ratio  $f_{EL}^{cluster}/f_{EL}^{field}$ .

The curve in Fig. 2.8 is below one for all magnitudes, indicating that cluster galaxies at all  $M_R$  are not as likely to be star forming as field galaxies. This is consistent with other recent studies of larger samples (Gómez et al., 2003; Lewis et al., 2002). The effect is stronger at the bright end in our data. To confirm this, we have applied a Spearman rank correlation test. The coefficient is  $r = 0.93$ , with a high level of significance ( $p < 0.0005$  for an accidental correlation). This rules out a zero slope. The best fit slope is positive, indicating that the star formation properties of giant galaxies ( $M_R \approx M_R^*$ ) vary more with the environment than those of fainter galaxies ( $M_R \geq M_R^*$ ).

de Propris et al. (2003), in their study of the  $b_J$ -band GLF in 60 clusters from the 2dF Galaxy Redshift Survey, obtain results that are complementary to ours. They find a small difference of  $\sim 0.1$  in the faint end slope  $\alpha$  between the field GLF by Madgwick et al. (2001) and their cluster GLF (which is steeper). Such a difference cannot be ruled out from our sample, but could also be due to the greater sensitivity of  $b_J$  magnitudes to dust and star formation. With  $\alpha_{b_J} \approx -1.28$ , their faint end slope for the cluster GLF is not significantly different from our  $\alpha = -1.21$ . A small difference in  $\alpha$  is expected if there is a color gradient with absolute magnitude ( $\alpha_R - \alpha_{b_J} \approx d(b_J - R)/dM_R$ ). If we assign colors of  $(B - R) = 1.0$  to EL galaxies and  $(B - R) = 1.5$  to NEL galaxies and use Fig. 2.8 to estimate the mixing ratio between these two populations at the faint end, the difference in the faint end slope between  $b_J$ -band and R-band GLFs should be  $\Delta\alpha \approx 0.05$ ,

the blue GLF being steeper than the red GLF. This is entirely consistent with the comparison between these two studies. De Propris et al., comparing their results to the field GLF of Madgwick et al. (2001) also confirm that variations of the GLF in subsamples selected by spectral types are stronger than in the overall GLF, with early spectral types showing a significant steepening of the faint end slope when going from the field to clusters.

### 2.5.5 The Bright End

The bright end of the GLF in clusters has been the subject of many studies. Interest has focused on the brightest cluster galaxies, often suggested as a special population (Nelson et al., 2002) of standard candles (Sandage, 1972; Postman & Lauer, 1995) inconsistent with the Schechter function fit to fainter galaxies (Tremaine & Richstone, 1977; Dressler, 1978). A related question is whether these galaxies exist in environments other than the densest clusters (Morgan, Kayser & White, 1975).

Our cluster overall and NEL GLFs exhibit an apparent excess of galaxies over the Schechter function in the brightest bins ( $M_R \leq -22.5$ ). The field GLF also exhibits what may be a bright end excess, but at a lower level. If these galaxies are unique to rich clusters, then there may be a significant difference between the field and cluster bright ends that our previous tests have not been sensitive enough to detect.

To address this question, we normalize the cluster GLF to predict the correct total number of sampled field galaxies with  $M_R \leq -20$  when integrated over redshift (excluding the redshift ranges associated with the clusters),  $M_R$ ,  $\mu_R$ , and the area of the survey. At this bright magnitude limit, any faint end slope differences not affect our normalization. We then use the normalization factor to determine how many field galaxies the cluster GLF would predict in the bright



tip with  $M_R \leq -22.5$ . The answer is  $\sim 22$  galaxies. The observed number of bright field galaxies is 10. The reverse test shows that, when normalized to the number of galaxies in the cluster sample, the field GLF predicts the number of bright galaxies in the clusters to be  $\sim 4$ , whereas the actual number is 10.

We estimate the significance of this discrepancy in the following way: we assume a certain fraction  $\int \phi(M < -22.5)dM / \int \phi(M < -20)dM$  of galaxies in the parent distribution to reside in the bright tip (essentially, we assume a GLF with two bins). Because the number of galaxies with  $M_R < -22.5$  is negligible compared to the number of galaxies with  $M_R < -20$ , this fraction scales approximately as the fraction of bright galaxies in the spectroscopic sample,  $N(M < -22.5)/N(M < -20)$ . We then construct 1000 mock samples, each with the same number of galaxies as the actual sample, and use the fraction above (scaled from the bright end fraction of the parent distribution to the bright end fraction of the sample) as the probability that a given galaxy has  $M_R < -22.5$ . From the results, we estimate the probability that the observed number of bright end galaxies or fewer (for the field sample) or the observed number of bright end galaxies or more (for the cluster sample) would have been generated from a luminosity function with this bright end fraction. We then multiply the probabilities for the field and the cluster samples to obtain the probability that the observed bright ends of both samples could have been generated from this luminosity function. We repeat this procedure for a number of bright end fractions.

We find that a common bright end fraction  $\int \phi(M < -22.5)dM / \int \phi(M < -20)dM$  for the field and clusters is ruled out at a level of  $2\sigma$ . The maximum probability (for the assumption that  $\sim 1\%$  of galaxies with  $M_R < -20$  are in the bright tip) is 0.02. Therefore, *the bright end in the cluster GLF is significantly enhanced compared to that of the field.*

Applying the same test to the NEL GLF, we cannot rule out that the bright end fractions in the field and clusters are the same (the probability for drawing the observed numbers of bright galaxies from a GLF with a bright end fraction of  $\sim 1\%$  is  $\sim 0.06$ ).

de Propriis et al. (2003) find  $M_{b_J}^* = -20.07$ , but do not explicitly investigate the possibility of an enhancement of the bright end of the GLF. Assuming  $(b_J - R) \approx 1.5$  for early type galaxies, this  $M_{b_J}^*$  is brighter than we would expect from our results. The reason for this discrepancy is not clear.

## 2.6 Conclusions

Using a stepwise maximum likelihood algorithm that we have modified to account for sampling fractions that vary both in magnitude and surface brightness, we have calculated R-band galaxy luminosity functions for six nearby clusters, as well as composite cluster and field GLFs, from deep spectroscopic samples. The deepest GLF for an individual cluster, A1060, extends to  $M_R = -14 (M^* + 7)$ , making this the deepest spectroscopic survey of the cluster GLF to date. The composite GLF in clusters is consistent with a Schechter function with  $M_R^* = -21.14^{+0.17}_{-0.17}$  and  $\alpha = -1.21^{+0.08}_{-0.07}$ . Employing the same methods and the same cluster fields, we calculate the composite field GLF using non-cluster members, which allows for a homogeneous comparison of cluster and field environments. The field GLF is best fit with a Schechter function with  $M_R^* = -21.15^{+0.16}_{-0.16}$  and  $\alpha = -1.28^{+0.12}_{-0.11}$ , in agreement with other recent determinations of the field GLF.

1. There is a considerable degree of uniformity among the GLFs of our six clusters. We estimate that, at a  $2\sigma$  confidence level, our overall composite cluster GLF is representative for at least 60% of all clusters obeying similar selection criteria down to  $M_R = -18.35$ . Our composite GLFs for emis-

sion line and non-emission line galaxies are representative of at least 40% of similar clusters.

2. The GLFs of non-emission line (quiescent) galaxies vary significantly among clusters.
3. The overall and emission line (star forming/active) GLFs are indistinguishable between the field and clusters in our sample, except for a significant enhancement in the luminous tip of the cluster GLF ( $M_R < -22.5$ ) relative to the field.
4. The GLF of non-emission line (quiescent) galaxies varies significantly between the field and all the clusters, corresponding to a steepening of the faint end slope for  $M_R < -17$ .
5. The fraction of star forming galaxies varies more strongly with environment for giant galaxies ( $\approx M_R^*$ ) than for fainter galaxies ( $\approx M_R^* + 2$ ).

## CHAPTER 3

### THE U-BAND GALAXY LUMINOSITY FUNCTION OF NEARBY CLUSTERS

#### 3.1 Chapter Summary

Despite the great potential of the  $U$ -band galaxy luminosity function (GLF) to constrain the history of star formation in clusters, to clarify the question of variations of the GLF across filter bands, to provide a baseline for comparisons to high-redshift studies of the cluster GLF, and to estimate the contribution of bound systems of galaxies to the extragalactic near-UV background, determinations have so far been hampered by the generally low efficiency of detectors in the  $U$ -band and by the difficulty of constructing both deep and wide surveys. In this paper, we present  $U$ -band GLFs of three nearby, rich clusters to a limit of  $M_U \approx -17.5$  ( $M_U^* + 2$ ). Our analysis is based on a combination of separate spectroscopic and  $R$ -band and  $U$ -band photometric surveys. For this purpose, we have developed a new maximum-likelihood algorithm for calculating the luminosity function that is particularly useful for reconstructing the galaxy distribution function in multi-dimensional spaces (e.g., the number of galaxies as a simultaneous function of luminosity in different filter bands, surface brightness, star formation rate, morphology, etc.), because it requires no prior assumptions as to the shape of the distribution function.

The composite luminosity function can be described by a Schechter function with characteristic magnitude  $M_U^* = -19.82 \pm 0.27$  and faint end slope  $\alpha_U = -1.09 \pm 0.18$ . The total  $U$ -band GLF is slightly steeper than the  $R$ -band GLF, in-

dicating that cluster galaxies are bluer at fainter magnitudes. Quiescent galaxies dominate the cumulative  $U$ -band flux for  $M_U < -14$ . The contribution of galaxies in nearby clusters to the  $U$ -band extragalactic background is  $< 1\%$   $\text{Gyr}^{-1}$  for clusters of masses  $\sim 3 \times 10^{14}$  to  $2 \times 10^{15} M_\odot$ .

### 3.2 Introduction

The galaxy luminosity function (GLF) is one of the most fundamental statistics of galaxy populations. Its shape and variation with environment provide a crucial constraint on any model of galaxy evolution. Recent studies of the GLF in clusters of galaxies, based on spectroscopic surveys, have been carried out in the  $R$ -band (Christlein & Zabludoff, 2003) and the  $b_J$ -band (de Propris et al., 2003). However, there is a large amount of additional information to be gleaned from  $U$ -band GLFs of cluster galaxies.

Why do existing luminosity functions in the  $R$ - and  $b_J$ -band not provide us with a complete picture of the galaxy population in clusters? Different filter bands are sensitive to different stellar populations. For example, blue and near-UV bands are most sensitive to recent star formation, while red and near-IR bands better approximate the total stellar mass. Determinations of the GLF in different magnitude bands are therefore complementary, and the  $U$ -band in particular promises insight into a number of important questions:

First, several studies (Bromley et al., 1998; Madgwick et al., 2002; Christlein & Zabludoff, 2003) show that star-forming and quiescent galaxies have very different GLFs. It is therefore reasonable to expect that star formation would affect the shape of the overall GLF, and that it would do so differently in different filter bands. Other phenomena, such as the presence of dust or the metallicity of a galaxy population, also affect galaxy colors and could introduce inhomogeneities

in comparisons between different filter bands. Understanding how strongly such color-dependent effects influence the GLF is crucial to evaluating the significance of discrepancies between observational determinations of the GLF in different filter bands, and between observed GLFs and model predictions.

Second, with the availability of large telescopes and new detection techniques (Madau et al., 1996), recent years have seen growing interest in observations of clusters at high redshift. Surveys of the GLF in such systems reach to approximately 2 mag fainter than the characteristic bright magnitude  $M^*$  at redshifts up to  $z \approx 1$  (Stanford et al., 2002; Massarotti et al., 2003). High-redshift observations in red and near-IR bands typically observe blue or near-UV rest-frame wavelengths (Fig. 3.1). Measuring the evolution of such fundamental statistics as the GLF in these bands has been complicated by the lack of corresponding low-redshift GLFs. Our  $U$ -band GLF for clusters at  $z \approx 0$  provides a baseline for comparisons to high-redshift clusters that are observed in the  $U$ -band rest-frame band with a comparable depth ( $< M^* + 2$ ).

Third, there is an extragalactic background in the  $U$ -band at a level of 2 to  $4 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ \AA}^{-1}$  (Bernstein, Freedman & Madore, 2002; Henry, 1999). It is thought that a majority of this light is produced by stellar nucleosynthesis rather than AGN or other non-stellar sources. Therefore, the  $U$ -band extragalactic background is of cosmological interest as a cumulative constraint on stellar nucleosynthesis. Furthermore, it is related to the far-UV background, knowledge of which is crucial to understanding the ionization state of the intergalactic medium. From galaxy number counts in deep images, numerical estimates of the contribution of resolved and unresolved normal galaxies to this background have been made (Bernstein, Freedman & Madore, 2002b). Our present work makes it possible to calibrate a relation between cluster mass and cumulative  $U$ -band lu-

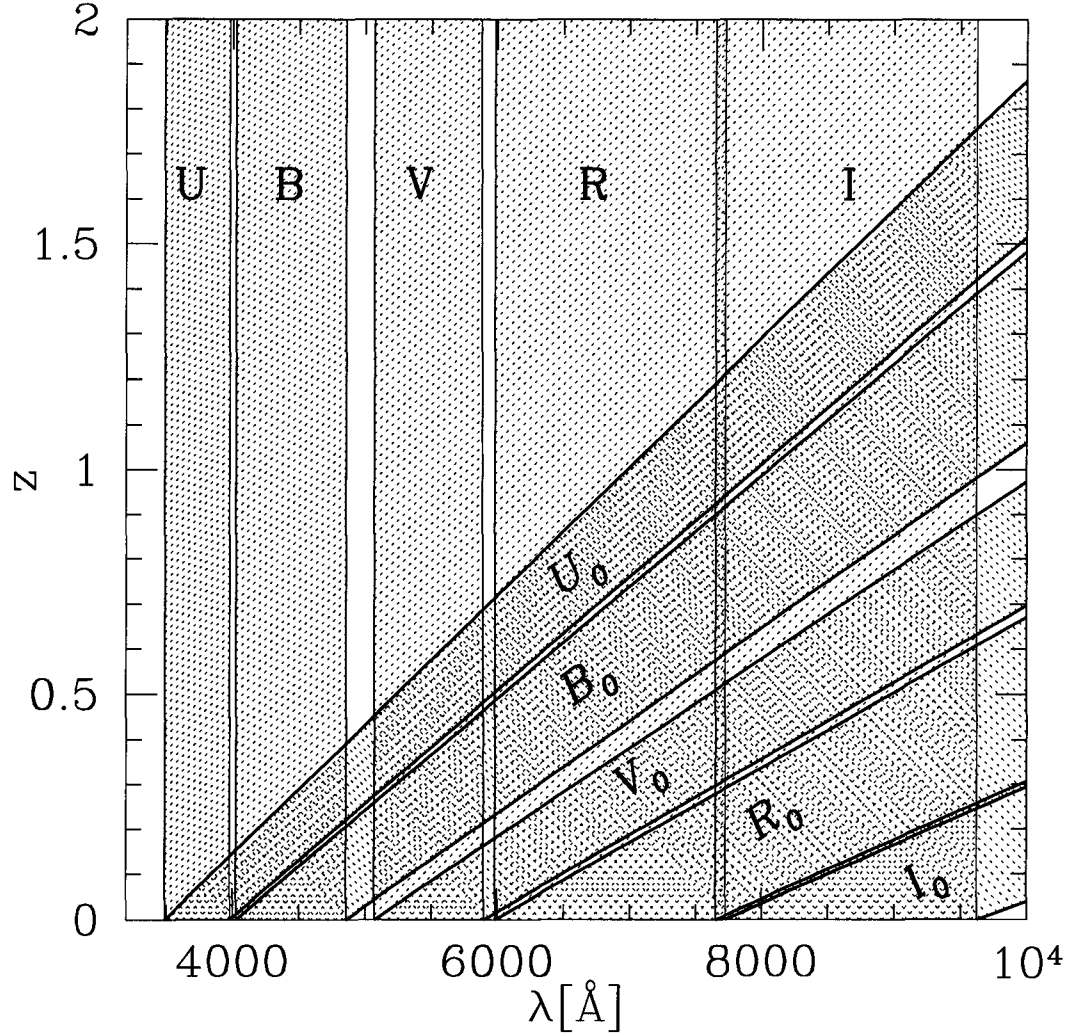


Figure 3.1: Comparison between rest-frame UBVRI and observed  $U_0B_0V_0R_0I_0$  filter bands at different redshifts. Tilted lines show how the rest-frame bands are mapped onto observed bands (vertical solid lines) for any given redshift. At  $z=0.8$ , observations in the  $R$ -band detect the rest-frame  $U$ -band.

minosity. By coupling this with a cluster mass function (Jenkins et al., 2001), we can make an independent estimate for the contribution of nearby clusters (with a mass range of  $10^{14}$  to  $10^{15} M_{\odot}$ ) to the extragalactic  $U$ -band background light. Furthermore, assuming that there is no break in the mass-luminosity relation, we can extrapolate it to roughly estimate the contribution from lower-mass, group-like systems ( $10^{12}$  to  $10^{14} M_{\odot}$ ).

Determinations of the  $U$ -band GLF in clusters at low redshift have so far been complicated by the generally low efficiency of detectors in the  $U$ -band and by the challenges of surveying both wide and deep. Beijersbergen et al. (2002) give a deep determination in the Coma cluster, but use statistical background subtraction rather than a spectroscopically selected sample to account for contamination of the sample by fore- and background galaxies. As Valotto, Moore & Lambas (2001) have demonstrated, this technique can be subject to large systematic errors if the background is inhomogeneous. Spectroscopic samples allow for superior background subtraction.

The recent availability of blue-sensitive wide-field detectors now makes measurement of  $U$ -band GLFs possible. The combination of a  $U$ -band photometric survey (McIntosh, in prep.) with a spectroscopic sample of cluster galaxies from the Las Campanas Nearby Cluster Survey (Christlein & Zabludoff, 2003) enables us to, for the first time, present  $U$ -band GLFs of three nearby clusters from a spectroscopically selected galaxy sample. The availability of  $R$ -band photometric data for the same set of galaxies allows us to make self-consistent comparisons between  $R$ -band and  $U$ -band GLFs from the same sample.

Our procedure of combining additional photometric data in the  $U$ -band with an existing survey whose completeness is known in the  $R$ -band requires us to deal with at least three quantities in the calculation of the GLF: the apparent



$U$ -band magnitude ( $m_U$ ), the  $R$ -band magnitude ( $m_R$ ), and the  $R$ -band surface brightness ( $\mu_R$ ). We have therefore developed a new algorithm for the calculation of the GLF that is particularly suited for such multi-variate analyses. This new algorithm is a variant of maximum likelihood estimators and retains their advantage of being unbiased by density inhomogeneities due to large scale structure. In addition, it offers the benefit that no analytical or binned form for the galaxy parent distribution needs to be assumed *a priori*, a great advantage in multi-variate problems, where analytical forms often do not exist and binning the distribution in a multi-dimensional space is inefficient.

We discuss our cluster sample and the spectroscopic and photometric surveys in §2. In §3 and App. D, we introduce our new GLF algorithm, the Discrete Maximum Likelihood method, and we discuss the completeness of each survey, as well as systematic corrections to account for biases related to color terms in the sampling fraction in App. E. §3.5 gives our results and discussion. We present the GLFs for the three individual clusters in §3.5.1. We determine the composite GLF for all galaxies as well as for emission line (star forming and active) and non-emission line (quiescent) galaxies in §3.5.2. We then compare the  $U$ - and  $R$ -band GLFs, which are calculated from the same sample and from the same processing pipeline (§3.5.3). In §3.5.4, we examine the contribution of clusters to the metagalactic  $U$ -band background. In Appendices F and G, we discuss the effects of the spatial and magnitude limits of our survey on our ability to sample the total  $U$ -band flux from clusters. Our conclusions are presented in §3.6.

Table 3.1. The Cluster Sample

Cluster	N	$\bar{cz}$	$\Delta m$	$cz$ range	$\sigma$	center	area
Cluster	N	[km/s]	[mag]	[km/s]	[km/s]	(J2000.)	[arcmin]
A496	241	$9910 \pm 48$	35.78	7731 - 11728	$728 \pm 36$	04 : 33 : 37.84, -13 : 15 : 44.5	$59.9 \times 58.7$
A754	415	$16369 \pm 47$	36.90	13362 - 18942	$953 \pm 40$	09 : 08 : 32.00, -09 : 37 : 00.0	$60.1 \times 58.5$
A85	280	$16607 \pm 60$	36.94	13423 - 19737	$993 \pm 53$	00 : 41 : 50.46, -09 : 18 : 11.6	$60.2 \times 58.3$

Notes:  $N$  is the number of sampled galaxies per cluster.  $cz$  is the mean redshift,  $\Delta m$  the distance modulus. Center coordinates and survey area are for the  $U$ -band photometric survey, which has a smaller coverage than the  $R$ -band photometric or the spectroscopic survey.

### 3.3 The Data

#### 3.3.1 The Cluster Sample

Our sample consists of three clusters (Abell 496, Abell 754, Abell 85) from the original spectroscopic and  $R$ -band imaging sample of Christlein & Zabludoff (2003). The Christlein & Zabludoff clusters were selected based on 1) their visibility from Las Campanas, 2) the availability of some prior spectroscopic and X-ray data in the literature, 3) their redshifts, which allowed us to sample the cluster to at least one virial radius with the  $1.5^\circ \times 1.5^\circ$  field of the fiber spectrograph field, and 4) their range of velocity dispersions, which suggest a wide range of virial masses. The properties of these clusters (for  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.27$  and  $\Omega_\Lambda = 0.73$ , as applied throughout this paper) are given in Table 1. We refer readers to Christlein & Zabludoff (2003) for details on the spectroscopic survey and data reduction.

#### 3.3.2 $R$ -band Survey

We derive our master galaxy catalog from a photometric survey of the clusters in the  $R$ -band (Christlein & Zabludoff, 2003). This catalog is complete within certain

magnitude and surface brightness limits, so we use it as the reference to estimate the completeness of the spectroscopic and the  $U$ -band photometric catalogs as a function of  $(m_R, \mu_R)$ . Details of this survey, the image reduction, photometry, and the construction of the master catalog are in Christlein & Zabludoff (2003).

### 3.3.3 $U$ -band survey

The  $U$ -band cluster galaxy data come from wide-field ( $1^\circ \times 1^\circ$ ) imaging of A85, A496, and A754 during a January 2000 run using the NOAO Mosaic Imager on the Kitt Peak National Observatory (KPNO) 0.9-meter Telescope. Complete details of the sample selection, observations, reductions and photometric calibrations are reported in McIntosh, Rix, Caldwell & Zabludoff (2004, in prep.). The final sample contains  $U$ -band and  $V$ -band magnitudes for a total of 631 spectroscopically confirmed cluster member galaxies, and is comparable in depth and in membership coverage to the most comprehensive study of the Coma cluster by Terlevich, Caldwell, & Bower (2001). In physical units, the  $U$ -band survey region covers about  $1.7 \times 1.6 \text{ Mpc}^2$  in A496,  $2.7 \times 2.6 \text{ Mpc}^2$  in A754, and  $2.7 \times 2.6 \text{ Mpc}^2$  in A85.

The data are well-flattened and carefully corrected to ensure uniformity of the photometric zero point using a customized reduction pipeline that follows standard image reduction techniques and uses the IRAF<sup>1</sup> environment. We perform the basic reduction of the  $U$ -band data using the IRAF MSCRED package. The images are spatially and spectrally flattened ( $\lesssim 1\%$  deviations globally) using an optimized night-sky flat-field frame constructed from individual science exposures with all objects masked. We calibrate the data astrometrically to the USNO-v2.0 system (Monet et al., 1996) and photometrically to Johnson (1966)  $U$

---

<sup>1</sup>IRAF is distributed by the National Optical Astronomical Observatories, which are operated by AURA, Inc. under contract to the NSF.

magnitudes on the Landolt (1992) system.

We perform source detection and extraction using SExtractor (Bertin & Arnouts, 1996) with the following configuration parameters defining our definition of an imaged source: minimum of 5 detected pixels (DEBLEND\_MINAREA) above a background threshold of  $3\sigma_{\text{bkg}}$  (DETECT\_THRESH), with overlapping sources deblended if the contrast between flux peaks associated with each object is  $\geq 0.05$  (DEBLEND\_MINCONT). We confirm that these parameters provide good source detection and deblending by visually inspecting random regions from each image. We reject sources flagged ( $\text{FLAGS} \geq 4$ ) as saturated or otherwise bad, and we exclude all sources within 1 arcmin of image edges. The empirical magnitude limits where the  $U$ -band source count distributions flatten and turn over are roughly 20.7 (A85), 21.0 (A496), and 20.5 (A754) mag. We separate stellar and extended sources by fitting a PSF-convolved bulge+disk model to the light profiles of detected sources using GIM2D (Simard et al., 2002) following the method described in McIntosh et al. (in prep.). This method is robust to  $V = 19$  mag, which corresponds to  $\sim 20.5$  in  $U$  for most cluster members.

From 975 redshift coordinates within the coverage of the three cluster  $U$  images, we find 727 cluster members (with  $cz_i = \langle cz \rangle_{\text{cl}} \pm 3\sigma_{\text{cl}}$ ) and 248 with recession velocities outside of these bounds. We cross-correlate the coordinates of member galaxies from the redshift data with the  $U$ -band source positions and achieve 631 members from our imaging catalog. We define image/redshift matches to be the nearest within  $5''$  and find the mean coordinate separation is  $< 2''$ . Only 2.4% (12) of the 88 redshift positions without a  $U$  image detection are brighter than  $U = 19.5$ . Finally, 112  $U < 19.5$  extended sources have no redshifts due to the fractional incompleteness of the spectroscopic sample at  $R > 16$  (Christlein & Zabludoff, 2003). Therefore, our  $U$ -band galaxy catalog from clusters A85, A496,

and A754 is  $\sim 85\%$  complete down to  $U \sim 19.5$  mag.

We use the dust maps by Schlegel, Finkbeiner & Davis (1998) for extinction corrections and apply a k-correction of  $K_U = 0.065(z/0.02)$ , as appropriate for early-type galaxies (Pence, 1976), to the  $U$ -band magnitudes. We apply this correction to all galaxies in this sample, because cluster galaxies are predominantly early types, individual morphological classifications are not available, and the k-correction for extreme late types differs from the value for early types by less than 0.1 mag even for the most distant sample cluster.

### 3.4 Calculating the Luminosity Function

#### 3.4.1 A New Method

For reconstructing the luminosity function of our survey, we use a new statistical method, which we refer to as the *Discrete Maximum Likelihood* Method. The DML retains the advantages of Maximum Likelihood estimators in being unbiased by density inhomogeneities in a sample and is therefore easily applicable to field as well as cluster samples. In contrast to most maximum likelihood-based LF algorithms (Sandage, Tamman & Yahil, 1979; Efsthathiou, Ellis & Peterson, 1988; Blanton et al., 2001), which were developed when luminosity alone was the primary variable of interest in determining galaxy distribution function, the DML algorithm does not assume an ansatz (binned or analytic) for the distribution function *a priori*, and is therefore independent of the dimensionality of the parameter space of interest. It is therefore ideally applicable to multivariate distribution functions, i.e., the abundance of galaxies as a simultaneous function of luminosity in different filter bands, surface brightness, environment, star formation, morphology, etc. The derivation of the Discrete Maximum Likelihood Method is discussed in detail in Appendix D.

### 3.4.2 The Sampling Fraction

The reconstruction of the GLF, as described in App. D, requires knowledge of the sample completeness,  $f(\vec{x}_h \mid F_i)$ .  $f$  is the probability that we have both  $U$ -band photometry and spectroscopic information for a given object with certain physical properties  $\vec{x}_n$  (e.g., absolute magnitude and surface brightness) if it is in a particular field  $F_i$  (characterized by redshift, Galactic foreground extinction, position on the sky). Due to the design of our survey, the sampling fraction is not known analytically, but has to be reconstructed from the data. Our  $R$ -band detection catalog serves as the reference against which we calculate the completeness of the spectroscopic and  $U$ -band photometric samples. Appendix E describes the determination of the sampling fraction, as well as systematic corrections that are necessary to account for color selection effects.

### 3.4.3 Consistency checks

The introduction of a new method for the calculation of the luminosity function requires us to demonstrate its consistency with established algorithms. We confirm this by comparing the composite  $R$ -band GLF obtained from the SWML method (as used in Christlein & Zabludoff (2003)) with our new algorithm. Fig. 3.2 shows these two GLFs (upper curves, unnormalized), determined for the three clusters in our survey. The GLFs are statistically indistinguishable. The results of both methods are equivalent for calculating distribution functions in low-dimensional parameter spaces, but the DML offers greater efficiency and convenience in treating multi-dimensional problems.

We also test whether our sampling fraction model accounts correctly for the incompleteness of the  $U$ -band photometric sample. This is particularly important because we model the completeness of the  $U$ -band catalog as a function of  $(m_R, \mu_R)$ , while the strongest dependence should be on  $(m_U, \mu_U)$ . We calculate

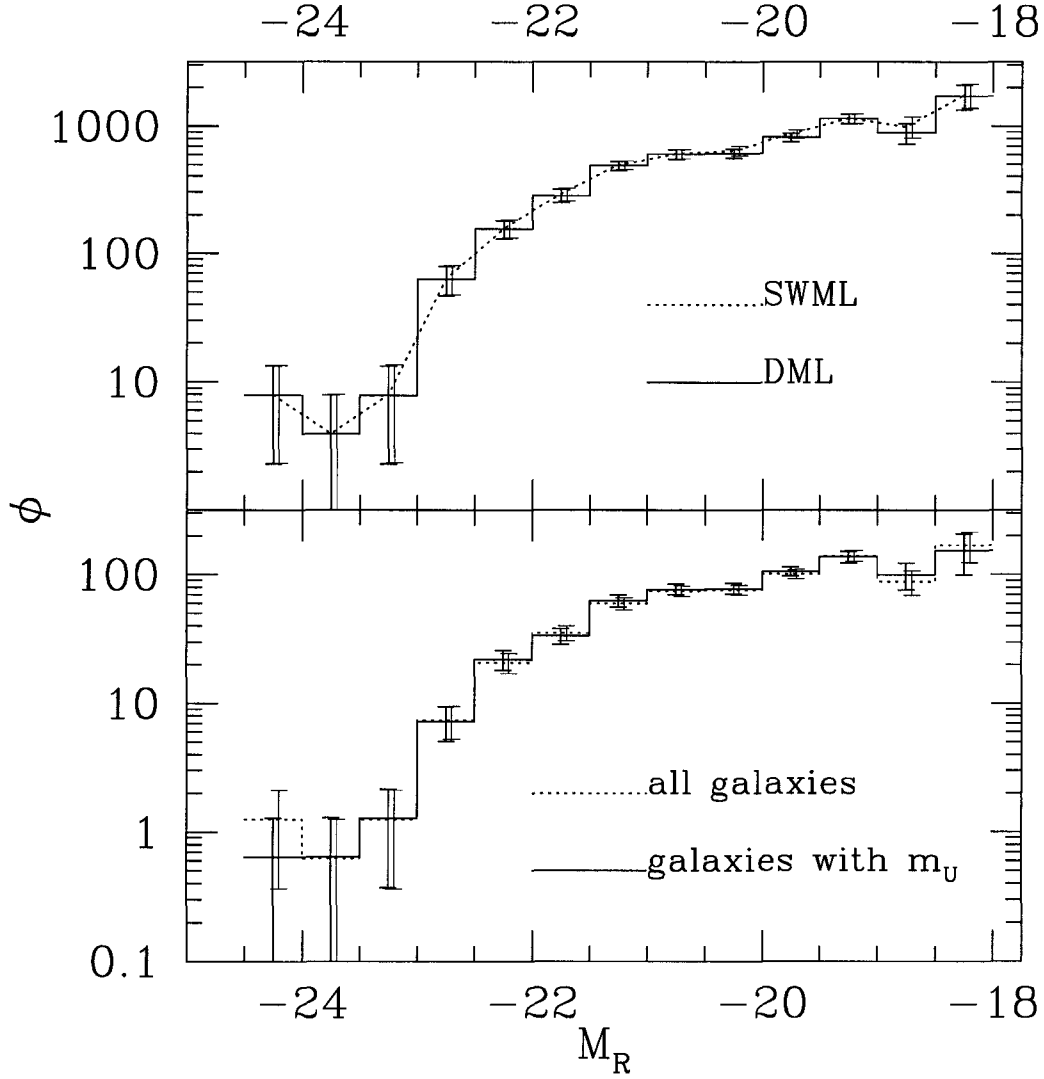


Figure 3.2: Upper pair of curves: Comparison between composite  $R$ -band cluster GLFs using the Stepwise Maximum Likelihood method (dashed line; Efstathiou, Ellis & Peterson 1988) and our Discrete Maximum Likelihood method (solid line), determined over the area of the  $R$ -band photometric sample. Normalization is arbitrary. Lower pair of curves: Comparison between GLFs from all galaxies within  $U$ -band survey area (dashed line) and from those galaxies with  $U$ -band photometry only, after completeness correction (solid line). Normalization is galaxies / magnitude / cluster.

the composite  $R$ -band GLF from all galaxies in the  $R$ -band master catalog that lie within the spatial boundaries of the  $U$ -band survey, regardless of whether we know their  $U$ -band magnitude. We then calculate an  $R$ -band GLF from galaxies for which we have  $U$ -band photometry, and apply the appropriate completeness corrections. The agreement between the two GLFs is again excellent, indicating that the incompleteness of the  $U$ -band catalog does not introduce systematic biases. We emphasize that this  $R$ -band GLF is plotted using the same weighting and normalization factors as the  $U$ -band GLF that we present below.

#### 3.4.4 Coverage of $(M_R, M_U, \mu_R)$ Parameter Space

Our survey has magnitude limits in two different filter bands,  $R$  and  $U$ . At each  $M_U$ , the  $R$  band magnitude limit introduces an incompleteness in the calculated  $U$ -band GLF, because, for any given  $M_U$ , some fraction of galaxies are too faint in  $M_R$  to have been sampled. To avoid this problem, we calculate  $U$ -band GLFs only down to a limit at which the fraction of lost galaxies is likely to be small.

We use Fig. 3.3 to identify this completeness limit for our  $U$ -band GLF. The figure shows the completeness-corrected galaxy distribution (greyscale) in the  $(M_U, M_R)$  plane. Dots represent the individual galaxies in the sample. The bold solid line shows the faintest  $M_R$  to which a galaxy with a given  $M_U$  and any  $\mu_R$  typical for galaxies at that  $M_U$  could have been sampled. The effective absolute magnitude limit at the faint end is  $M_R \approx -17.3$ . Fig. 3.3 shows that, for  $M_U > -18$ , the galaxy distribution has not been sampled completely; in any  $M_U$  bin substantially fainter than that limit, a substantial fraction of galaxies are likely to be missing due to the  $M_R$  limit. We adopt  $M_U = -17.5$  as the absolute magnitude limit for our analysis of the  $U$ -band GLF; at this limit, most of the galaxy distribution function has been sampled, and only a small fraction of galaxies ( $\sim 5\%$ , if the galaxy distribution in  $M_R$  for a given  $M_U$  is approximately



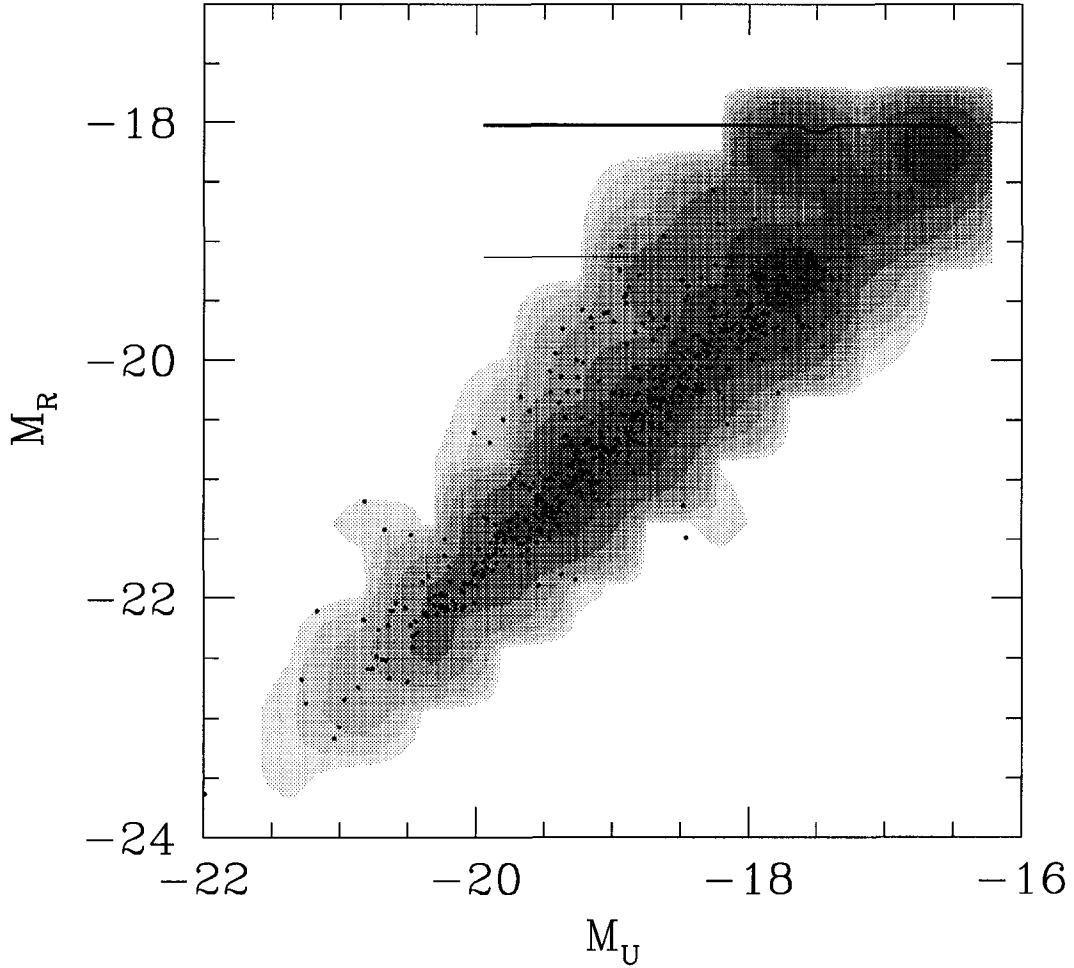


Figure 3.3: Galaxy distribution in the  $(M_U, M_R)$  plane in greyscale as calculated using the Discrete Maximum Likelihood method. Points denote cluster members with both redshifts and  $U$ -band photometry available. The bold horizontal line at  $M_R \approx -18.0$  gives the effective absolute magnitude limit in  $R$ , i.e., the  $M_R$  at which the sampling fraction is zero for a galaxy of a given  $M_U$  and typical  $\mu_R$  at that  $M_U$ . Galaxies fainter than  $M_U \approx -17.5$  are too faint in  $M_R$  to have been sampled, thus motivating our absolute magnitude limit of  $M_U = -17.5$ . The more distant clusters are only sampled to  $M_R \approx -19.0$  (thin horizontal line), motivating a limit of  $M_U = -18.5$ .

Gaussian) in the faint- $M_R$  tail are likely to be lost. The two more distant clusters, A754 and A85, have only been sampled to  $M_R = -19.0$ . Therefore, we consider the individual  $U$ -band GLFs for these clusters to be reliable down to  $M_U = -18.5$ .

The projection of the distribution onto the  $(M_U, M_R)$  plane is noticeably broader at fainter magnitudes than at the bright end, a result of the greater fraction of blue, emission line galaxies.

### 3.5 Results and Discussion

#### 3.5.1 $U$ -band GLFs for Individual Clusters

Fig. 3.4 shows the individual GLFs for the three clusters (calculated by the method in §3.4.1) superimposed on the composite GLF, which we calculate in §3.5.2. Table 3.2 gives the numerical values for the cluster GLFs down to the effective sample limits.

The shapes of the individual cluster GLFs are consistent with each other within  $1\sigma$  under a  $\chi^2$  test if the normalization is adjusted to optimize the agreement. However, the cluster GLFs are significantly offset in normalization. By drawing 1000 Monte Carlo samples of galaxies from the three cluster GLFs, we find the expected number of galaxies at  $M_U < -18.5$ :  $91.8 \pm 11.4$  in A496,  $222.0 \pm 17.4$  in A754, and  $143.1 \pm 13.8$  in A85 over their respective survey regions. If we truncate the survey regions to comparable physical radii (taking into account both the angular diameter distance and the fact that characteristic length scales tend to increase linearly with  $\sigma$  (Girardi et al., 1998)), the numbers are  $91.8 \pm 11.4$  in A496,  $189.2 \pm 15.8$  in A754, and  $120.9 \pm 12.9$  in A85. The differences between A754 and the other two clusters are significant and most likely reflect the higher mass of A754.

In the next section, we will form a composite GLF from the three individual

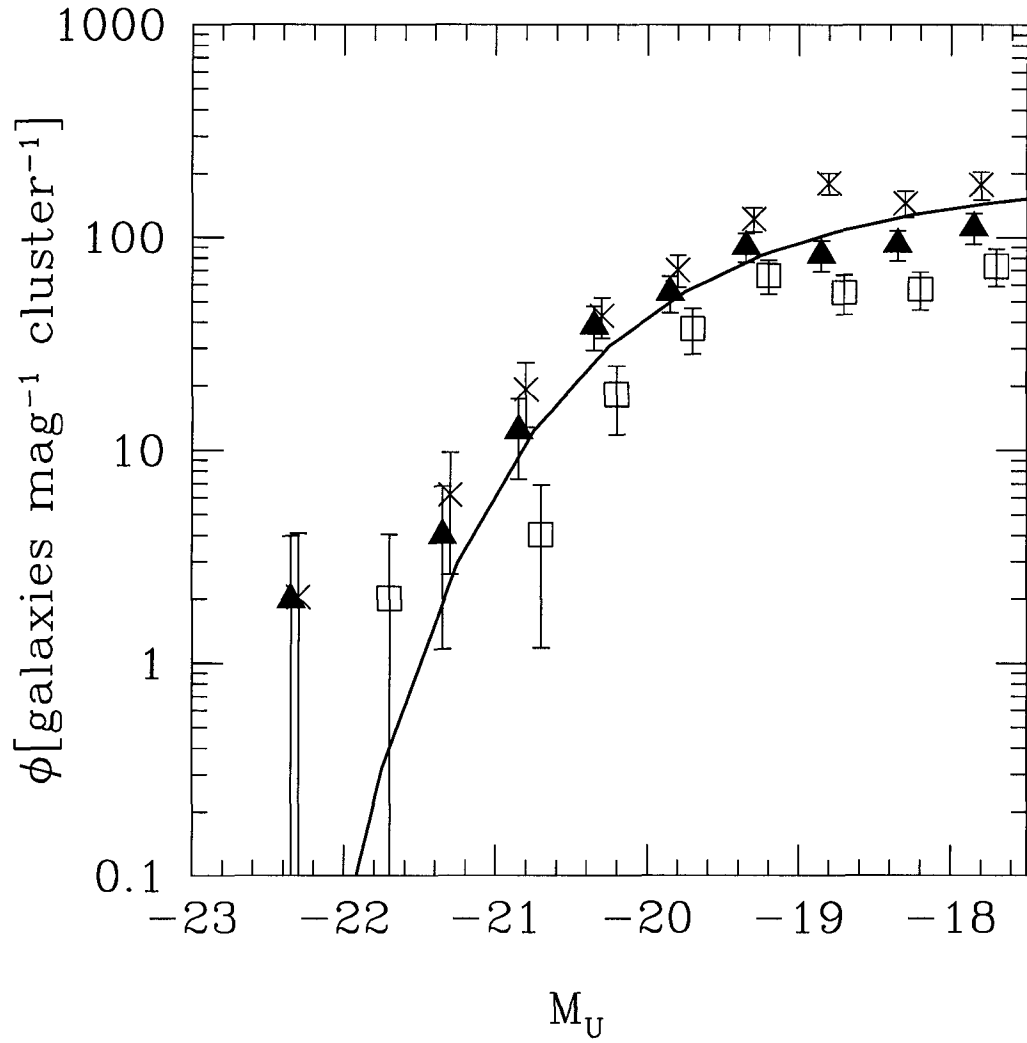


Figure 3.4: *U*-band GLFs for individual clusters. Normalization of all GLFs is preserved. We show the GLFs for A496 (open squares), A754 (crosses), and A85 (triangles). Solid line shows best fit Schechter function to composite GLF.

Table 3.2. *U*-band GLFs.

$M_U$	A496	A754	A85
-22.25	...	2.04(1)	1.98(1)
-21.75	2.02(1)	...	...
-21.25	...	6.21(3)	3.98(2)
-20.75	4.03(2)	19.28(9)	12.38(6)
-20.25	18.31(8)	42.88(21)	38.42(18)
-19.75	37.28(17)	70.49(34)	55.15(26)
-19.25	66.42(30)	122.43(56)	90.64(42)
-18.75	55.13(22)	180.50(73)	82.94(36)
-18.25	57.23(24)	145.09(50)	92.78(38)
-17.75	73.61(25)	177.61(44)	111.43(36)

Note. — Corrected number of galaxies per mag and cluster. Number of sampled galaxies per bin in parentheses.

clusters. de Propriis et al. (2003) find evidence that the shape of the GLF in clusters varies as a function of distance from the cluster center, raising the question whether it is legitimate to form a composite GLF from three clusters with inhomogeneous sampling radii. In our sample, we find no statistically significant differences in the shapes of the GLFs of A754 and A85 between the truncated and untruncated samples, which legitimizes our approach to forming the composite GLF.

### 3.5.2 Composite $U$ -band GLFs

We apply the DML algorithm to the complete sample of three clusters and then bin the recovered galaxy distribution function over  $M_U$ . We repeat this process for subsamples of emission line (EL) and non-emission line (NEL) galaxies separately. EL galaxies are defined as having an equivalent width of the [OII]  $\lambda 3727$  doublet of  $5\text{\AA}$  or more and are thus star-forming or active galaxies. NEL galaxies have an equivalent width of less than  $5\text{\AA}$  and are hereafter classed as quiescent galaxies. Fig. 3.5 shows all three composite  $U$ -band GLFs. We have applied corrections for B-R color terms, as described in App. E, in all cases.  $1\sigma$  error bars are based on Poisson errors. Table 3.3 gives the numerical values for the three GLFs, including the numbers of galaxies in each bin. All GLFs are normalized as described in App. D to represent the average number of galaxies per magnitude and cluster within the region of the  $U$ -band photometric survey.

We have fitted Schechter functions (Schechter, 1976) to these GLFs. The functional form is

$$\phi(M) = \phi^* (10^{0.4(M^* - M)})^{1+\alpha} e^{-10^{0.4(M^* - M)}}, \quad (3.1)$$

where the nomenclature  $\phi$  is used to distinguish this function, which is only a function of absolute magnitude, from the  $\varphi$  in Eqn. D.1, which is a function of an argument of arbitrary dimensionality.

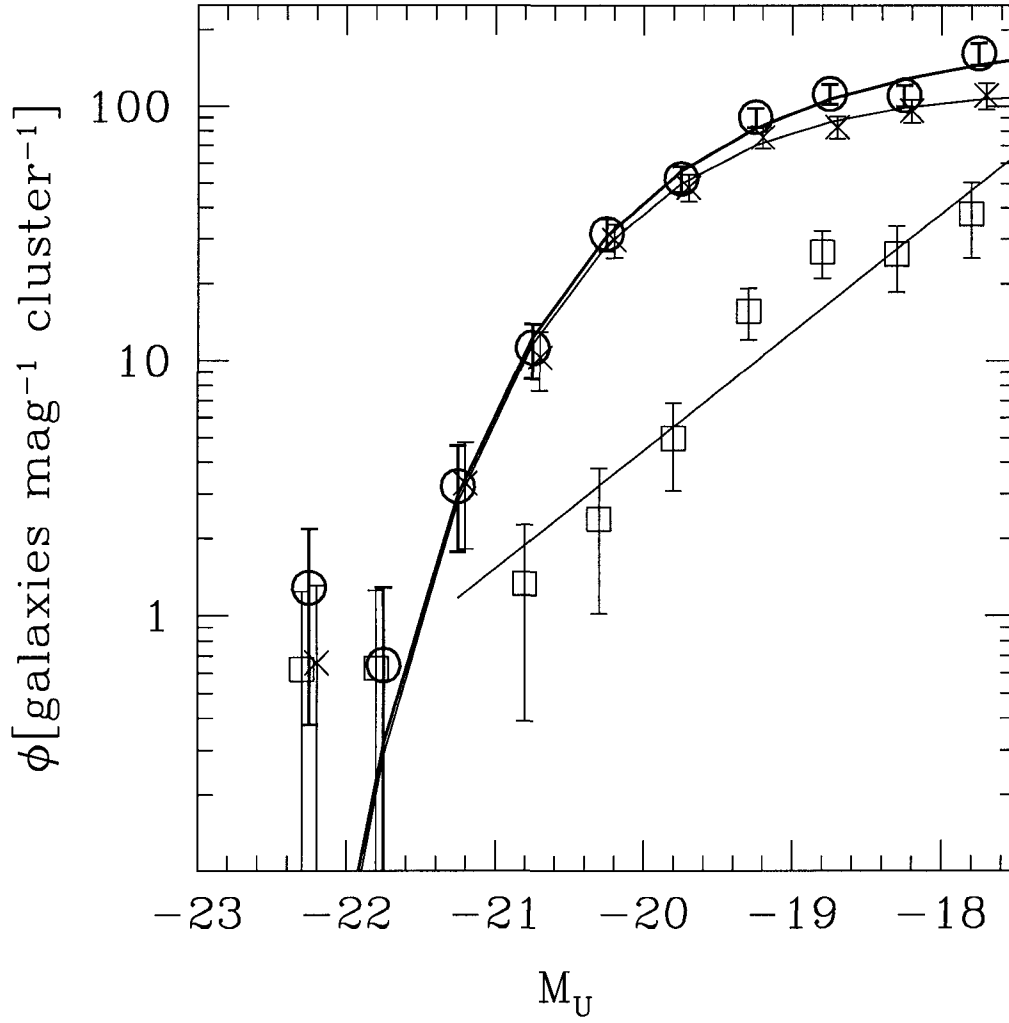


Figure 3.5: *U*-band GLFs for all galaxies (circles), non-emission line galaxies (crosses) and emission line galaxies (open squares). Data points for emission and non-emission line galaxies are displaced by  $\pm 0.05$  mag for display purposes only. Schechter functions for overall (bold) and non-emission line galaxies, as well as the power law fit for emission line galaxies, are also shown.

Table 3.3. Composite  $U$ -band GLFs.

$M_U$	all	EL	NEL
-22.25	1.29(2)	0.62(1)	0.66(1)
-21.75	0.64(1)	0.63(1)	...
-21.25	3.22(5)	...	3.31(5)
-20.75	11.22(17)	1.33(2)	10.29(15)
-20.25	31.53(47)	2.40(3)	29.78(43)
-19.75	52.03(77)	4.98(7)	48.11(68)
-19.25	90.73(128)	15.65(19)	75.71(104)
-18.75	112.20(131)	26.65(22)	83.11(100)
-18.25	110.55(112)	26.15(12)	96.40(89)
-17.75	161.86(105)	37.70(9)	110.45(70)

Note. — Corrected number of galaxies per mag and cluster. Number of sampled galaxies per bin in parentheses.

The faint end slope is not well-constrained in this sample because the GLFs are only sampled to a fairly bright magnitude limit. The best fit Schechter parameters are  $M_U^* = -19.82_{-0.27}^{+0.27}$ ,  $\alpha_U = -1.09_{-0.18}^{+0.18}$ , and  $\phi_U^* = 142 \pm 5$  for the overall GLF. Errors in  $\phi_U^*$  are for the best fit  $\alpha$  and  $M^*$  held fixed. Corrections for the B-R color terms in the sampling fraction, as described in App. E, are  $\ll 1\%$  in  $\phi$  and thus negligible for the overall GLF.

The NEL GLF is also described well by a Schechter function. The best fit parameters are  $M_U^* = -19.77_{-0.30}^{+0.28}$ ,  $\alpha_U = -0.97_{-0.18}^{+0.22}$  and  $\phi_U^* = 133 \pm 6$  for the NEL GLF. Corrections for the B-R color terms are more important here, because the color difference between EL and NEL galaxies is larger than between field and cluster galaxies. Without our color corrections, the faint end slope would be shallower by  $\Delta\alpha \approx 0.11$ .

For the EL galaxies, both a Schechter function and a power law provide acceptable fits. To reduce the number of free parameters, and because the Schechter parameters are very weakly constrained by the EL GLF, we decide to fit a power law of the functional form

$$\phi(M) = \phi^* (10^{0.4(-(M+20))})^{1+\alpha} \quad (3.2)$$

The best fit parameters are  $\alpha_U = -2.16_{-0.19}^{+0.16}$  and  $\phi^* = 4.45_{-0.52}^{+0.52}$ . Again, color corrections are important here: the uncorrected EL GLF is steeper by  $\Delta\alpha \approx 0.12$ .

### 3.5.3 Comparison to *R*-band GLFs

Earlier attempts to determine the GLF in clusters have yielded conflicting results, with reported faint end slopes varying from  $\alpha \approx -1.0$  to  $\alpha \approx -2.2$ . One possible explanation of these discrepancies could lie in the use of different filter bands. Comparisons between two recent, spectroscopically selected studies of the cluster GLF (Christlein & Zabludoff, 2003; de Propriis et al., 2003) do not show evidence



for systematic variations in the faint end slope  $\alpha$  between the  $R$ -band and the  $b_J$ -band. Our present work provides an even more stringent test for the uniformity of the GLF across different filter bands, because the  $U$ -band is even more sensitive to star formation and dust than the  $b_J$  band. If differences in the GLFs exist, they are likely to be revealed in a comparison between  $U$  and  $R$ .

We first compare the  $U$ -band and  $R$ -band GLFs using a  $\chi^2$  test, allowing for one degree of freedom (the shift in  $M$ ) to minimize  $\chi^2$ . This approach not only tests the GLFs for consistency in shape, but also provides a numerical estimate of the magnitude offsets between the  $U$ -band and the  $R$ -band. The magnitude offset that optimizes  $\chi^2$  is  $U - R = 1.56 \pm 0.04$ . The maximum probability of consistency is  $p = 0.10$ . Although the  $R$ - and  $U$ -band GLFs are formally consistent under this  $\chi^2$  test, a comparison of the ratio of galaxies with  $M_U^* < M_U < -17.5$  to galaxies with  $M_U \leq M_U^*$  reveals that the  $U$ -band GLF is slightly steeper; the ratio is  $8.3 \pm 0.9$  for the  $U$ -band versus  $5.2 \pm 0.5$  for the  $R$ -band. Fig. 3.6 shows a superposition of the  $U$ -band GLF and the shifted  $R$ -band GLF.

The difference in the shape of the two GLFs indicates that there is a color gradient,  $\frac{\partial(U-R)}{\partial R}$ , in the sample: fainter cluster galaxies are bluer. What is the cause of this color gradient? Factors that could influence  $(U - R)$  colors are star formation, nuclear activity, dust, and metallicity. Of these, AGN do not appear to play a significant role in our study: the NASA Extragalactic Database lists only one Seyfert galaxy in our sample. We now test whether differences in star formation activity alone could account for the color gradient. For this purpose, we separate our sample into star-forming (EL) and quiescent (NEL) galaxies. If star formation activity is solely responsible for the color gradient, then allowing for star-forming and quiescent galaxies to have different colors should account for the discrepancy in the overall GLF.

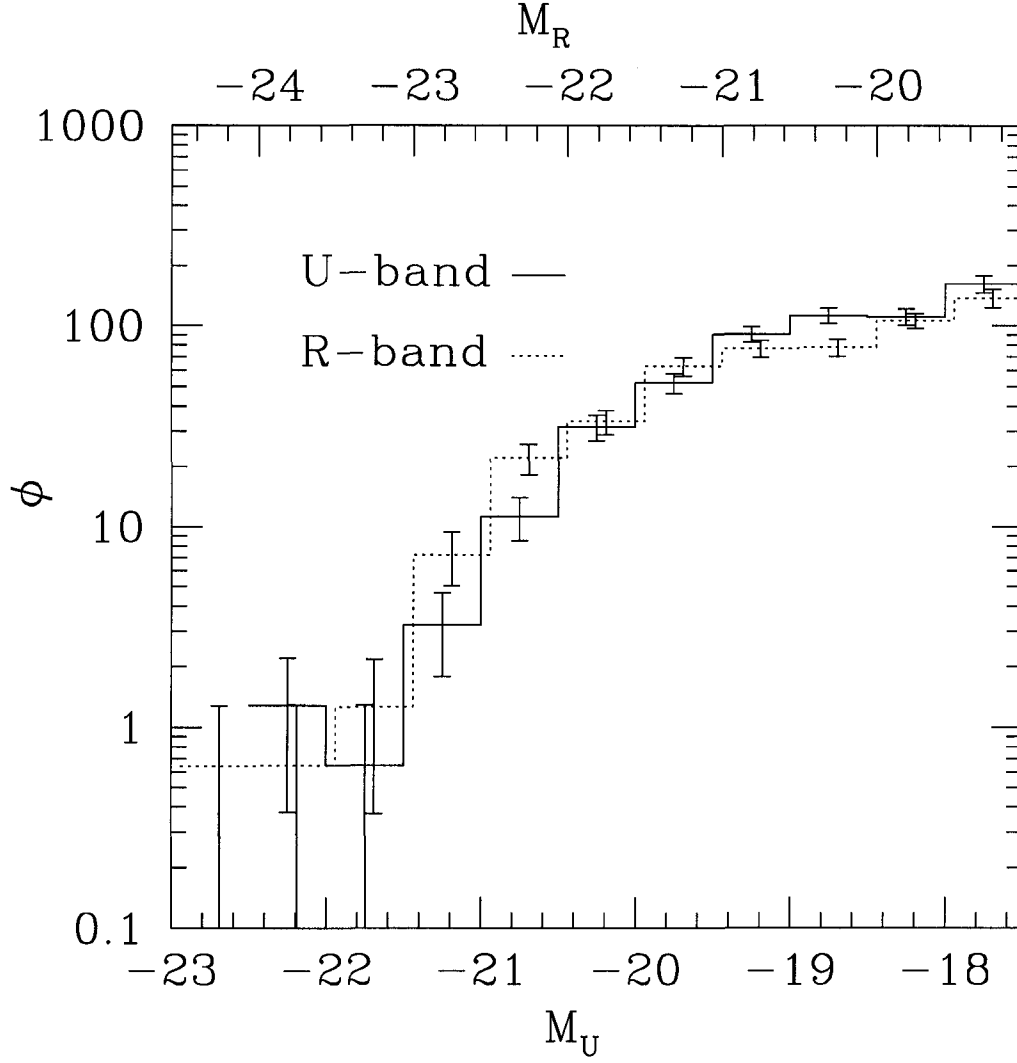


Figure 3.6: Superposition of  $U$ -band and  $R$ -band GLFs. The  $R$ -band GLF (dotted line) has been shifted by  $\Delta = U - R = 1.56$  mag to maximize the agreement between both GLFs. Detailed analysis shows the  $U$ -band GLF to be steeper than the  $R$ -band GLF.

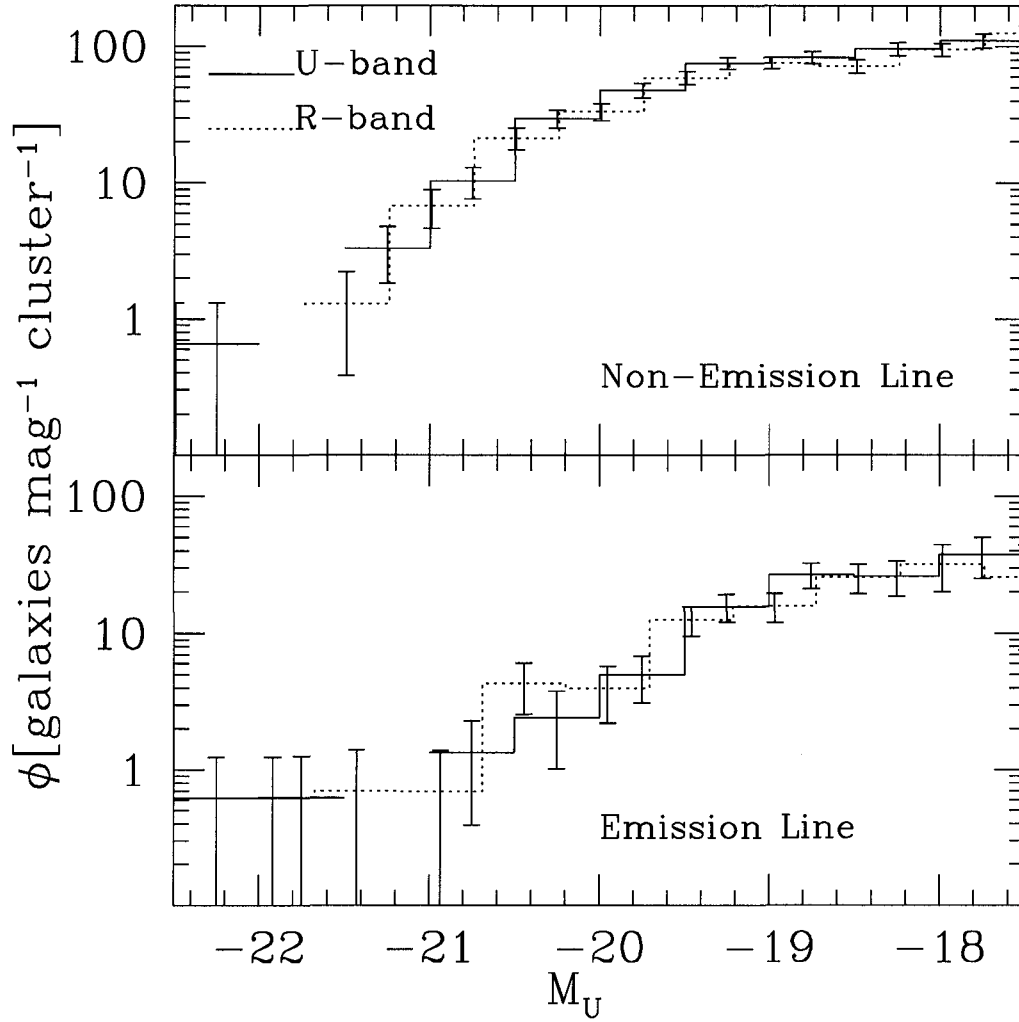


Figure 3.7: Same as Fig. 3.6 for emission line (top) and non-emission line (bottom) galaxies. The  $R$ -band emission line GLF has been shifted by  $\Delta = U - R = 0.76$ , and the non-emission line GLF by  $\Delta = U - R = 1.65$ . Within each subsample, we find no significant discrepancy in the GLF shapes between the  $U$ - and  $R$ -band.

Fig. 3.7 compares the  $U$ -band and  $R$ -band GLFs of the EL and NEL subsamples. The optimal magnitude shift that minimizes  $\chi^2$  for the EL GLFs is  $0.76 \pm 0.30$ . The maximum  $\chi^2$  probability is 0.89, indicating that these GLFs are consistent with each other in shape. For the NEL galaxies, the  $\chi^2$  comparison yields a magnitude shift of  $U - R = 1.65 \pm 0.13$  with a probability of  $p = 0.97$ . The ratio of faint to bright galaxies, determined as above and relative to the same absolute magnitude,  $M_U^*$ , is not significantly different between the  $U$ - and the  $R$ -band in either the EL or NEL subsample, suggesting that  $U - R$  colors are more homogeneous in the subsamples than in the overall sample.

If we shift each of the  $R$ -band EL and NEL GLFs by their respective best  $U - R$  magnitude offset and combine them, the resulting GLF has a steeper faint-to-bright ratio than the original  $R$ -band GLF, indicating that star formation accounts for some of the discrepancy between the  $U$ - and  $R$ -band GLFs. However, the shifted  $R$ -band and the  $U$ -band GLF are still significantly different with regard to the ratio of faint to bright galaxies. Therefore, our simple bisection of the sample into star-forming and quiescent galaxies by [OII] EW does not completely explain the color gradient. Other effects, such as dust and metallicity gradients with luminosity, are probably responsible for the residual color gradient, or else the above binning is not fine enough to homogenize colors in each subsample. Clarification of this question will have to wait for future investigations with larger surveys.

Color gradients can introduce differences in the faint end slope between different filter bands. Can these differences account for the wide range of  $\alpha$  that has been reported in the literature? To answer this question, we simultaneously fit Schechter functions to both the  $U$ - and  $R$ -band GLFs to determine the difference  $\alpha_U - \alpha_R$  and its uncertainty interval. If we fix  $M_U^* - M_R^*$  at its best fit value of 1.93, we obtain  $\alpha_U - \alpha_R = +0.03_{-0.11}^{+0.14}$ . This is consistent with the expectation

that color gradients within the galaxy population introduce only a small difference of  $\alpha_U - \alpha_R = -\frac{\partial(U-R)}{\partial R}(\alpha_R + 1)$ . With an estimated  $\frac{\partial(U-R)}{\partial R} = -0.08$ , based on the  $U - R$  color offsets for EL and NEL galaxies above and the EL/NEL ratios from Christlein & Zabludoff (2003), the expectation value for our sample is  $\alpha_U - \alpha_R = -0.01$ . Therefore, the faint end slope  $\alpha$  in clusters is only a weak function of the filter band. This is qualitatively consistent with Paolillo et al. (2001), who found the faint end slope  $\alpha$  to be nearly identical in the  $i$ -,  $r$ -, and  $g$ -bands.

### 3.5.4 Contribution from Clusters and Groups to Near-UV Background

Understanding the contribution from normal galaxies to this background light provides important constraints on the star formation history of the universe and the relative importance of non-stellar sources, such as AGN or Ly- $\alpha$  recombination radiation from intergalactic gas (Tyson, 1995). Observational estimates for the extragalactic background intensity in the near-UV are on the order of  $2 - 4 \times 10^{-9} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ \AA}^{-1}$  (Bernstein, Freedman & Madore, 2002; Henry, 1999). Past studies (Pozzetti et al., 1998; Bernstein, Freedman & Madore, 2002b) have estimated the contributions from field galaxies to this background radiation from number counts in the Hubble Deep Field. Our data provide an independent estimate of the contribution of nearby massive clusters, which are different from the field with regard to their constituent galaxy populations and evolutionary history.

What is the contribution of clusters such as those in our sample to the extragalactic background in the  $U$ -band? The rate of increase of the  $U$ -band background intensity from a population of sources with differential luminosity  $L_{U,\lambda}$  and differential spatial density  $dn/dM$  (where  $M$  is the mass) is

$$\frac{dI_\lambda}{dt} = \frac{c}{4\pi} \int L_{U,\lambda}(M) \frac{dn}{dM} dM. \quad (3.3)$$

An estimate of the mass function,  $dn/dM(M)$ , is given by Jenkins et al. (2001) from numerical simulations. We use their mass function for cosmological parameters of  $H_0 = 70$ ,  $\Omega_m = 0.3$  and  $\Omega_\lambda = 0.7$ .

$L_{U,\lambda}(M)$  can be constrained from our data. Table 3.4 shows the sampled  $U$ -band luminosity within the common magnitude limit of  $M_U = -18.5$  for all galaxies, for EL (star forming) galaxies and for NEL (quiescent) galaxies. All clusters have been sampled to at least this limit. We discuss possible contributions from fainter galaxies in App. F, but find that, unless the GLF shows a strong upturn at the faint end, we have sampled most of the  $U$ -band luminosity of the clusters. The  $U$ -band luminosity is not corrected for radial sampling limits either, but represents the total emission from a cluster's galaxies within our survey area, which is approximately one harmonic radius. The table also gives the cluster mass,  $M_{200}$ , taken from Reiprich & Böhringer (2002) and based on ROSAT and ASCA measurements of the intracluster gas profile. We estimate errors in the sampled  $U$ -band luminosity with a Monte Carlo algorithm.

Beijersbergen et al. (2002) have determined a  $U$ -band GLF for the Coma Cluster from statistical background subtraction. Integrating their GLF to  $M_U = -19$  gives a cumulative luminosity of  $L_U \approx 369 \times 10^{39} \text{ erg s}^{-1} \lambda^{-1}$ . At a mass of  $M_{200} \approx 13.6 \times 10^{14} M_\odot$ , this luminosity is lower than suggested by the extrapolation from the three clusters in our sample, but on the same order of magnitude. Given various systematic differences between these two surveys — most notably, the different survey areas and the use of statistical background subtraction — we refrain from imposing this data point as an additional constraint on our  $L_U(M_{200})$  relation.

To derive an estimate for  $L_{U,\lambda}$  over a continuous range of cluster masses, we plot the total luminosities from Table 3.4 in Fig. 3.8 and fit an analytical expres-

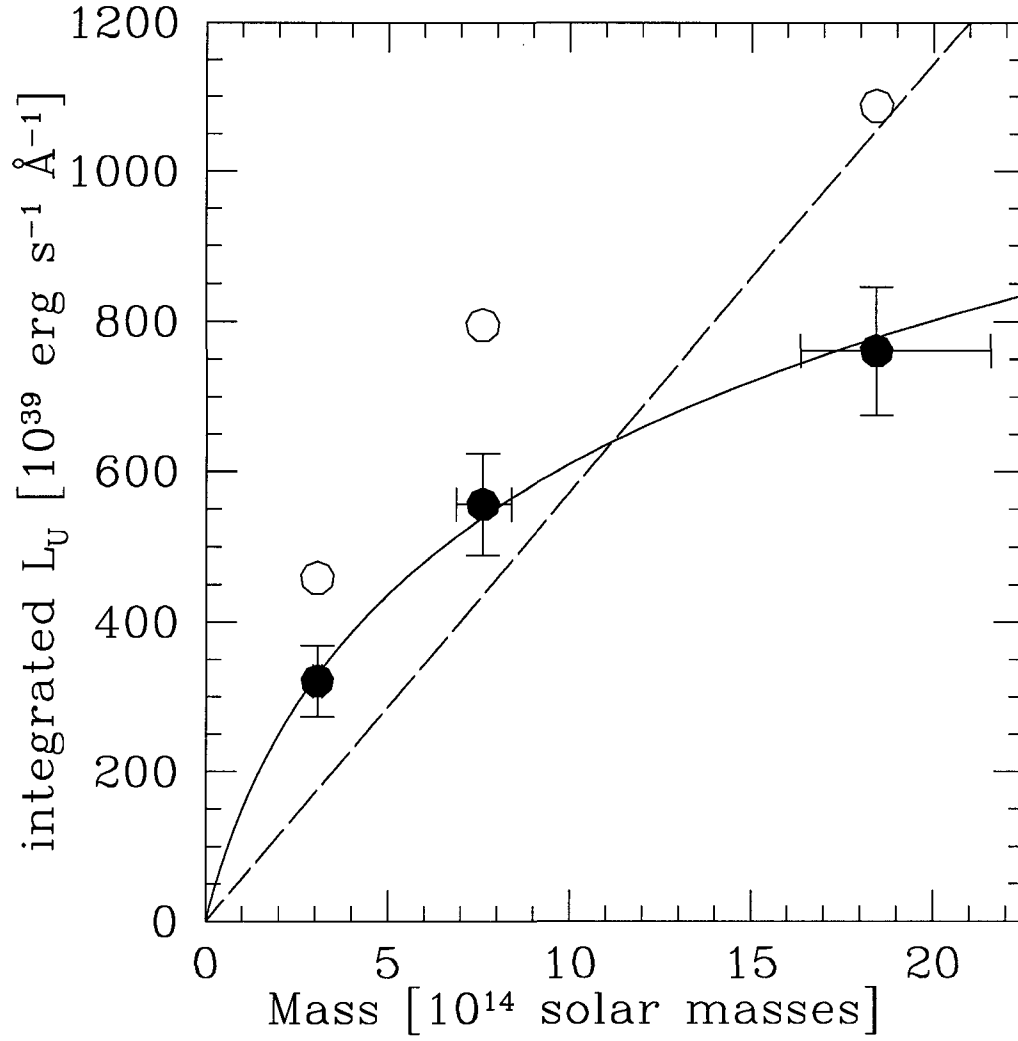


Figure 3.8: Integrated  $U$ -band luminosity as a function of cluster mass. Filled circles show contribution from the sampled magnitude range ( $M_U < -18.5$ ), empty circles show extrapolation to  $M_U < -10$  using the Schechter fit to the composite GLF. Solid and dashed lines are various analytical fits, which are discussed in §3.5.4.

Table 3.4. Integrated  $U$ -band Luminosities as a Function of Absolute Magnitude Limit

cluster	mass [ $10^{14} M_{\odot}$ ]	$L_U$ sampled	$L_U$ extrapolated	$L_U$ extrapolated from EL/NEL	$L_U^{EL}$ sampled	$L_U^{NEL}$ sampled
composite		$533^{+39}_{-39}$	$761^{+80}_{-70}$	$1048^{+1714}_{-322}$	$89^{+20}_{-20}$	$447^{+29}_{-29}$
A496	$3.06^{+0.13}_{-0.12}$	$321^{+48}_{-48}$	$458^{+76}_{-72}$	$706^{+1453}_{-311}$	$76^{+36}_{-36}$	$241^{+36}_{-36}$
A754	$18.44^{+3.13}_{-2.08}$	$761^{+85}_{-85}$	$1087^{+146}_{-135}$	$1271^{+1408}_{-301}$	$71^{+19}_{-19}$	$682^{+74}_{-74}$
A85	$7.61^{+0.79}_{-0.73}$	$556^{+68}_{-68}$	$795^{+113}_{-106}$	$1156^{+2085}_{-431}$	$109^{+44}_{-44}$	$454^{+54}_{-54}$

Notes: Luminosities in  $10^{39} \text{ erg s}^{-1}$ . Errors are based on  $1\sigma$  uncertainties in analytical fits. Sampled luminosities are for galaxies with  $M_U < -18.5$ , extrapolated  $M_U < -10$ . Masses are from Reiprich & Böhringer (2002).

sion to the three data points for the uncorrected cumulative luminosity. For similar analyses in the  $K$ -band, Kochanek et al. (2003) get  $L_K \propto M^{1.10 \pm 0.09}$ , consistent with a linear relation. Girardi et al. (2000) find a weak, but significant departure from linearity for the  $B_j$ -band. Their estimate is  $L_{B_j} \propto M^{0.8} - M^{0.9}$ . Therefore, previous work suggests both linear and curved relations. To take this uncertainty into account, we attempt to fit both to the data in Fig. 3.8, and provide below the range of results from both approaches.

A linear relation does not fit the three clusters in our sample well. The best fit slope in units of  $10^{24} \text{ erg s}^{-1} M_{\odot}^{-1}$  is  $57 \pm 0.7$ , with  $\chi^2 = 95$ .

For the curved relation, in order to avoid a singularity of the mass-to-light ratio at the origin, we do not fit a power law, but opt instead for the functional form

$$L_{U,\lambda} = a_1 \lg(a_2 M + 1). \quad (3.4)$$

This functional form is not theoretically motivated, but it reproduces the data points and provides a plausible interpolation between them, passes through the



origin, and is differentiable there. The fit is excellent. With  $a_2 = 0.63 \times 10^{-14} M_\odot^{-1}$ , the best-fit value, we find  $a_1 = 709_{-94}^{+88} * 10^{39} \text{ erg s}^{-1} \text{ \AA}^{-1}$ .

With these two fits, we obtain luminosity densities from clusters in the mass range from A496 ( $M_{200} \approx 3 \times 10^{14} M_\odot$ ) to A754 ( $M_{200} \approx 1.8 \times 10^{15} M_\odot$ ) from  $1.7$  to  $2.8 \times 10^{36} \text{ erg s}^{-1} \text{ \AA}^{-1} \text{ Mpc}^{-3}$ . The quoted range comprises the difference between the two analytical forms as well as the  $1\sigma$  errors associated with each. This luminosity density contributes to the  $U$ -band background intensity by  $\frac{dI_\lambda}{dt} = 4.3 - 6.4 \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ \AA}^{-1} \text{ Gyr}^{-1}$ , which corresponds to an increase of  $0.1 - 0.3\%$  relative to the current background per Gyr. The higher estimate is associated with the curved functional form.

Estimating the cumulative contribution of clusters to the  $U$ -band background would require models for the evolution of the cluster mass function, the star formation rate, and the spectral energy distribution of star forming galaxies at high redshifts. A detailed discussion exceeds the scope of this paper, and for this reason, we only give the rate of change in the  $U$ -band background due to present-day clusters here. However, even if star formation rates in clusters and their progenitors were higher by an order of magnitude in the past (approximately the difference between star formation in present-day field galaxies and the peak of star formation history around  $z \approx 1$  (Madau, Pozzetti & Dickinson, 1998; Steidel et al., 1999)), the total cumulative contribution of clusters at  $z \leq 3$  to today's  $U$ -band background would still be only a few percent.

Clusters do not contribute much to the mass density of the universe either. Integrating the mass function from Jenkins et al. (2001) shows that clusters in the mass range covered by our sample contribute  $\sim 2\%$  of the critical density, or  $\sim 8\%$  of the matter density of the universe. On the other hand, a significant fraction of the  $U$ -band background may be contributed by galaxies in virialized systems of

lower mass, by isolated field galaxies, or by non-stellar sources. Integrating the mass function above shows that groups and clusters with  $M_{200} > 10^{12} M_{\odot}$  account for a mass density of  $\Omega_m = 0.21$ , and additional contributions could come from isolated galaxies and groups with  $M_{200} < 10^{12} M_{\odot}$ .

To estimate the contribution from groups with  $M_{200} > 10^{12} M_{\odot}$ , we extrapolate the mass- $L_U$  relations found above and integrate the luminosity density between  $M_{200} = 10^{12} M_{\odot}$ , the lower bound of the mass of the Local Group, and a variable upper mass cutoff. Fig. 3.9 shows the resulting integrated luminosity density as a function of the upper integration limit. The total luminosity density from groups and clusters with  $M_{200} \geq 10^{12} M_{\odot}$  approaches  $15.0 - 53.7 \times 10^{36} \text{ erg s}^{-1} \text{ \AA}^{-1} \text{ Mpc}^{-3}$ . Groups of  $\sim 10^{12}$  to  $10^{13} M_{\odot}$  contribute most to this emission. The contribution to the  $U$ -band background from the entire mass range is  $\frac{dI_{\lambda}}{dt} = (42.7 - 138) \times 10^{-12} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1} \text{ \AA}^{-1} \text{ Gyr}^{-1}$ , corresponding to an increase of 1–6% of the current  $U$ -band background over 1 Gyr, depending on the functional form of the fit and the value of the current  $U$ -band background used.

### 3.6 Conclusions

We have calculated galaxy luminosity functions (GLFs) from total  $U$ -band magnitudes in three nearby clusters. Our analysis is based on a spectroscopic sample, providing cluster membership confirmation for each galaxy and avoiding the need to resort to statistical background subtraction.

We have introduced a new variant of the maximum likelihood method for calculating luminosity functions. Conventional maximum likelihood methods use binned or analytic distributions in just one or two variables (e.g., absolute magnitude and surface brightness) as an *ansatz* and require the algorithm to be customized and rerun every time the form of this *ansatz* is changed, making it in-

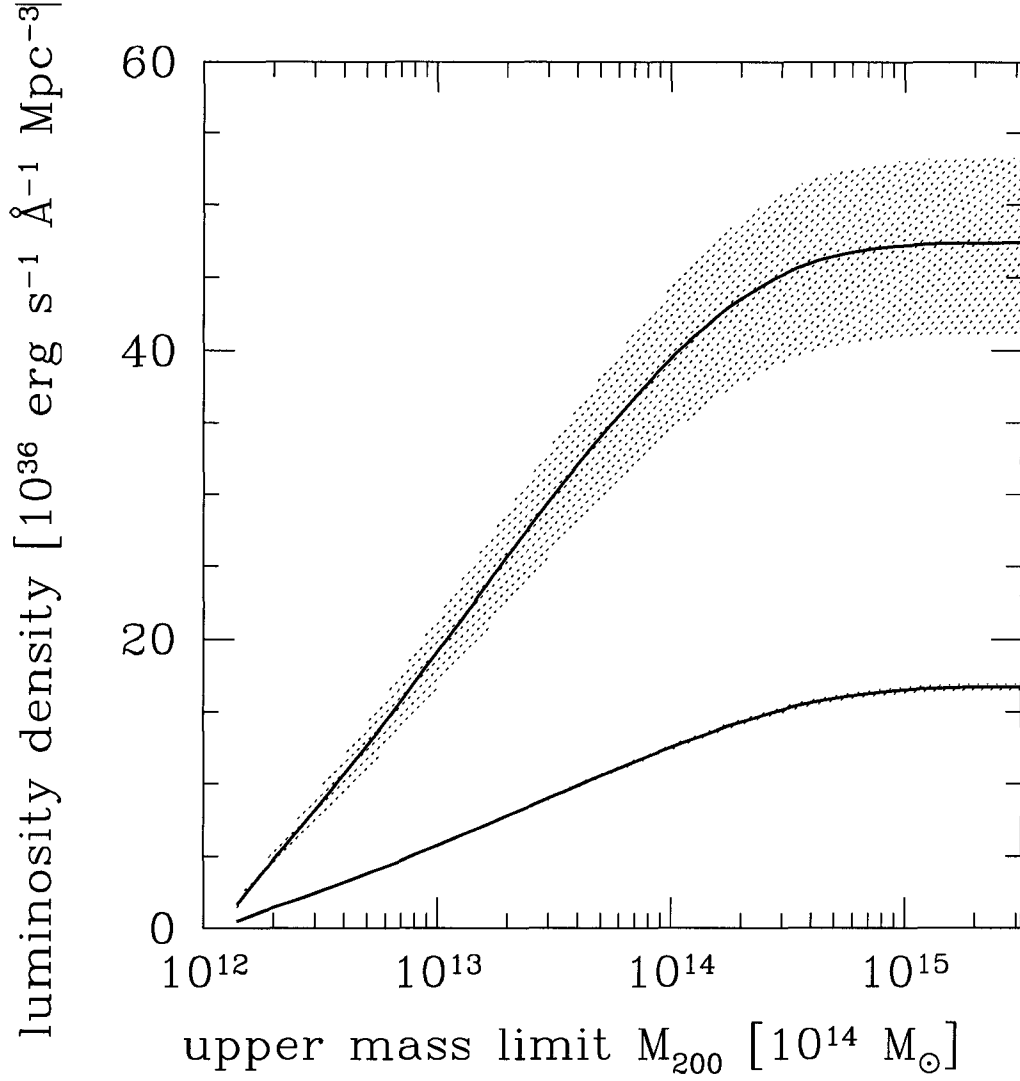


Figure 3.9: *U*-band luminosity density from groups and clusters as a function of the upper mass limit  $M_{200}$ . Lower mass limit for the integration is  $10^{12} M_{\odot}$ . Mass function is from Jenkins et al. (2001). The upper line corresponds to the extrapolation using a curved functional form, as discussed in the text. The lower line corresponds to a linear relation between  $M_{200}$  and  $L_U$ . Uncertainties (shaded regions) are based on the uncertainties in the normalization of  $L_U - M_{200}$  relation.

convenient to analyze galaxy properties in higher-dimensional parameter spaces (e.g., as a function of absolute magnitude in different filter bands, surface brightness, environment, morphology, star formation, etc., simultaneously). Our *Discrete Maximum Likelihood* method, on the other hand, does not assume a specific form or dimensionality for the *ansatz* a priori, but assigns a statistical weighting factor to each sampled galaxy. This weighting factor contains all completeness and volume corrections. Therefore, the full range of observables of each galaxy can be retained throughout the calculation and subsequently analyzed without having to customize and rerun the algorithm with a new *ansatz* for every new analysis. The DML method is therefore ideal for analyzing modern surveys that gather a large number of photometric, spectroscopic and morphological parameters about each object. Like other maximum likelihood estimators, the DML has the advantage of being unbiased by density inhomogeneities in the sample.

The results of our GLF analysis are summarized below:

- The  $U$ -band GLF in clusters down to  $M_U < -17.5$  ( $\approx M_U^* + 2$ ) can be described by a Schechter function with  $M_U^* = -19.81 \pm 0.27$ ,  $\alpha_U = -1.09 \pm 0.18$ , and  $\phi_U^* = 142 \pm 5$  galaxies cluster $^{-1}$  mag $^{-1}$ .
- We compare the  $U$ -band and  $R$ -band GLFs and find that, although the difference is too subtle to be reflected in the Schechter parameter  $\alpha$ , the ratio of faint ( $M > M^*$ ) to bright ( $M < M^*$ ) galaxies is significantly larger in the  $U$ -band. This indicates that cluster galaxies are bluer at fainter magnitudes.
- For quiescent galaxies, we find best fit Schechter parameters of  $M_U^* = -19.77^{+0.28}_{-0.30}$ ,  $\alpha_U = -0.97^{+0.22}_{-0.18}$ , and  $\phi_U^* = 133 \pm 6$ . The star forming GLF can be fit with a power law with a slope of  $\alpha = -2.16^{+0.16}_{-0.19}$ .
- If the Schechter fit to the overall GLF can be extrapolated to faint magni-

tudes ( $M_U \leq -10$ ), we have sampled  $\sim 85\%$  of the total  $U$ -band light from the clusters within the limit of  $M_U < -17.5$ . Quiescent galaxies dominate the  $U$ -band flux at  $M_U < -14$ . If there is a faint end upturn in the overall GLF (as suggested by extrapolating the star forming GLF) that continues past  $M_U = -14$ , we cannot rule out that dwarf emission line galaxies dominate the total  $U$ -band flux. The uncertainty is primarily in the shape of the star forming GLF.

- Rich clusters in the mass range of our sample ( $\sim 3 \times 10^{14}$  to  $1.8 \times 10^{15} M_\odot$ ) account for a  $U$ -band luminosity density of  $(1.71 \pm 0.02)$  to  $(2.46 \pm 0.33) \times 10^{36} \text{ erg s}^{-1} \text{ \AA}^{-1} \text{ Mpc}^{-3}$ . This corresponds to an increase of the current  $U$ -band background of  $0.1 - 0.3\%$  per Gyr. Additional contributions from galaxies outside our spatial survey limits could increase this estimate by a factor of 2-4. Galaxies beyond our faint magnitude limit could add another  $\sim 20\%$  if the overall GLF can be extrapolated to  $M_U \approx -10$ .

## CHAPTER 4

### CAN EARLY TYPE GALAXIES EVOLVE FROM FADING THE DISKS OF LATE TYPES?

#### 4.1 Chapter Summary

We examine whether early-type galaxies in clusters may have evolved from later types by the fading of their disks (e.g., as a result of ram-pressure stripping or strangulation) or by enhancement of the bulge luminosity (e.g., due to tidal interactions and mergers). For this purpose, we compare the bulge and disk luminosities of early- and late-type galaxies and of galaxies at different radial distances from the cluster center. We find that, in order for early-type galaxies, including S0s, to have evolved from late-type galaxies, their bulge luminosities must have been physically enhanced. Disk fading models cannot explain the differences observed. We then show that galaxy bulges are systematically brighter at small projected distances from the cluster center, while disk luminosities are uncorrelated with cluster-centric distance. Our results suggest that bulge enhancement, not disk fading, distinguishes early from late types and is thus at least partially responsible for the morphology-environment relation of bright cluster galaxies.

#### 4.2 Introduction

Galaxies have a wide range of morphologies, and the mechanisms that are responsible for generating this diversity are not completely understood. Elliptical galaxies may be the result of major mergers (Barnes & Hernquist, 1992). Spirals may represent a more pristine population undisturbed by major disruptive

events. The origin of intermediate types — particularly S0s — has been the subject of much debate.

One proposed scenario to explain the apparent increase in the fraction of S0s in rich clusters over time (Dressler et al., 1996) is that, as spirals enter dense environments, their reservoirs of neutral gas are cut off either by disrupting further accretion (i.e., by strangulation; Larson, Tinsley & Caldwell 1980; Balogh, Navarro & Morris 2000; Bekki, Couch & Shioya 2002), or by removing gas directly from the disk (i.e., via ram-pressure stripping; Gunn & Gott 1972). Star formation in the disk would then cease, and the disk would evolve passively and fade as high-mass stars burn out. Studies of clusters at higher redshift (Couch et al., 1998; Dressler et al., 1999; Poggianti et al., 1999) have uncovered a population of recently star forming ( $\sim 1\text{--}2$  Gyr earlier), but now quiescent galaxies with normal disk morphologies that those authors cite as possible examples of systems whose star formation was disrupted as described above.

Another possibility is that galaxy-galaxy interactions, including close tidal encounters and mergers, increase the luminosity of the bulge component by heating the central parts of the disk or triggering star formation in the center (Barnes, 1999; Bekki, 1998; Mihos & Hernquist, 1994). Extensive observations demonstrate that such interactions do increase central star formation rates (Lonsdale, Persson & Matthews, 1984; Kennicutt et al., 1987; Liu & Kennicutt, 1995) and can produce bulge-enhanced merger remnants (Yang et al., 2004). These interactions are likely in clusters with substructure (i.e., poor groups accreted from the field), where the relative speeds and internal velocity dispersions of subcluster members are similar.

Both scenarios increase the bulge fraction,  $B/T$ , the fraction of the total luminosity of a galaxy associated with the bulge. The first scenario increases  $B/T$

by reducing the luminosity contribution from the disk. The second scenario increases  $B/T$  by increasing the luminosity of the bulge. Could either mechanism be responsible for generating the morphological sequence in clusters?

Past observational work has attempted to address this question by comparing bulge fractions or bulge luminosities to Hubble type and/or environment (Dressler, 1980; Boroson, 1981; Simien & de Vaucouleurs, 1986; Solanes, Salvador-Solé & Sanromà, 1989), with often conflicting conclusions. Dressler (1980) finds that the bulges of S0s are more luminous than those of spirals, and that their luminosity is increased in denser environments. In contrast, Solanes, Salvador-Solé & Sanromà (1989), working from the same sample, argue that disk luminosities are decreased in dense environments. Boroson (1981) finds bulge fractions of S0s to be similar to those of spirals above a certain  $B/T$ , while Simien & de Vaucouleurs (1986) find S0s to have systematically larger bulge fractions.

Several problems affect these studies: 1) While Dressler (1980) and Solanes, Salvador-Solé & Sanromà (1989) consider a large, apparent magnitude-limited sample of cluster galaxies, they are not able to spectroscopically confirm cluster members, which raises the possibility that correlations between environment and galaxy structural properties could arise from background contamination, especially at large radii. 2) Hubble types are generally used as morphological quantifiers, a measure that is subjective and not easily reproducible. 3) The bulge-disk decompositions in Dressler (1980) (also used by Solanes et al.) are done visually and thus likely to be affected by large errors and incompleteness arising from unresolved bulges and disks. 4) None of the studies above quantify how their morphological classifications may be influenced by the bulge fraction. This bias, whether conscious or otherwise, may dramatically affect the interpretation of those results, because it could introduce intrinsic correlations between bulge and



disk luminosities and morphological type. 5) The impact of bulge enhancement and disk fading is on the luminosity distributions of the bulge and disk components, and the luminosity functions of bulges and disks are thus the best discriminators between the two evolutionary hypotheses outlined above, but only Solanes, Salvador-Solé & Sanromà (1989) determine a luminosity function. This determination is semi-empirical and not calculated directly from the observed luminosity distribution of bulges. More recently, Benson, Frenk & Sharples (2002) have calculated bulge and disk luminosity functions directly, but their sample is too small to hold any discriminatory power over the evolutionary mechanisms described above.

In this paper, we compare the luminosity functions for bulges and disks, calculated directly from a sample of several hundred spectroscopically-confirmed cluster galaxies (Christlein & Zabludoff, 2003) to determine whether early type galaxies could have evolved from late type galaxies by disk fading alone. We use the bulge fraction,  $B/T$ , rather than Hubble type, as a quantitative, reproducible measure of morphology, which enables us to account for any bias that could introduce intrinsic correlations between bulge luminosity and morphology. By examining bulge and disk luminosities as a function of projected radial distance from the cluster center, we test whether the morphology-environment relation (Dressler, 1980) could have been generated principally by bulge-enhancing or disk-diminishing effects.

### 4.3 The Data

Our sample consists of cluster galaxies from a spectroscopic and  $R$ -band imaging survey of six nearby clusters (A85, A496, A754, A1060, A1631, A3266). The spectroscopic survey ensures that contamination of the sample by background field

Table 4.1. The Cluster Sample

Cluster	#	$\overline{cz}$	$\Delta m$	$cz$ range	$\sigma$	$r_{sampling}$
	Galaxies	[km/s]	[mag]	[km/s]	[km/s]	[Mpc]
A1060	252	$3683 \pm 46$	33.59	2292 - 5723	$724 \pm 31$	0.67
A496	241	$9910 \pm 48$	35.78	7731 - 11728	$728 \pm 36$	1.76
A1631	340	$13844 \pm 39$	36.53	12179 - 15909	$708 \pm 28$	2.42
A754	415	$16369 \pm 47$	36.90	13362 - 18942	$953 \pm 40$	2.83
A85	280	$16607 \pm 60$	36.94	13423 - 19737	$993 \pm 53$	2.87
A3266	331	$17857 \pm 69$	37.10	14129 - 21460	$1255 \pm 58$	3.07

Note. — # is the number of sampled galaxies per cluster.  $\overline{cz}$  is the mean velocity,  $\Delta m$  the distance modulus (for  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ). “ $cz$  range” is the velocity range spanned by cluster members,  $\sigma$  is the line-of-sight velocity dispersion, and  $r_{sampling}$  is the projected physical sampling radius (center-to-edge).

galaxies is minimized. Table 4.1 lists the kinematic properties of these six clusters for  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ , as applied throughout this paper. For details regarding the survey and data reduction, see Christlein & Zabludoff (2003).

#### 4.4 Bulge-Disk Decomposition

We use the GIM2D software (Simard et al., 2002) to perform a two-dimensional decomposition of the galaxy images into a bulge component, described by a de Vaucouleurs surface brightness profile (de Vaucouleurs, 1948), and a disk component with an exponential surface brightness profile. Prior to the fit, we transform all galaxy images to a fiducial rest frame  $cz = 17858 \text{ km/s}$  (corresponding to mean the velocity of the most distant cluster, A3266) by fading the surface brightnesses by  $(z_{cosmological} + 1)^{-2}(z_{total} + 1)^{-2}$ , smearing the images to achieve a consistent

FWHM of 2 arcsec, and rebinning their pixels with the new angular diameter distance to achieve the same physical resolution per pixel. This approach ensures that determinations of  $B/T$  are internally consistent among the clusters, which span a mean velocity range from 3682 to 17858 km/s. The final catalog contains bulge-disk decompositions of 1637 galaxies (1304 of them for galaxies with  $M_R \leq -19.2$ ). The full catalog, including  $B/T$  values, appears in a subsequent paper (Christlein & Zabludoff 2004, in prep.).

Fig. 4.1 shows the distribution of  $B/T$  for galaxies that are identified in the NASA Extragalactic Database (NED) as S, S0, or E, as well as the  $B/T$  distribution for all sample galaxies with and without literature classifications. Because of the inhomogeneity of the literature sources, this figure should only be regarded as an approximate orientation. The distribution of  $B/T$  in our sample is roughly comparable to those of other studies such as Tran et al. (2001), although we find a lower fraction of extreme late-type galaxies ( $B/T < 0.1$ ) and a higher fraction of intermediate-type galaxies ( $B/T \approx 0.3$ , typical of early-type spirals) than their sample, which consists only of field and group galaxies. Spirals are mostly confined to small  $B/T$ . There are very few pure bulge systems in our sample, and we find significant disk components even in galaxies classified as ellipticals (Rix & White, 1990). Such low-surface brightness features are typically more difficult to identify on photographic plates, which may explain why such systems have been classified as ellipticals. The fact that our classifications are in good agreement with galaxies typed as spirals in the literature supports this hypothesis.

To test our quantitative morphological classifications, we visually examine all literature-classified “ellipticals” with the lowest  $B/T$  values ( $B/T < 0.4$ ). For the five most luminous among them, the reason for the low  $B/T$  is usually apparent (brightest cluster galaxies with extended envelopes or interacting systems with

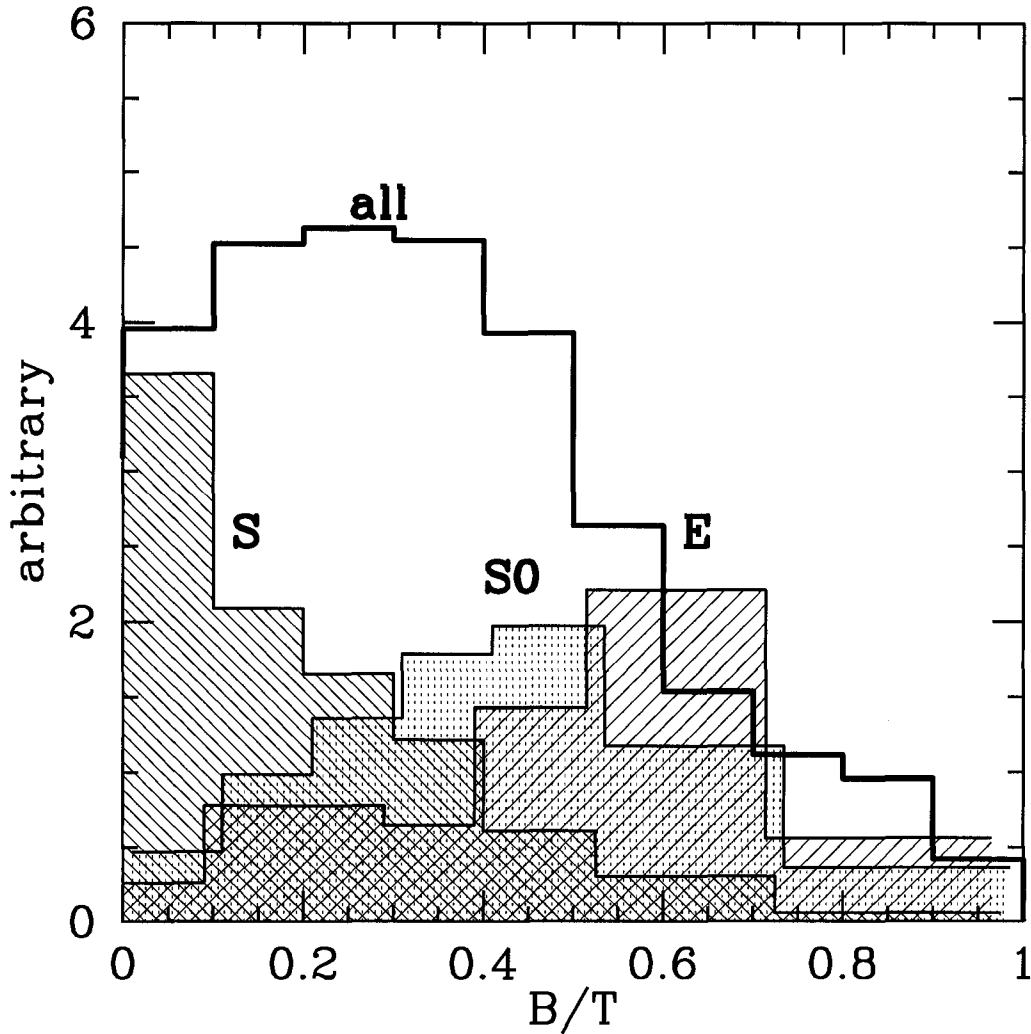


Figure 4.1: Distribution of measured  $B/T$  for galaxies identified in the NASA Extragalactic Database (NED) as spirals (S), S0s, and ellipticals (E) in units such that the area under each curve is unity. The bold, unshaded histogram shows the  $B/T$  distribution for all galaxies in the sample, regardless of whether Hubble types are available for them in NED (scaled by a factor of 3.5 for better legibility). For display purposes, the histograms have been slightly displaced along the  $B/T$  axis. Spirals have low  $B/T$ , with a median  $B/T < 0.2$ . S0s and Es have very broad distributions with medians  $B/T \approx 0.45$  and  $B/T \approx 0.6$ , respectively. Many galaxies identified in the literature as ellipticals have significant disk components.

extended features that could be disks or tidal debris). Although these “disk” features may be of a different nature than normal spiral disks, they are reproduced well by GIM2D, and so we do not discard them in order to avoid introducing subjective selection criteria. Most low- $B/T$  “ellipticals” are fainter ( $M_R > -21$ ), and GIM2D also fits these systems very well. Attempts to fit these systems with pure bulge models typically result in an increase both in the reduced  $\chi^2_\nu$  and in the fraction of residual light after the model has been subtracted. Overall,  $B/T$  is well correlated with Hubble type for our sample (the Spearman rank correlation coefficient is  $r = 0.57$  at  $12\sigma$  deviation from the no-correlation hypothesis) and therefore suitable for our analysis.

#### 4.5 Calculating the Bulge and Disk Luminosity Functions

To calculate the distribution functions of bulge and disk luminosities, we need to consider not only bulge and disk luminosities, but total luminosity and surface brightness as well, because these quantities determine the completeness of the spectroscopic and morphological catalogs. This multi-variate problem is best solved with the Discrete Maximum Likelihood (DML) algorithm (Christlein, McIntosh & Zabludoff, 2004), which does not require any prior assumptions regarding the analytical form or dimensionality of the galaxy distribution function. This algorithm calculates statistical weighting factors that incorporate the completeness corrections and volume corrections for each galaxy. These corrections account statistically for the fact that the spectroscopic and morphological catalogs are not complete and that a galaxy of a given absolute magnitude may be observable in some, but not all the clusters in our sample.

The spectroscopic catalog is the major source of incompleteness; depending on surface brightness and cluster, the completeness of the catalog for  $m_R = 18$

galaxies is in the range  $\sim 0.25$  to  $\sim 0.5$ . The morphological catalog is mostly complete; the fraction of galaxies in the spectroscopic catalog that have bulge-disk decompositions is never smaller than  $\sim 85\%$  for any  $m_R < 18$ .

We use the weighting factors calculated by the DML to reconstruct the bulge and disk luminosity functions (LFs, §4.6.1) and to calculate correlation coefficients for the bulge and disk luminosity-environment relations (§4.6.2). Even for the faintest galaxies that we analyze in this paper ( $M_R = -19.2$ ), the final weighting factor (incorporating completeness and volume corrections as determined by the maximum likelihood method) is never larger than  $\sim 2.5\times$ .

We propagate uncertainties in  $B/T$  into the final bulge and disk luminosity functions with a Monte Carlo algorithm. For each of 100 realizations, the algorithm draws a value of  $B/T$  from the uncertainty interval for each galaxy. We then determine the mean and scatter of the value of the luminosity function in each magnitude bin due to these  $B/T$  uncertainties and add the scatter in quadrature to the Poisson errors. This is a conservative procedure, because the  $B/T$  uncertainty intervals are given at the 99% confidence level by GIM2D, and because errors in different magnitude bins due to  $B/T$  uncertainties are really correlated for any given realization (i.e., if a galaxy is presumed to be in one bin for a given choice of  $B/T$ , it cannot lie in any other bin).

## 4.6 Results and Discussion

### 4.6.1 Bulge- and Disk Luminosity as a Function of Morphology

To isolate the mechanism by which early-type galaxies may evolve from late types, we compare the bulge and disk luminosities of galaxies with different morphologies. We split the sample into six independent subsamples, selected by  $B/T$ . Table 4.2 shows the  $B/T$  intervals for each sample, the median  $B/T$ , and

Table 4.2. Morphological Subsamples and Schechter Fits

Subsample	B/T <sub>min</sub>	B/T <sub>max</sub>	B/T <sub>med</sub>	# Galaxies	% S	% S0	% E	$M_{bulge}^*$ ( $\alpha = -1$ )	$M_{disk}^*$ ( $\alpha = -1$ )
A	0.0	0.19	0.05	498 <sup>+31</sup> <sub>-4</sub>	57.4	14.6	10.4	-19.32 <sup>+0.39</sup> <sub>-0.37</sub>	-20.74 <sup>+0.12</sup> <sub>-0.11</sub>
B	0.2	0.29	0.25	247 <sup>+6</sup> <sub>-28</sub>	16.5	13.6	7.8	-19.83 <sup>+0.23</sup> <sub>-0.21</sub>	-20.99 <sup>+0.18</sup> <sub>-0.16</sub>
C	0.3	0.39	0.35	239 <sup>+11</sup> <sub>-11</sub>	12.2	17.8	6.5	-20.04 <sup>+0.22</sup> <sub>-0.22</sub>	-20.84 <sup>+0.17</sup> <sub>-0.17</sub>
D	0.4	0.49	0.45	239 <sup>+5</sup> <sub>-21</sub>	6.1	19.7	14.3	-20.91 <sup>+0.21</sup> <sub>-0.21</sub>	-21.13 <sup>+0.16</sup> <sub>-0.14</sub>
E	0.5	0.69	0.59	310 <sup>+19</sup> <sub>-22</sub>	6.1	23.5	44.1	-21.62 <sup>+0.13</sup> <sub>-0.12</sub>	-21.21 <sup>+0.13</sup> <sub>-0.12</sub>
F	0.7	1.0	0.91	185 <sup>+36</sup> <sub>-0</sub>	1.7	10.8	16.9	-21.23 <sup>+0.14</sup> <sub>-0.13</sub>	-20.00 <sup>+0.39</sup> <sub>-0.35</sub>

the number of galaxies in each bin. The next three columns show the percentages of all S, S0, and E-type galaxies that fall into each bin, based on literature classifications from NED. Because of the inhomogeneity of the literature sources and the problems discussed earlier that are associated with classifications based on photographic plates, these percentages should only be taken as a rough calibration of the distribution of galaxies in  $B/T$ .

Our analysis does not extend to deep magnitudes: in all cases,  $M_R < -19.2 + 5 \lg h$ , the absolute magnitude limit of the spectroscopic and morphological catalogs for the most distant cluster, A3266. We therefore focus our work on the bright end of the luminosity function and characterize it by fitting Schechter functions (Schechter, 1976) to the bulge and disk luminosity functions in each subsample. We impose the constraint that the faint end slope is flat ( $\alpha = -1$ ). By fixing  $\alpha$  at a constant value, the Schechter parameter  $M^*$  becomes a direct measure of the characteristic magnitude of the bright end exponential cutoff. We refer to the  $M_R^*$  obtained with this slope constraint as  $M_R^*(\alpha = -1)$ .

Figs. 4.2 and 4.3 show all twelve LFs (six bulge LFs and six disk LFs) and the

Schechter fits with  $\alpha = -1$ . With the exception of the disk LF for  $0.5 \leq B/T < 0.7$ , all fits are consistent with the LFs within  $2\sigma$ . Fig. 4.2 shows that the bright ends of the disk LFs are similar, except for the earliest type galaxies ( $B/T \leq 0.7$ ), whose LF is considerably fainter. In contrast to the disk LFs, the bright ends of the bulge LFs (Fig. 4.3) spread out over  $\sim 2$  magnitudes from the latest- to the earliest-type subsample.

The last two columns of Table 4.2 list the values of  $M_R^*(\alpha = -1)$  for bulges and disks in all six subsamples, and Fig. 4.4 shows this information as a function of  $B/T$ .

In Fig. 4.4,  $M_{bulge}^*(\alpha = -1)$  is negatively correlated with  $B/T$  over a wide range of  $B/T$ .  $M_{disk}^*(\alpha = -1)$  shows a weaker, but also negative, correlation with morphology. Note that the two slopes are not statistically independent, because, for any given  $B/T$ ,  $M_{bulge} - M_{disk}$  is fixed. Therefore, the only free parameter that places constraints on the evolution mechanisms of these galaxies is the absolute value of either slope.

Plotting a quantity  $M_{disk}^*(\alpha = -1)$  or  $M_{bulge}^*(\alpha = -1)$  that is explicitly dependent on  $B/T$  versus  $B/T$  introduces a bias, which we refer to as the “B/T bias”. In selecting high- $B/T$  galaxies to determine  $M_{bulge}^*(\alpha = -1)$ , we are typically selecting galaxies with brighter bulges, but also galaxies with fainter disks, which creates an intrinsic correlation between  $B/T$  and the luminosities of bulges and disks at a given  $B/T$ . This effect depends on parameters such as the bin size and the sample’s  $B/T$  distribution, and cannot be quantified analytically. To account for this bias and determine if the trends in Fig. 4.4 favor disk fading or bulge enhancement, we use a Monte Carlo analysis to calculate the range of expected slopes of the  $M_{bulge}^*(\alpha = -1)$ - $B/T$  and  $M_{disk}^*(\alpha = -1)$ - $B/T$  relations under the following null hypotheses: 1) B/T bias is solely responsible for the observed trends;



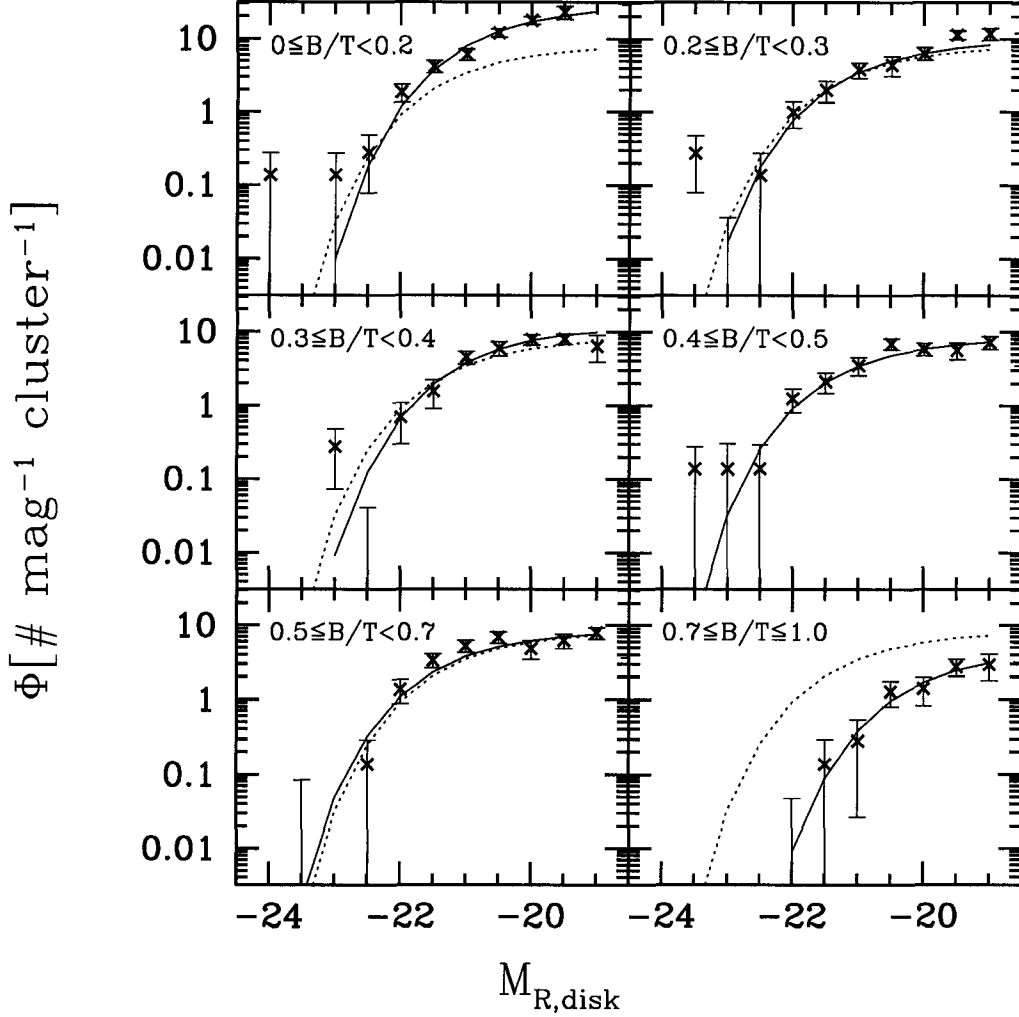


Figure 4.2: Luminosity functions and Schechter fits (with  $\alpha = -1$ ) for disks for six subsamples selected by their bulge fraction,  $B/T$ . See text and Table 4.2 for the definition of the subsamples. For orientation, the Schechter fit for an intermediate subsample ( $0.4 \leq B/T < 0.5$ ) has been marked in each panel. The bright end of the disk LF shows little variation along the morphological sequence (with the exception of the disk LF of the earliest-type galaxies, which is significantly fainter).

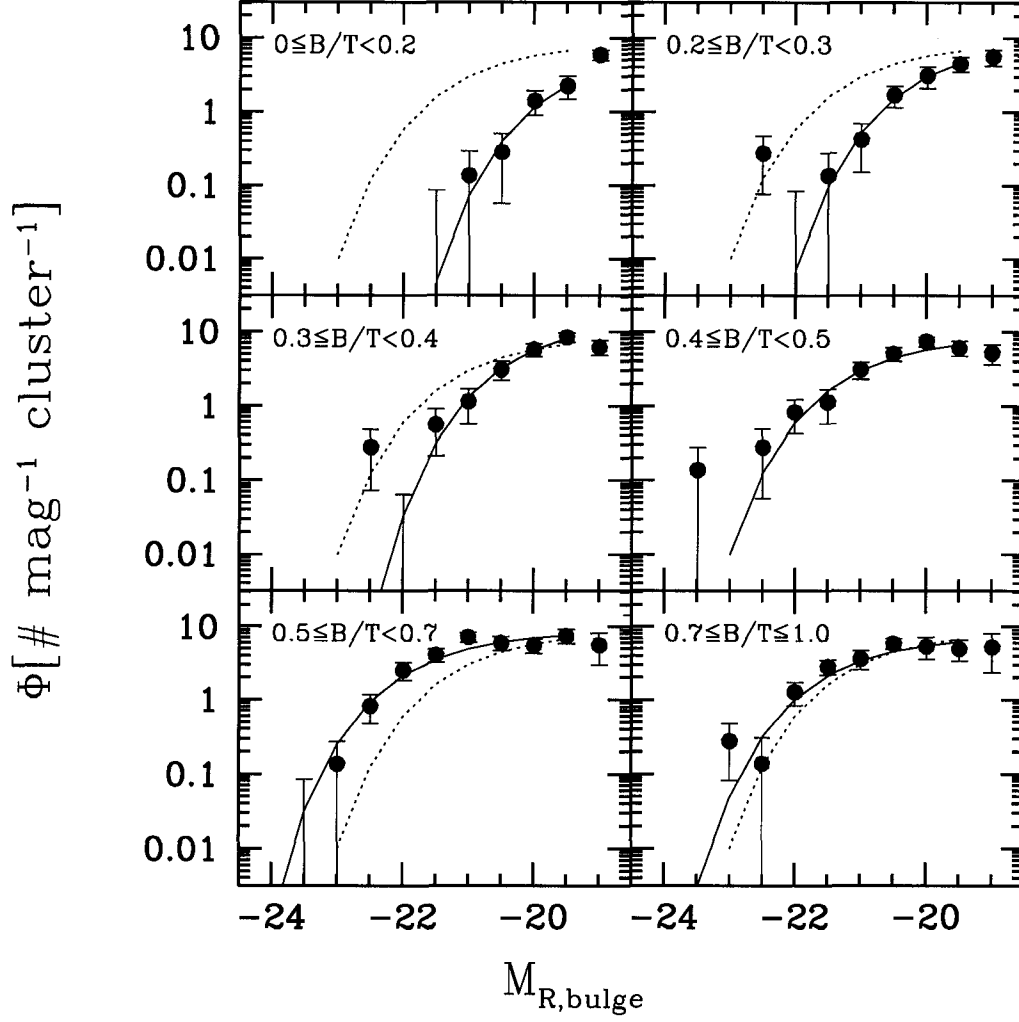


Figure 4.3: Luminosity functions and Schechter fits (with  $\alpha = -1$ ) for bulges for six subsamples selected by their bulge fraction,  $B/T$ . See text and Table 4.2 for the definition of the subsamples. For orientation, the Schechter fit for an intermediate subsample ( $0.4 \leq B/T < 0.5$ ) has been marked in each panel. From the late to early types, the bright end of the bulge LF grows more luminous by  $\sim 2$  mag.

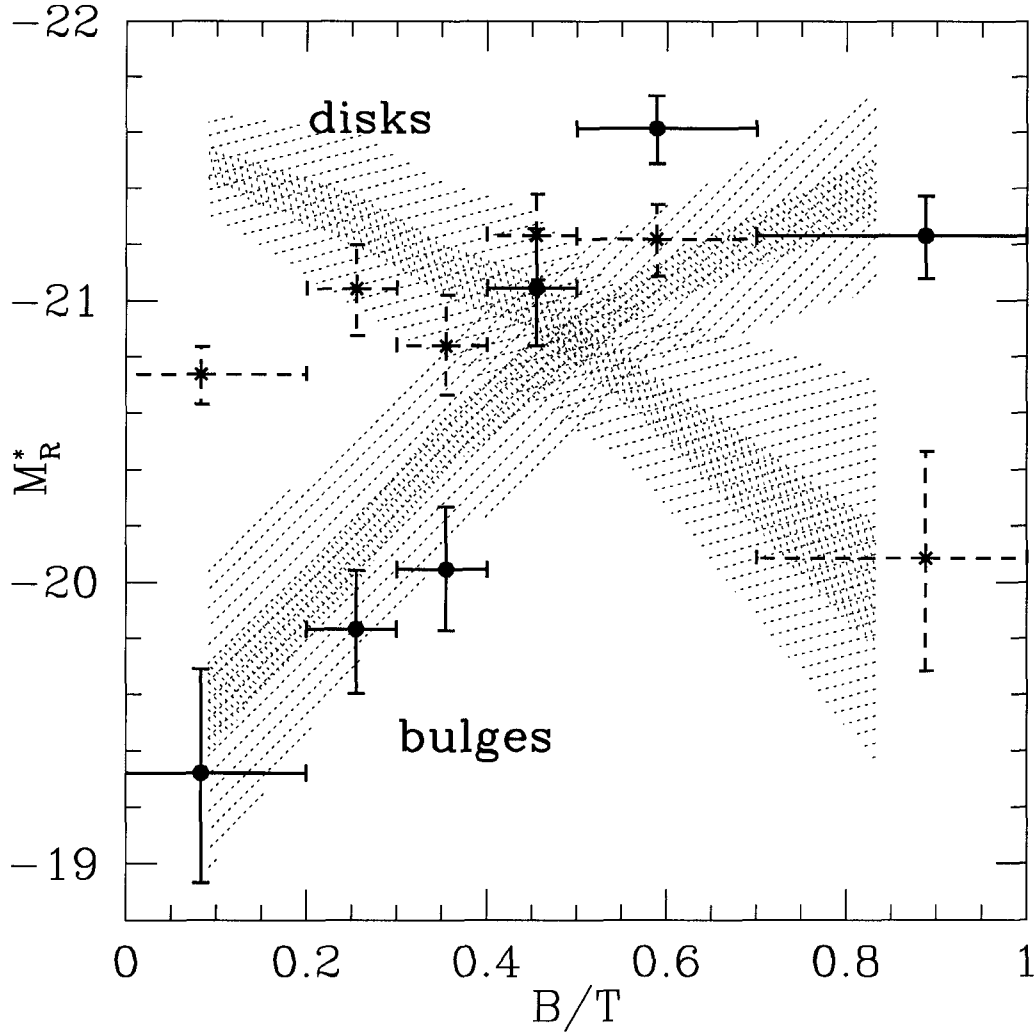


Figure 4.4: Characteristic bright magnitude,  $M_R^*(\alpha = -1)$ , as a function of  $B/T$  for bulge (circles) and disk luminosity functions (crosses). Error bars in  $M_R^*$  show  $1\sigma$  uncertainties from the Schechter fits. Error bars in  $B/T$  indicate bin width. Shaded areas show the  $1\sigma$  and  $2\sigma$  uncertainty intervals from the “B/T bias” Monte Carlo sample discussed in the text. Over most of the morphological sequence, the correlations are more negative than expected, favoring bulge enhancement over disk fading as the process that may transform late-type into early-type galaxies.

2) disk fading occurs in addition to the effects of B/T bias; 3) bulge enhancement operates in addition to the bias. We construct the Monte Carlo samples for these null hypotheses in the following ways:

(a) “B/T Bias”: Can B/T bias alone explain the observed trends? In this case, the observed trends in  $M_{disk}^*$  and  $M_{bulge}^*$  with  $B/T$  should arise even if the  $B/T$  values are uncorrelated with the total absolute magnitude  $M_T$ . The most straightforward test of this hypothesis is therefore to scramble the observed  $B/T$  values with respect to  $M_T$ , using them to calculate a new  $M_{bulge}$  and  $M_{disk}$  for each galaxy. B/T bias is then the only source of correlations between  $M_{bulge}$  and  $B/T$  and between  $M_{disk}$  and  $B/T$ .

We construct the Monte Carlo sample by associating random values of  $M_T$  from the completeness-corrected galaxy luminosity function with values of  $B/T$  drawn from the non-completeness-corrected  $B/T$  distribution of the sample. This ensures that the progenitor population is drawn from a realistic luminosity function, but that the  $B/T$  distribution of the Monte Carlo sample is the same as in the observed sample.

For each Monte Carlo sample (of several thousand), we then calculate LFs and Schechter functions for the bulge and disk components. We compare the trends of  $M_{disk}^*(\alpha = -1)$  and  $M_{bulge}^*(\alpha = -1)$  with  $B/T$  that we recover from the Monte Carlo samples to the observed trends to decide whether B/T bias can account for the latter.

The shaded regions in Fig. 4.4 show the range of  $M_{disk}^*$  and  $M_{bulge}^*$  recovered from this null hypothesis in each  $B/T$  bin. Even from visual inspection, it is clear that the observed slopes are more negative than predicted by the B/T bias hypothesis.

(b) “Disk Fading”: With this null hypothesis, we examine whether a disk fading

ing mechanism, in combination with B/T bias, can reproduce the observed trends in  $M_{disk}^*$  and  $M_{bulge}^*$  with  $B/T$ . Under the Disk Fading null hypothesis, any galaxy with a  $B/T$  characteristic of early-type galaxies has been generated from a progenitor population of low- $B/T$  galaxies by a reduction in the disk luminosity.

We incorporate the disk fading mechanism into the Monte Carlo catalog in the following way: we begin by scrambling all  $B/T$  values in the catalog with respect to  $M_T$ , as we did for the “B/T Bias” hypothesis. Then, for each mock galaxy with  $B/T > 0.3$ , we calculate the amount of disk fading necessary to generate it from a progenitor with  $B/T = 0.3$ , and apply it to the disk and total luminosity of the galaxy. As we did for the “B/T bias” hypothesis, we then calculate  $M_{disk}^*$  and  $M_{bulge}^*$  again in each B/T bin.

In the absence of any bias, this hypothesis would generate a positive correlation between  $M_{disk}^*$  and  $B/T$  (i.e., galaxies with larger  $B/T$  have fainter disks), while it would not introduce a correlation between  $M_{bulge}^*$  and  $B/T$ .

Our choice of  $B/T \approx 0.3$  for the hypothetical progenitor population is a conservative one, because 1) 73% of spirals have  $B/T < 0.3$ , and 2) the peak of the spiral distribution is at  $B/T < 0.2$ , while the peak of the S0 distribution is at  $B/T \approx 0.45$ . If the progenitors of today’s cluster galaxies were high-redshift field galaxies, it is likely that they had even lower  $B/T$  values than today’s population of cluster spirals.

(c) “Bulge Enhancement”: Under this hypothesis, galaxies with high  $B/T$  are generated from low- $B/T$  galaxies by an increase in the bulge luminosity. We construct the Monte Carlo samples for this hypothesis analogously to the “Disk Fading” hypothesis, except that, instead of fading the disks of our Monte Carlo sample members to adjust  $B/T$ , we increase the luminosity of the bulges.

Without B/T bias, this effect would lead us to expect a negative correlation

between  $M_{bulge}^*$  and  $B/T$ , and no correlation between  $M_{disk}^*$  and  $B/T$ .<sup>1</sup>

We compare the Monte Carlo samples generated by each of the three null hypotheses to the observations by fitting regression lines to the  $M_{bulge}^*(\alpha = -1)$ - $x$  and  $M_{disk}^*(\alpha = -1)$ - $x$  relations both in the observed and Monte Carlo samples, where  $x = \ln[(B/T)/(1 - B/T)]$ . The coordinate transformation from  $B/T$  to  $x$  ensures that the difference between the slopes of the relations does not depend on  $B/T$ , but it does not change the sense of the correlation. We perform the linear regression over a variety of ranges in  $B/T$  and compare models and observations separately for each of them. This allows us to test whether a given null hypothesis might still describe a viable transformation mechanism over a smaller range of morphologies (e.g., from spirals to S0s), even if it cannot explain the observed trends over the full range  $0 \leq B/T \leq 1$ .

Fig. 4.5 shows the results of these comparisons for all three null hypotheses over two different ranges in  $B/T$ . Plotted as contour lines in  $(\partial M_{bulge}^*(\alpha = -1)/\partial x, \partial M_{disk}^*(\alpha = -1)/\partial x)$  space is the distribution of regression slopes recovered from the null hypotheses. The contour lines in each panel encompass the results from 95% and 68%, respectively, of the Monte Carlo runs for a given null hypothesis. The observed slopes are marked by a data point in each panel. If the observed slopes fall outside the 95% contour line, we consider the observations inconsistent with the null hypothesis for that panel.

We find the following results:

- (1) None of our three null hypotheses alone, including B/T bias, can explain

---

<sup>1</sup>The predictions of these three null hypotheses could also be described in terms of a relation between total galaxy luminosity and  $B/T$  (uncorrelated for the “B/T bias” hypothesis, positive for the “Bulge Enhancement” hypothesis, and negative for the “Disk Fading” hypothesis). This approach is mathematically equivalent to the one we pursue here. We choose our approach because disk fading and bulge enhancement act independently on disk and bulge luminosities, so it is more straightforward to use the correlations of these quantities, rather than total luminosity, with  $B/T$  to discriminate among the three null hypotheses.

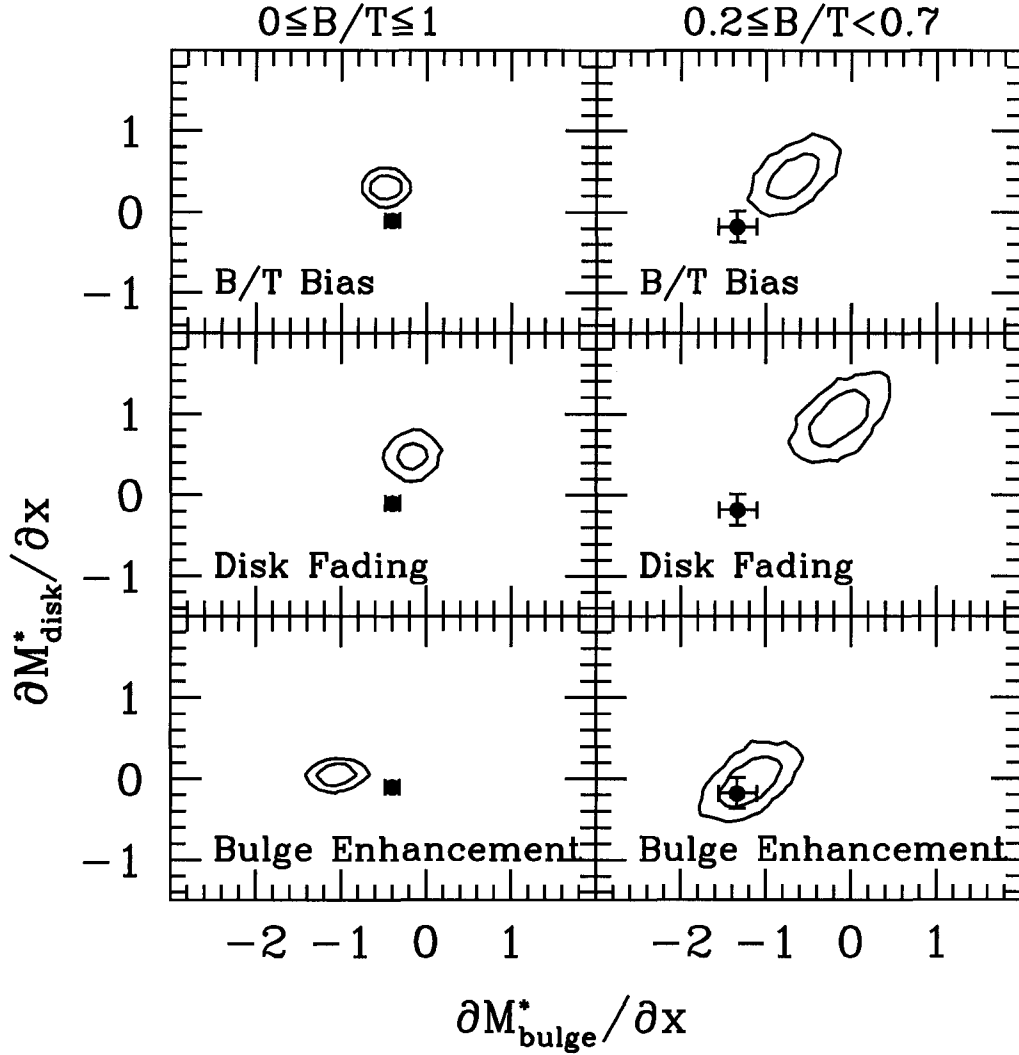


Figure 4.5: Observed slopes for the  $M_{bulge}^*(\alpha = -1)$ - $x$  and  $M_{disk}^*(\alpha = -1)$ - $x$  correlations (filled circles), compared with 1- and  $2\sigma$  error contours from the three null hypotheses. The quantity  $x$  is a parametrization of  $B/T$ . Left panels show fits over the full range of  $0 \leq B/T \leq 1$ , right panels show fits over the range  $0.2 \leq B/T < 0.7$ , which extends from early-type spirals to S0s and ellipticals. No single null hypothesis can explain the entire morphological sequence. Over the range  $0.2 \leq B/T < 0.7$ , bulge enhancement is clearly favored.

the observed  $M^*(\alpha = -1)$ - $x$  correlations over the entire morphological sequence. The left panels of Fig. 4.5 show the observed slopes relative to the distribution expected from each of the three null hypotheses. Excluding the  $B/T < 0.2$  subsample (which is not affected by the disk fading or bulge enhancement prescriptions anyway because it lies below the progenitor cutoff value of  $B/T = 0.3$ ) does not alter this conclusion.

(2) The “B/T Bias” and “Disk Fading” hypotheses are clearly ruled out over the range  $0.2 \leq B/T < 0.7$ , which excludes the most extreme late-type spirals and early-type ellipticals. In contrast, the “Bulge Enhancement” hypothesis is in excellent agreement with the observed correlations. The right panels of Fig. 4.5 show the observed slopes over this range, compared to the expected distributions.

(3) If we tighten the range further to  $0.2 \leq B/T < 0.5$ , the “Disk Fading” hypothesis is still ruled out. This range of  $B/T$  covers the range between the peaks of the distribution of spirals and S0s, and thus provides the most specific constraints on the possible morphological transformation mechanisms affecting this sample.

Thus, our Monte Carlo approach does more than account for  $B/T$  bias and establish the presence of a physical effect — it also demonstrates that even a very simplistic bulge enhancement model reproduces not only the sense, but also the magnitude of the observed  $M_{disk}^*-B/T$  and  $M_{bulge}^*-B/T$  relations over all  $B/T$ .

We emphasize that our analysis does not assume that early types evolve from a progenitor population like the late types in our sample. We only assume that both early and late types evolved from the same progenitor population in the past, prior to any subsequent environmental transformations. Relaxing this assumption might make it possible to reproduce the observed  $M_{disk}^*-B/T$  and  $M_{bulge}^*-B/T$  relations.



$B/T$  relations over a limited  $B/T$  range, but it is hard to imagine a model that would generate the observed correlations over all  $B/T$ .

#### 4.6.2 Bulge- and Disk Luminosity as a Function of Environment

The results of the previous section suggest that the morphologies of cluster galaxies are differentiated by bulge enhancement, not by disk fading. Is the morphology-environment relation in clusters (Dressler, 1980) thus due mostly to changes in bulge, not disk luminosity? In other words, is there evidence that the bulges of galaxies grow more luminous towards the denser centers of clusters? If disk fading does occur as a galaxy falls through a cluster, we should also see differences in disk luminosity with projected radial distance from the cluster center.

We test whether the morphology-environment relation is due to bulge enhancement or disk fading by splitting the sample, in analogy to the analysis in §4.6.1, into six subsamples selected by radial distance from the cluster center. We then calculate  $M_{bulge}^*(\alpha = -1)$  and  $M_{disk}^*(\alpha = -1)$  for each subsample. To account for the fact that the six clusters in our sample have different intrinsic physical radii, we scale the radial distances by  $800 \text{ km s}^{-1} \sigma^{-1}$ , where  $\sigma$  is the line-of-sight cluster velocity dispersion. We refer to this scaled distance as  $R_\sigma$ . This scaling is motivated by the fact that the size of virialized systems increases approximately as  $\sigma$  (Girardi et al., 1998). The constant of  $800 \text{ km s}^{-1}$  serves to provide an orientation for the physical scale of a typical cluster.

In this analysis, the uncertainties in the Schechter fits are typically larger than those found in §4.6.1, and the effect is too subtle to determine whether there is a significant correlation of  $M_{bulge}^*$  or  $M_{disk}^*$  with radius. Nonetheless, we obtain constraints on the slopes of the  $M_{bulge}^*-R_\sigma$  and  $M_{disk}^*-R_\sigma$  relations of  $\partial M_{bulge}^*/\partial R_\sigma = (0.31 \pm 0.27) \text{ mag Mpc}^{-1}$  and  $\partial M_{disk}^*/\partial R_\sigma = -0.02 \pm 0.20 \text{ mag Mpc}^{-1}$ .

Does a more sensitive test allow us to establish whether the  $M_{bulge}^*-R_\sigma$  rela-

tion is significantly different from zero? We use a Spearman Rank Correlation test (Gibbons, 1976) to examine correlations between the bulge luminosities of individual galaxies (with a total absolute magnitude  $M_{total} \leq -19.2 + 5 \lg h$ ) and the radial distance  $R_\sigma$ . A rank correlation test is appropriate in this case, because the distribution of galaxy properties such as luminosity is very uneven and non-linear. In calculating the coefficients, we weight each data point by the weighting factor assigned to the galaxy by the DML solution. To additionally improve the significance of our conclusions, we exclude A1060, because of its small sampling radius ( $R_\sigma \approx 0.7h^{-1}$  Mpc centroid-to-corner), and analyze galaxies only within  $R_\sigma = 1.8h^{-1}$  Mpc, the largest radius to which the other five clusters are fully sampled.

The rank correlation coefficient for the correlation between bulge luminosity and projected radius for 937 galaxies is  $r_{M_{bulge}, R_\sigma} = 0.16$ , which is significant at the  $5\sigma$  level. The correlation is weak because, for any given radius, galaxies have a wide range of absolute magnitudes. The effect is nonetheless significant, implying that bulges are systematically brighter at smaller projected radii.

After performing a similar analysis on the disk luminosities, we find no correlation between disk luminosity and radius; the correlation coefficient is  $r_{M_{disk}, R_\sigma} = -0.02$ , a deviation of only  $0.5\sigma$  from zero. Does this result imply that the disk luminosities of individual galaxies remain constant as a function of radial distance? On its own, the fact that the correlation coefficient is close to zero does not permit such a conclusion, because effects at the bright and at the faint end of the disk luminosity function could cancel each other. For example, a simultaneous brightening of the brightest  $M_{disk}$  and increase in the number of faint disks could conspire to keep the median disk luminosity constant. Likewise, the fact that  $M_{disk}^*$  itself does not vary significantly with radius could also be explained if disk

galaxies (which are likely to dominate the bright end of the disk LF) lie far from the cluster center in three-dimensional space and are only seen close to the core in projection. However, if we consider both observations together — that neither the characteristic bright-end disk magnitude nor the median disk luminosity of all galaxies with  $M_R < -19.2$  vary significantly as a function of projected radius — then we can conclude that radial distance has no significant effect on the luminosities of individual galaxy disks.

The observed correlation of  $M_{bulge}$  with  $R_\sigma$ , in which bulges are brighter closer to the projected cluster center, may at least partially explain the morphology-environment relation. This result is consistent with those in §4.6.1, and thus provides additional support for the hypothesis that the morphology-environment relation is affected by galaxy-galaxy interactions that enhance bulge luminosities, rather than by disruption of star formation with subsequent passive evolution and fading of the disks.

#### 4.7 Conclusions

Using a new survey of six clusters of galaxies (Christlein & Zabludoff, 2003), we constrain the mechanisms by which early-type galaxies may evolve from late-type galaxies in two ways: 1) by comparing the luminosities of bulges and disks as functions of overall galaxy morphology, as measured by the bulge fraction  $B/T$ , and 2) by comparing the trends of bulge and disk luminosity with radial distance from the cluster center.

Models that generate early type galaxies by fading the disks of late-type galaxies are ruled out for a wide range of morphologies. Specifically, the bulges of galaxies with  $B/T \approx 0.45$ , typical of S0 galaxies, are significantly brighter than expected if disk fading were the primary mechanism transforming late-type into

early-type galaxies. Bulge Enhancement models, which increase  $B/T$  by increasing the luminosity of the bulge, are in excellent agreement with our observations over a wide range of  $B/T$ .

This result is strengthened by our comparison of bulge and disk luminosities as a function of projected radial distance from the cluster center. There is a significant tendency for bulges to be brighter towards the cluster center, while galaxy disks have a similar luminosity regardless of their position within the cluster. This test reveals no evidence that cluster-specific processes reduce the luminosity of disks, and thus does not support disk fading mechanisms such as ram pressure stripping or strangulation as shapers of the morphological sequence in clusters. Instead, our results favor processes like galaxy-galaxy interactions and mergers, which can enhance galaxy bulges. Such processes are more efficient in lower-density environments such as poor groups of galaxies, which, when accreted hierarchically by clusters, may have played a major role in generating the morphological sequence that we observe in rich clusters today.

## CHAPTER 5

### DISENTANGLING MORPHOLOGY, STAR FORMATION, STELLAR MASS, AND ENVIRONMENT

#### 5.1 Chapter Summary

We present a study of the spectroscopic and photometric properties of galaxies in six nearby clusters. We perform a partial correlation analysis on our dataset to investigate whether the correlation between star formation rates in galaxies and their environment is merely another aspect of correlations of morphology, stellar mass, or stellar age with environment, or whether star formation rates vary independently of these other correlations. We find a residual correlation of ongoing star formation with environment, indicating that even galaxies with similar morphologies, stellar masses, and stellar ages have lower star formation rates in denser environments. Thus, the star formation gradient in clusters is not just another aspect of the morphology-density relation. Furthermore, the star formation gradient cannot be solely the result of initial conditions, but must partly be due to subsequent evolution through a mechanism (or mechanisms) sensitive to environment. Our results constitute a true “smoking gun” pointing to the effect of environment on the later evolution of galaxies.

#### 5.2 Introduction

Galaxies exhibit a great diversity in mass, luminosity, morphology, and star formation activity. It is now generally recognized that these properties are correlated with the environment wherein a given galaxy resides. Such correlations are ex-

pected in hierarchical models of galaxy evolution, which predict the oldest, most massive, least star-forming galaxies to reside in the densest regions of the universe. Subsequent evolution brings less massive, still star-forming galaxies into these denser regions from the field. This leads us to expect gradients in galaxy properties such as morphology and star formation with environment, even from initial conditions alone. If galaxies are affected by environmentally-dependent mechanisms later in their lives, these gradients may be altered.

Observationally, one of the best-established among these correlations is the morphology-environment relation (Dressler, 1980), in which the morphological composition of the galaxy population is shifted towards earlier types — S0s and ellipticals — in denser environments, e.g., in groups or clusters compared to the field, near the center of a cluster, or in regions of locally enhanced galaxy number density.

Then there is the correlation between star formation and environment: galaxies have lower star formation rates in dense environments than elsewhere. This correlation has also been known for a long time (Gisler, 1978) and recently confirmed by large modern surveys (Lewis et al., 2002; Gómez et al., 2003), which demonstrate that the star formation rates in galaxies near clusters vary as a function of the distance of the galaxy from the cluster center. Such results have been interpreted as evidence for an effect of the cluster environment on star formation. However, these studies do not provide unambiguous proof for such an effect, because three important questions remain:

- 1) Is the star formation gradient in clusters simply a consequence of the morphology-environment relation? Early-type galaxies generally have lower star formation rates than late-type galaxies. Does this morphology-star formation correlation, combined with the morphology-environment relation, explain the star

formation gradient in clusters? Or is the evolution of star formation rates in different environments, at least to some extent, independent of the morphological evolution, either because different mechanisms act in different environments, or because the same mechanisms affect star formation and morphology differently in different environments?

The former result would indicate that the initial conditions and/or subsequent evolutionary processes that populate the morphological sequence are, in principle, the same in every environment, but have over time produced more early-type galaxies in dense environments. The latter alternative would demonstrate that an environment-dependent mechanism is actually at work, so that even galaxies with similar morphologies have different star formation rates in dense environments. Only this latter result would constitute proof for a direct and environment-specific effect on star formation.

2) Is the star formation gradient simply a consequence of variations in the galaxy luminosity or mass function with environment? Some evidence (de Propris et al., 2003) exists that the faint end of the luminosity function is shallower in the cores of clusters than on the outskirts. Since low-mass, faint galaxies are typically more strongly star-forming, the lower rates of star formation in clusters could be due to the fact that samples of dense environments contain a larger fraction of giant galaxies. As long as this question is unresolved, it is not clear whether environment has a direct effect on the star formation rates of individual galaxies, or whether the star formation gradient is simply a result of variations of the luminosity or mass function with environment (which could be due to initial conditions or subsequent evolution).

3) Does the star formation gradient in clusters simply arise because clusters contain a large fraction of long-quiescent galaxies? Although individual objects

are known that could be in the process of being transformed by mergers (Yang et al., 2004) or ram pressure stripping (Vollmer et al., 2004a,b), this does not constitute evidence that such ongoing transformations are primarily responsible for the gradients in morphology and star formation with environment. Many galaxies in clusters contain very old stellar populations and may have been quiescent for many Gyr, possibly since their formation. The star formation gradient could simply reflect initial conditions: these old galaxies constitute an increasing fraction of all galaxies towards the cluster center, as star-forming galaxies from the field dominate the cluster outskirts. Claiming an ongoing effect of environment on star formation would require us to detect a star formation gradient even for galaxies with comparable star formation histories (i.e., mean stellar ages).

Testing whether dense environments have a direct and ongoing effect on star formation rates in clusters therefore requires us to examine whether an independent correlation of star formation with environment exists that cannot be accounted for by the correlations of morphology, stellar age, and stellar mass with environment. Such an analysis must be based on measurements of galaxy spectra and structural parameters in a large sample. The observational capabilities for such a program, particularly for multi-fiber spectroscopy of a large number of objects, as well as the computational capabilities to quantitatively characterize the morphologies of a large sample of galaxies, have not existed until recently. A few studies (Gómez et al., 2003; Balogh et al., 1998) have attempted to address this question by splitting galaxies into broad morphological classes and examining the dependence of star formation within each class on environment. This procedure suffers from the problem that a coarse binning in morphology may leave residual correlations between star formation and morphology within each bin, and it is therefore not clear whether the residual correlations reported in



these studies are really independent of the morphology-environment relation. Hashimoto et al. (1998) use the concentration index as a quantitative measure of morphology, which allows for a finer and more reproducible binning in morphology. However, none of these studies provide a constraint on whether the lower star formation rate measured in denser environments is due to recent environmental effects or instead due to galaxies that have been quiescent for a long time, perhaps since their formation.

In this paper, we present the results of a multi-variate analysis of a spectroscopic sample of galaxies in six nearby clusters. We use a partial correlation analysis to examine whether there is a residual correlation between star formation and environment even when holding morphology, stellar mass, and stellar age fixed. Using partial correlation coefficients allows us to account for a large number of variables simultaneously and thus to disentangle the effect of environment from all other correlations, in order to isolate any ongoing, direct impact of environment on star formation. With this approach, we also avoid the problems associated with the usual method of representing morphology by coarse binning.

This paper is organized as follows: In §5.3 we briefly review the sample upon which our analysis is based. In §5.4, we discuss the variables — star formation, morphology, stellar mass, mean stellar age — that we use in our search for residual correlations between star formation and environment. We represent ongoing star formation with the equivalent width of the [OII]  $\lambda 3727$  doublet, and recent star formation with the equivalent width of H $\delta$  in absorption in the absence of [OII] emission. We characterize morphology with the bulge fraction, stellar mass with a combination of near-IR photometry and the 4000Å break strength, and mean stellar age with the 4000Å break strength. In §5.5, we review the mathematical tools used for the analysis. We discuss how we calculate partial corre-

lation coefficients to address the question of whether star formation varies differently with environment than morphology, stellar mass, or mean stellar age. We also present our procedure for calculating completeness corrections using the Discrete Maximum Likelihood method (Christlein, McIntosh & Zabludoff, 2004; Christlein & Zabludoff, 2004) and discuss the limits of our survey and its completeness within these limits. In §5.6, we present the results of our analysis. First, in §5.6.1, we explore different measures of environment and choose the projected radial distance of a galaxy from the cluster center as our environmental variable. We then verify that we can reproduce the morphology-environment and star formation-environment relations. In §5.6.2, we examine whether residual correlations of current and recent star formation with environment remain when accounting for the effects of morphology, stellar mass, and stellar age, and quantify the magnitude of the residual star formation gradient relative to other effects. In §5.6.3, we estimate the effect that observational uncertainties in our determinations of morphology, stellar mass, and stellar age have on the residual star formation gradient. In §5.7, we present our conclusions.

### 5.3 The Cluster Survey

Our sample consists of cluster galaxies from a spectroscopic and  $R$ -band imaging survey of six nearby clusters. The spectroscopic survey ensures that contamination of the sample by background field galaxies is minimized. Table 5.1 lists some parameters of these six clusters for  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ , as applied throughout this paper. For details regarding the survey and data reduction, we refer readers to Christlein & Zabludoff (2003).

Table 5.1. The Cluster Sample

Cluster	centroid (RA, Dec; J2000)	N	$\bar{cz}$ [km/s]	$\Delta m$ [mag]	$\Delta cz$ [km/s]	$\sigma$ [km/s]	$r_{sampling}$ [Mpc]
A1060	10 36 51.29 -27 31 35.3	252	$3683 \pm 46$	33.59	2292 - 5723	$724 \pm 31$	0.67
A496	04 33 37.09 -13 14 46.3	241	$9910 \pm 48$	35.78	7731 - 11728	$728 \pm 36$	1.76
A1631	12 52 49.84 -15 26 17.1	340	$13844 \pm 39$	36.53	12179 - 15909	$708 \pm 28$	2.42
A754	09 08 50.08 -09 38 11.8	415	$16369 \pm 47$	36.90	13362 - 18942	$953 \pm 40$	2.83
A85	00 41 37.81 -09 20 33.2	280	$16607 \pm 60$	36.94	13423 - 19737	$993 \pm 53$	2.87
A3266	04 31 11.92 -61 24 22.7	331	$17857 \pm 69$	37.10	14129 - 21460	$1255 \pm 58$	3.07

Note. —  $N$  is the number of sampled galaxies per cluster.  $\bar{cz}$  is the mean velocity,  $\Delta m$  the distance modulus (for  $H_0 = 71 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ).  $\Delta cz$  is the velocity range spanned by cluster members,  $\sigma$  is the line-of-sight velocity dispersion, and  $r_{sampling}$  is the projected physical radius sampled (centroid-to-edge).

## 5.4 Galaxy Properties

### 5.4.1 Star Formation Indices

Because the spectra for our sample are not flux calibrated, it is not possible to calculate absolute line fluxes and star formation rates. Instead, we use the equivalent width of the  $[OII]\lambda 3727$  doublet as a measure of current star formation. Although sensitive to dust and metallicity effects (Kewley, Geller & Jansen, 2004),  $EW([OII])$  is a sufficiently accurate indicator of star formation rate (Kennicutt, 1998) for our purposes, which only require an approximate relative measure, not a precise calibration.

A suitable measure of recent, but not on-going, star formation is the  $H\delta$  Balmer absorption line in the absence of significant  $[OII]$  emission. Balmer absorption lines are typically strongest in galaxies whose light is dominated by stars of spectral type A, and so their equivalent widths are particularly sensitive to a galaxy's

star formation activity over the timescale of the past  $\sim$  Gyr. A caveat is that, in strongly star-forming galaxies, Balmer emission from HII regions can fill the Balmer absorption lines and thus affect our measurement of the  $H\delta$  equivalent width. To avoid this, we only consider galaxies that have no significant [OII] emission (defined as having [OII] line emission detectable at less than  $2\sigma$ ; see Zabludoff et al. 1996).

To calculate an equivalent width, the local continua are fit over the 100 pixels ( $\sim 250\text{\AA}$ ) on either side of the line that exclude the line itself and the nearby sky lines. Beginning at the line center, the line is integrated outward until reaching the continuum level. That uncalibrated flux and the interpolated value of the continuum at line center are used to calculate the equivalent width. The equivalent width uncertainties, which are typically less than  $1\text{\AA}$ , are calculated using counting statistics (the detector is a photon counter with approximately zero read noise), the local noise in the continuum, and standard propagation of errors. Equivalent widths are cosmologically corrected.

Values of  $EW([OII])$  and  $EW(H\delta)$  for our sample are tabulated in Table 5.2.

#### 5.4.2 Morphologies

We use the GIM2D software (Simard et al., 2002) to perform a two-dimensional decomposition of the galaxy images into a bulge component, described by a de Vaucouleurs surface brightness profile (de Vaucouleurs, 1948), and a disk component with an exponential surface brightness profile. We then calculate the fraction  $B/T$  of the total luminosity associated with the bulge. Prior to the fit, we transform all galaxy images to a fiducial rest frame  $cz = 17858\text{ km/s}$  (corresponding to the mean velocity of the most distant cluster, A3266) by fading the surface brightnesses by  $(z_{\text{cosmological}} + 1)^{-2}(z_{\text{total}} + 1)^{-2}$ , smearing the images to achieve a consistent FWHM of 2 arcsec, and rebinning their pixels with the new angular di-

Table 5.2. Spectroscopic Catalog (Example)

ID	RA (J2000)	Dec (J2000)	$D4000$	$\Delta D4000$	$EW([OII])$ [Å]	$\Delta EW([OII])$ [Å]	$EW(H\delta)$ [Å]	$\Delta EW(H\delta)$ [Å]
1060A_295[11]	10 36 13.46	-27 31 54.90	1.828	0.073	1.80	2.10	-0.10	0.57
1060A_295[26]	10 35 37.40	-27 10 49.60	1.893	0.103	3.13	2.87	1.53	1.34
1060A_295[41]	10 34 48.52	-26 53 30.40	1.604	0.075	2.65	2.21	2.02	1.22
1060A_393[1]	10 36 49.04	-27 23 18.60	2.257	0.057	-0.16	0.56	0.30	0.59
1060A_393[4]	10 36 35.60	-27 08 56.20	1.852	0.050	-0.28	0.63	0.07	0.42
1060A_393[6]	10 36 26.76	-27 23 24.80	1.741	0.056	0.48	1.24	0.08	0.97
1060A_393[7]	10 35 55.70	-27 14 11.90	1.997	0.048	0.72	0.97	1.16	1.01
1060A_393[10]	10 36 19.30	-27 28 45.90	1.786	0.055	-1.28	0.74	0.69	0.67
1060A_393[12]	10 35 46.79	-27 38 48.70	1.515	0.037	8.43	1.22	2.76	1.14
1060A_393[13]	10 35 38.27	-27 31 34.30	1.892	0.034	0.99	0.79	1.21	0.94
1060A_393[14]	10 35 21.68	-27 23 26.20	1.365	0.026	4.44	0.77	7.42	1.32
1060A_393[21]	10 36 29.47	-27 45 29.10	1.727	0.046	0.04	0.65	1.32	0.61
1060A_393[22]	10 36 21.38	-27 46 30.60	1.721	0.052	0.92	1.60	1.96	0.96
1060A_393[23]	10 36 03.81	-27 55 08.30	1.787	0.047	-0.26	0.65	0.79	0.61
1060A_393[24]	10 36 01.81	-27 41 07.00	1.955	0.039	0.68	0.81	0.33	0.51
1060A_393[25]	10 35 55.21	-27 45 16.60	2.110	0.053	3.91	1.88	1.50	1.04
1060A_393[27]	10 35 53.88	-27 22 19.90	2.126	0.053	-0.70	0.60	-0.07	0.39

Note. — The full table is published in Christlein & Zabludoff (2004b). The  $R$ -band magnitudes are published in Christlein & Zabludoff (2003).

ameter distance to achieve the same physical resolution per pixel. This approach ensures that determinations of  $B/T$  are internally consistent among the clusters, which span a mean velocity range from 3682 to 17858 km/s. The final catalog contains bulge-disk decompositions of 1637 galaxies (1304 of them for galaxies with  $M_R \leq -19.2$ ). This procedure is also discussed in Christlein & Zabludoff (2004).

The limit of  $m_R = 18$  that we adopt for the spectroscopic catalog as well as for the bulge-disk decompositions corresponds to  $M_R \approx -19.2$  in the fiducial field. For fainter galaxies, the signal-to-noise ratio is too low to permit reliable fits, and spectroscopic information is not available for all clusters. We therefore do not analyze the sample beyond  $M_R = -19.2$ .

Table 5.3 contains the coordinates and  $B/T$  values used in this paper.

Is  $B/T$  an adequate proxy for galaxy morphology? The null hypothesis that we test in §5.5 is that the star formation-environment relation and the morphology-environment relation are completely interdependent. For testing this hypothesis, it is sufficient to fix a variable that is strongly correlated with morphology. All environmentally-dependent transformation processes suggested in the literature — including mergers (Barnes, 1999; Bekki, 1998; Mihos & Hernquist, 1994) and ram-pressure stripping (Gunn & Gott, 1972) — that could affect star formation also have a strong impact on  $B/T$ . The bulge fraction is therefore an appropriate variable to use in this analysis.

#### 5.4.3 Luminosities and Stellar Masses

We estimate stellar masses in the following way: given a measure of the color of a galaxy, it is possible to reconstruct an approximate mass-to-light ratio and, with a corresponding luminosity measurement, its stellar mass. Infrared magnitudes are best suited for this purpose, because the mass-to-light ratio is most

Table 5.3. Morphological Catalog (Example)

ID	RA (J2000)	Dec (J2000)	$B/T$	$B/T_{min}$	$B/T_{max}$	$B/T_f$	$B/T_{f,min}$	$B/T_{f,max}$	$\omega$
1060A_295[11]	10 36 13.46	-27 31 54.90	0.03	0.00	0.14	0.82	0.00	1.00	4.872
1060A_295[26]	10 35 37.40	-27 10 49.60	0.13	0.00	0.25	0.40	0.00	1.00	5.690
1060A_295[41]	10 34 48.52	-26 53 30.40	0.00	0.00	0.09	0.78	0.00	1.00	10.319
1060A_393[1]	10 36 49.04	-27 23 18.60	0.68	0.64	0.73	0.60	0.12	1.00	1.872
1060A_393[4]	10 36 35.60	-27 08 56.20	0.09	0.08	0.12	0.04	0.00	0.29	0.754
1060A_393[6]	10 36 26.76	-27 23 24.80	0.29	0.20	0.34	0.54	0.00	1.00	4.717
1060A_393[7]	10 35 55.70	-27 14 11.90	0.30	0.26	0.33	0.80	0.00	1.00	2.699
1060A_393[10]	10 36 19.30	-27 28 45.90	0.52	0.45	0.59	0.81	0.36	1.00	4.734
1060A_393[12]	10 35 46.79	-27 38 48.70	0.00	0.00	0.05	0.54	0.00	1.00	4.741
1060A_393[13]	10 35 38.27	-27 31 34.30	0.00	0.00	0.00	0.04	0.00	0.18	1.160
1060A_393[14]	10 35 21.68	-27 23 26.20	0.27	0.10	0.43	0.83	0.03	1.00	4.661
1060A_393[21]	10 36 29.47	-27 45 29.10	0.37	0.23	0.41	0.36	0.00	0.84	4.737
1060A_393[22]	10 36 21.38	-27 46 30.60	0.23	0.17	0.29	0.44	0.00	1.00	4.571
1060A_393[23]	10 36 03.81	-27 55 08.30	0.20	0.16	0.24	0.24	0.00	0.64	1.404
1060A_393[24]	10 36 01.81	-27 41 07.00	0.26	0.24	0.28	0.10	0.00	0.49	2.571

Note. —  $B/T$  values are given as the best fit values and the minimum and maximum values within the 99% uncertainty interval, as determined by GIM2D (Simard et al., 2002). Values with subscript  $f$  refer to the fits to images that have been artificially faded to the fiducial redshift of 17858 km/s. The ' $\omega$ ' column lists the statistical weighting factors, calculated by the Discrete Maximum Likelihood method, which corrects for the spectroscopic, morphological and radial incompleteness of the sample. The table contains galaxies down to  $m_R = 18$ , not all of which have been used for our analysis. The full table is published in (Christlein & Zabludoff, 2004b). The radial velocities and  $R$ -band magnitudes are published in Christlein & Zabludoff (2003).

stable in these magnitude bands and least affected by ongoing star formation or dust attenuation.  $R$ -band magnitudes are available for all galaxies in the sample (Christlein & Zabludoff, 2003), and  $J$ - and  $K$ -band magnitudes for  $\sim 80\%$  of them from the 2MASS survey. A rough measure of color that is available for all our galaxies is  $D4000$ , the strength of the  $4000\text{\AA}$  break. This is defined as

$$D4000 = \frac{\int_{4050}^{4250} f_{\lambda} d\lambda}{\int_{3750}^{3950} f_{\lambda} d\lambda}. \quad (5.1)$$

We use the GALAXEV spectrophotometric evolution code (Bruzual & Charlot, 2003) to determine a relation between mass-to-light ratios in the  $R$ -,  $J$ -, and  $K$ -band, and  $D4000$ . Because this relation is dependent on a galaxy's star formation history, we calculate it for three different models: a single-burst model (reflecting a galaxy with an old stellar population), a model with exponentially declining star formation rate (with an e-folding timescale of 2 Gyr, chosen to provide an intermediate scenario between a population dominated by a single burst and one dominated by near-constant star formation), and a model with a constant star formation rate of  $1 M_{\odot} \text{ yr}^{-1}$  (representative of a field galaxy with on-going star formation). These three models cover a baseline of different, albeit schematic, evolutionary histories and thus allow us to estimate the impact that the choice of model has on the determination of the stellar mass. In all cases, we use the Padova 1994 models with solar metallicity and a Salpeter Initial Mass Function (Salpeter, 1955).

From each of the three bands (where available) and each of the three population models, we calculate stellar masses for the galaxies in our sample. We do not know a priori which of the three models best represents a given galaxy, but an accurate population model should yield the same stellar mass estimate from any magnitude band. For this reason, if  $J$ - and  $K$ -band photometry are available for a given galaxy, we pick the population model for which the stellar masses



calculated from these two bands show the best agreement and adopt the mean of the two masses as the true stellar mass of the galaxy, referred to as  $M_*$  (this is equivalent to choosing the model that best matches the  $J - K$  color of a galaxy).

For comparison, we also calculate stellar masses from the  $R$ -band photometry and the single-burst model for each galaxy.  $R$ -band-derived stellar masses are usually less reliable because the  $R$ -band mass-to-light ratio is more sensitive to stellar age. By comparing the  $R$ -band-derived stellar masses to the estimates based on  $J$ - and  $K$ -band photometry, we find that, for large stellar masses, the  $R$ -band yields values up to  $\sim 50\%$  larger than does the infrared photometry. Therefore, for the 11% of galaxies in the sample that do not have  $J$ - and  $K$ -band measurements, we use the  $R$ -band-derived stellar masses, but apply a correction to compensate for the systematic deviation from the  $J$ - and  $K$ -band stellar masses. (An alternative reason for the 50% discrepancy could be that the *2MASS* photometry underestimates the total luminosity of a galaxy. However, for our purposes, the absolute calibration of the stellar mass scale is not important, and if such a systematic bias existed, it would not affect our conclusions.)

#### 5.4.4 Mean Stellar Age

We adopt the  $D4000$  index, as defined in Eq. 5.1, as a measure of mean stellar age. In reality,  $D4000$  and mean stellar age are not perfectly correlated. To estimate the reliability of  $D4000$  as a stellar age indicator, we calculate the correlation coefficient between mean stellar age ( $MSA$ ) and  $D4000$ ,  $r_{D4000,MSA}$ , from a set of GALAXEV population models. We run a variety of models with different star formation histories, consisting of an initial starburst and a variable number of later starbursts, interspersed with episodes of constant star formation, exponentially declining star formation, or quiescence. For all runs, we use the Padova 1994 models with solar metallicity and a Salpeter Initial Mass Function (Salpeter,

1955). By running a variety of models with different initial star formation histories and taking one data point at time  $t < 13.7$  Gyr from each galaxy that has been quiescent for at least 1 Gyr, we find

$$r_{D4000,MSA} = 0.91. \quad (5.2)$$

Thus  $D4000$  is a very good indicator of the mean stellar age of a galaxy, at least for galaxies that are not currently star forming. In the presence of current or recent star formation,  $D4000$  can be biased low. We discuss the impact of this systematic error on our results in §5.6.3.

The star formation history of any galaxy is more complex than can be summarized in a single variable, such as  $D4000$ . However, for our purpose of establishing whether the star formation gradient results from the recent or current effects of environment as opposed to initial conditions, using  $D4000$  as a proxy of star formation history and mean stellar age is adequate.

## 5.5 The Analysis

### 5.5.1 Partial Correlation Coefficients

In our search for trends between environment, star formation, and morphology, we use correlation analyses as our primary tool. The correlation coefficient  $r_{xy}$  is a simple and straightforward way of testing a sample for interdependences between any two variables  $x$  and  $y$ . Unlike linear regression slopes, which assume one variable is dependent on the other one, correlation coefficients are symmetric with regard to the variables whose correlation they describe. This is a particular advantage when we want to analyze a data set without inferring a causality beforehand.

However, there are many problems where two variables may be covariate without necessarily being causally related. Both could be dependent on a third,

as yet undiscovered variable, or a third variable could be a causal link between the two. If the third variable were somehow fixed at a constant value, no correlation would be observed between the two primary variables. In our analysis, the primary example is the correlation between star formation and environment. Is there a direct causal link between environment and star formation, or can the entire variation in star formation with environment be understood as a consequence of, say, the morphology-environment relation? Determining whether such a third variable exists, what impact it has on the observed correlations, and whether the correlations persist if the variable is held constant, is crucial in examining possible causal connections between environment, star formation, morphology, stellar mass, and mean stellar age.

To investigate whether there is a dependence of star formation on environment beyond that accounted for by these other variables, we use partial correlation coefficients. Partial correlation coefficients quantify the correlation that two variables, referred to here as 1 and 2, would exhibit if a third variable 3, or even a set of variables  $(3, 4, \dots, k)$ , was held constant, or “controlled”.

Following Kendall & Stuart (1977) in derivation and notation, the partial correlation coefficient between variables 1 and 2 after partialling out variables 3 through  $k$  is

$$r_{12.34\dots k} = \frac{-C_{12}}{(C_{11}C_{22})^{1/2}}. \quad (5.3)$$

Here, the cofactor  $C_{ij}$  is given by

$$C_{ij} = (-1)^{i+j} M_{ij}, \quad (5.4)$$

and  $M_{ij}$  is the minor of the correlation matrix  $(r_{ij})$ , i.e., the reduced determinant of the matrix that ensues from discarding row  $i$  and column  $j$  of  $(r_{ij})$ . The partial correlation coefficient is therefore only dependent on the total correlation coefficients  $r_{ij}$ .

The partial correlation coefficients can be tested for significant departures from the null hypothesis in the following way: We follow Kendall & Stuart (1977) in defining a statistic

$$Z = \frac{1}{2} \ln \left( \frac{1+r}{1-r} \right). \quad (5.5)$$

Unlike the partial correlation coefficient  $r$  itself, the probability distribution of  $Z$  approximates a normal distribution even for moderate ( $> 50$ ) numbers of data points  $N$ , so that its variance can be used to estimate the significance of the correlation. The variance is approximately given by

$$\text{var} Z = \sigma_Z^2 = \frac{1}{N-1-k}. \quad (5.6)$$

Controlling one variable only reduces the number of degrees of freedom by one. Because our sample contains hundreds of galaxies, our analysis retains discriminatory power even if several variables are partialled out. Thus, we can simultaneously control not only morphology, but also stellar mass and mean stellar age.

Correlation coefficients can be affected strongly by outlying data points, and in such cases, their physical significance is questionable. This is a particular problem for our data, which involve highly non-linear distributions of observables. For example, while most galaxies in the cluster sample have little star formation, and therefore small  $EW([OII])$ , the few emission-line galaxies may have  $EW([OII])$  values many times the typical standard deviation of the entire sample. Ideally, our analysis should be sensitive to correlations both within the bulk of the low- $[OII]$  population and among the outliers to this distribution.

This problem can be addressed by using the Spearman Rank Correlation Coefficient (Gibbons, 1976) instead of the Pearson Correlation Coefficient whose definition we have given above. For this purpose, we replace every variable  $x$  by its

rank  $rank(x)$ , which is the number of galaxies  $i$  with  $x < x_i$  plus  $1/2$  the number of galaxies with  $x = x_i$ . Then, the correlation coefficient is calculated from ranked quantities as described above. Unless explicitly noted otherwise, when we discuss correlations between two variables or provide a correlation coefficient  $r_{xy}$ , we always refer to rank correlation coefficients. Also note that we use a simplified notation in the subscripts of our correlation coefficients, e.g.,  $OII$  instead of  $EW([OII])$ .

### 5.5.2 Completeness-Corrected Correlation Coefficients

Our sample is incomplete for several reasons: the spectroscopic catalog is incomplete because some galaxies with faint apparent magnitudes and surface brightnesses were not targeted for spectroscopic observations, or observations failed to yield a redshift. The morphological catalog, which contains the results of the bulge/disk decomposition for each galaxy, is incomplete because we only fit galaxies to a limiting magnitude of  $m_R = 18$ , and a small number of fits are not successful even for brighter galaxies. Incompleteness introduces an unsystematic weighting into the calculation of the correlation coefficient. For example, if the sample is incomplete at faint magnitudes, and there is a stronger correlation between morphology and star formation for faint galaxies than for bright ones, the correlation coefficient for the incomplete sample will be weighted towards the bright end and artificially weakened.

To make our results more reproducible, we correct for this incompleteness. The solution is to calculate weighted correlation coefficients, where the weighting factors contain the corrections necessary to account for the sample incompleteness. Our sample spans a range of redshifts, so these weighting factors involve not only the completeness of the spectroscopic and morphological catalogs, but also a volume correction (e.g., a galaxy with  $M_R = -20$  can be observed in all

six clusters, while a galaxy with  $M_R = -18.5$  is below the detection limit in all but two of our clusters and therefore requires a larger correction). For the calculation of these weighting factors, we use the Discrete Maximum Likelihood method, which we describe in §5.5.3. Although initially developed to calculate luminosity functions, the DML is ideally suited for applications such as these, because it calculates individual weighting factors for each galaxy that comprise all completeness and volume corrections.

To calculate the corrected correlation coefficients, we make use of the definition of the total correlation coefficient as (apart from the sign) the geometric mean of the linear regression slopes of  $y$  on  $x$  and of  $x$  on  $y$ . To incorporate completeness corrections, we allow for each data point  $i$  to be weighted by a factor  $\omega_i$ . This factor, which we calculate in §5.5.3, contains all necessary completeness corrections. The regression slopes are then:

$$b_{yx} = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})\omega_i}{\sum_i (x_i - \bar{x})^2\omega_i}. \quad (5.7)$$

The means  $\bar{x}$  and  $\bar{y}$  are also computed as weighted means. Analogously, we calculate  $b_{xy}$ . The correlation coefficient is then given by

$$r_{xy} = \sqrt{b_{yx}b_{xy}} \text{SGN}(b_{xy}). \quad (5.8)$$

The sign function,  $\text{SGN}$ , ensures that the correlation coefficient has the same sign as either of the regression coefficients. Since partial correlation coefficients are calculated directly from the total correlation coefficients, this procedure is very straightforward to apply to partial correlation coefficients as well.

In general, these completeness-corrected coefficients do not obey the analytical probability distribution that underlies Eq. 5.6; the real probability distribution depends on the choice of  $\omega_i$  and can be recovered with a Monte Carlo algorithm. In our case, the distribution for the null hypothesis that the expectation value of

the correlation coefficient,  $\langle r \rangle$ , is zero turns out to be so similar to the analytical prediction for the unweighted case (the standard deviation for the corrected correlation coefficients is  $< 3\%$  larger) that we can use Eq. 5.6 to estimate the significance of our results.

### 5.5.3 The Discrete Maximum Likelihood Method

As we discuss in §5.5.2, statistical investigations of a sample of galaxies — whether by the use of luminosity or other distribution functions, or by multivariate techniques such as in this paper — require corrections to account for any incompleteness of the sample.

Maximum Likelihood methods have traditionally been used in studies of galaxy surveys to calculate luminosity functions from samples that are subject to incompleteness and limits in apparent magnitude. The procedure behind these algorithms is to assume a parent distribution function to predict the distribution of galaxies over absolute magnitude and possibly other variables (given the same observational selection effects) and to compare the result to the actual survey. The assumed parent distribution is then adjusted to maximize the probability that the observed sample has been drawn from it. Maximum Likelihood Estimators are very versatile and unbiased by large-scale density inhomogeneities, a particular advantage for field galaxy surveys. However, the functional form and dimensionality of the solution has to be chosen a priori, and the algorithm only solves for a number of pre-selected parameters, discarding information about individual galaxies. Since correlation coefficients are calculated from a set of data points, each of which represents an individual object, this approach is not helpful in our case.

We have therefore developed the Discrete Maximum Likelihood method specifically to determine completeness and volume corrections for galaxies in param-

eter spaces of arbitrary dimensionality. Unlike traditional algorithms, the DML associates its free parameters, the weighting factors  $\omega_i$ , not with fixed grid positions in a given parameter space (e.g., with fixed bins along the luminosity axis), but with the individual sampled galaxies. The DML thus combines the advantages of maximum likelihood methods with those of older methods such as the  $V/V_{max}$  method, which also associate individual weighting factors with galaxies but that calculate them in ways more susceptible to bias. Our method does not require us to assume a functional form for the distribution function and is, in fact, even independent of the dimensionality of the parameter space in which we are interested. Its disadvantage is that it does not probe regions of parameter space that do not contain galaxies and therefore does not automatically flag regions that could contain large numbers of galaxies outside the survey limits.

The derivation of the DML is described in detail in Christlein, McIntosh & Zabludoff (2004). We therefore only provide an overview here. The principle of the DML is to represent the *ansatz* for the distribution function that we wish to recover by the sampled galaxies themselves:

$$\varphi(\vec{x}) = C \sum_n \omega_n \delta(\vec{x}; \vec{x}_n), \quad (5.9)$$

where  $C$  is a normalization constant and  $\vec{x}_n$  is a parameter vector for galaxy  $n$  of arbitrary dimensionality. The weighting factors  $\omega_n$  are free parameters to be determined by the DML algorithm. This *ansatz* is similar to that of the  $C$  method (Lynden-Bell, 1971; Choloniewski, 1986), but the procedure for solving for the free parameters,  $\omega$ , is different, with our method retaining the benefits of ML estimators.

Applying this *ansatz* to a maximum likelihood approach yields the following



iteration formula for the weighting factors  $\omega$ :

$$\omega_h = \left( \sum_i \frac{f(\vec{x}_h | \vec{F}_i)}{\sum_g \omega_g f(\vec{x}_g | \vec{F}_i)} \right)^{-1}. \quad (5.10)$$

Here,  $f(\vec{x}_h | \vec{F}_i)$  is the completeness function of the survey — more specifically, it is the probability that a galaxy with the physical characteristics (e.g., luminosity, rest-frame surface brightness) of galaxy  $h$  would have been sampled by our survey if it were in the field (i.e., at the redshift and celestial coordinates, described by the parameter vector  $\vec{F}$ ) of galaxy  $i$ . While the DML algorithm is independent of the dimensionality of the parameter space that we want to investigate, the calculation of the completeness function is specific to a given survey. We calculate the completeness function of our survey in §5.5.4.

For the faintest galaxies that we consider in our analysis ( $M_R = -19.2$ ), we generally find weighting factors  $\omega_h < 3$ .

#### 5.5.4 The Completeness Function

In this section, we describe how we calculate the completeness function  $f(\vec{x} | \vec{F})$  of the sample. The calculation of the completeness function is the only part of the DML algorithm that has to be customized for a given survey.

We assume that our master catalog of photometric detections in the  $R$ -band is complete, and we determine the completenesses of the spectroscopic and morphological catalogs relative to it. There are three sources of incompleteness to consider.

##### 5.5.4.1 Spectroscopic Incompleteness

The first, and most severe, source of incompleteness is *spectroscopic incompleteness*,  $f_{spec}$ . For a discussion of how we calculate the spectroscopic completeness function, see Christlein & Zabludoff (2003). The completeness function (averaged

over all six clusters) is also plotted there as a function of apparent magnitude  $m_R$  and surface brightness  $\mu_R$ .

A source of uncertainty in the spectroscopic completeness function is the assumption that

$$\frac{N_{spec}}{N_{det}} = \frac{N_{spec,cl}}{N_{det,cl}}. \quad (5.11)$$

Here,  $N_{det}$  is the total number of photometrically detected galaxies (at a given magnitude and surface brightness), and  $N_{spec}$  is the number of spectroscopically sampled galaxies. The index  $cl$  indicates the same quantities only for cluster galaxies. This equation assumes that the theoretical sampling fraction for cluster galaxies, in which we are interested, is the same as the empirical sampling fraction for all galaxies, which we can recover from the data.

The assumption above is affected by statistical uncertainties. The number of sampled cluster galaxies follows a hypergeometric distribution, and the boundaries of the uncertainty interval are approximately given by

$$N_{spec,cl} = N_{spec} \frac{N_{det,cl}}{N_{det}} \left( 1 \pm \left( 1 - \frac{N_{det,cl}}{N_{det}} \right) \frac{N_{det} - N_{spec}}{N_{det} - 1} \right). \quad (5.12)$$

From this, it follows that the real sampling fraction for cluster galaxies is

$$f_{spec,cl} = \frac{N_{spec}}{N_{det}} \left( 1 \pm \left( 1 - \frac{N_{det,cl}}{N_{det}} \right) \frac{N_{det} - N_{spec}}{N_{det} - 1} \right), \quad (5.13)$$

and the approximate uncertainty in  $f_{spec}$  is

$$\Delta f_{spec} = f_{spec} \left( 1 - \frac{N_{det,cl}}{N_{det}} \right) \frac{N_{det} - N_{spec}}{N_{det} - 1}. \quad (5.14)$$

If we apply our assumption in Eq. 5.11 to this formula and substitute  $\frac{N_{spec,cl}}{N_{spec}}$  for  $\frac{N_{det,cl}}{N_{det}}$ , we obtain a first-order estimate for the uncertainties arising from this assumption. A stricter treatment should avoid this assumption altogether, but for our purposes, this procedure is sufficient, because, as we show in the next paragraph, the errors are negligible anyway.

Because of the various processing steps to which we subject the completeness function, the best way to estimate the impact of these statistical uncertainties on the final weighting factors  $\omega$  is by propagating them through the DML with a Monte Carlo approach. At this point, our intention is only to investigate whether the errors are small enough that they can be neglected safely. For that purpose, we calculate two sets of  $\omega$ , one using our best-estimate completeness function, and one for which we apply the uncertainty defined by Eq. 5.14 to the sampling fraction values before calculating  $\omega$ . We then compare the two sets of  $\omega$ , and find that the mean uncertainty in  $\omega$  is less than 5% even for the faintest galaxies considered here, and the most extreme outliers vary only by  $\sim 20\%$ . A more thorough analysis could estimate uncertainties on the individual  $\omega$  by testing a larger number of Monte Carlo realizations, but for our purposes, it is safe to conclude that statistical uncertainties in  $\omega$  are sufficiently small so as not to affect our analysis.

#### 5.5.4.2 Morphological Completeness

The second source of incompleteness of our sample is the morphological catalog, which contains bulge-disk decompositions for the majority of our galaxies. We determine this incompleteness empirically as we did for the spectroscopic completeness function, but use  $m_R$  as the only variable (as opposed to  $m_R$  and  $\mu_R$  for the spectroscopic completeness function). Note that bulge-disk decompositions for the artificially faded galaxy images are only available for cluster galaxies with known redshifts. The morphological completeness is therefore highly correlated with the spectroscopic completeness, and it is necessary to calculate it as a conditional completeness, i.e., only among cluster galaxies with spectroscopic information.

The completeness functions (only for spectroscopically confirmed cluster members) of all six clusters are shown in Fig. 5.1. The completeness is very high all

the way to our cutoff magnitude of  $m_R = 18$ . We do not distinguish between incompleteness due to a galaxy not being targeted for a bulge-disk decomposition or due to the bulge-disk decomposition failing.

#### 5.5.4.3 Radial Incompleteness

One selection effect for which we do not correct arises from the spatial boundaries of the survey. In some clusters, the sample includes galaxies at larger projected physical radii than in others. Although the DML allows us to correct for radial sampling incompleteness as well, such a correction would give undue weight (up to  $\omega \approx 17$ ) to a small subset of galaxies, severely distorting the probability distribution of the correlation coefficient. Because most of the clusters are sampled to physically comparable radii and radial incompleteness could only introduce a bias indirectly via secondary correlations with cluster global properties, this is not a serious problem for our analysis.

#### 5.5.4.4 Summarial Completeness Function

With contributions from both spectroscopic and morphological selection effects, the completeness function, defined as the probability that a galaxy with the physical parameters (absolute magnitude, rest frame surface brightness) of galaxy  $i$  would be observed in the survey if it were located in the field (redshift and spatial coordinates) of galaxy  $j$  is

$$f(\vec{x}_i | \vec{F}_j) = f_{\text{morphological}}(\vec{x}_i | \vec{F}_j, \text{spec}) f_{\text{spec}}(\vec{x}_i | \vec{F}_j). \quad (5.15)$$

We calculate the summarial completeness function only individually for each galaxy, so we do not plot it here. However, because there is no strong dependence of the morphological completeness function on  $m_R$  for  $m_R < 18$ , the shape of the summarial completeness function is similar to the spectroscopic completeness function shown in Christlein & Zabludoff (2003).

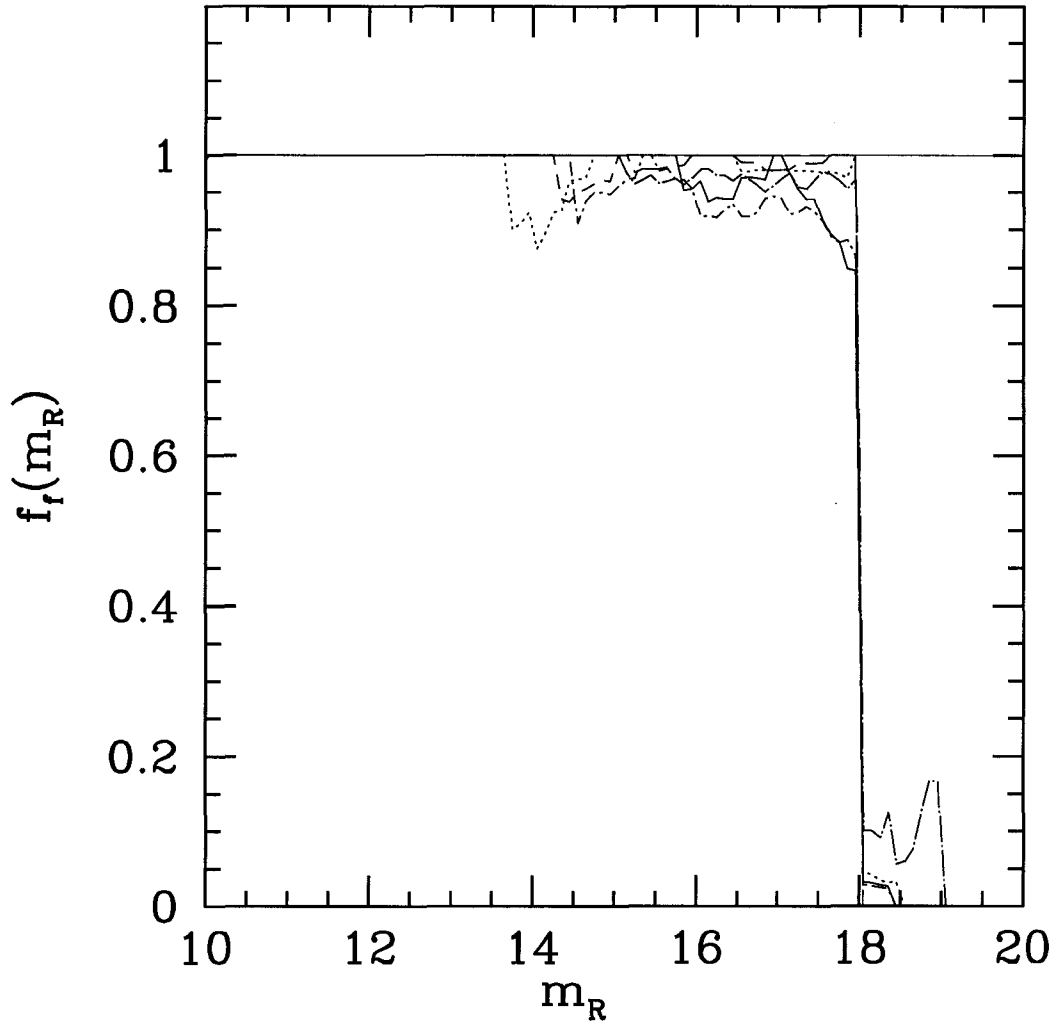


Figure 5.1: Completeness functions of the morphological catalog for all six clusters. Note that completeness is only measured among confirmed cluster members. The morphological catalog has a high level of completeness over all magnitudes brighter than our cutoff limit of  $m_R = 18$ .

## 5.6 Results

### 5.6.1 Morphology-Environment and Star Formation-Environment Relations

In this section, we check whether we can detect the morphology-environment relation (Oemler, 1974; Dressler, 1980) and the star formation-environment relation (Lewis et al., 2002; Gómez et al., 2003) in our sample. We have examined a range of environmental indices, including local density measures as used by Dressler (1980), different scalings of the radial distance from the cluster center, and spatial-kinematic measures such as the Dressler-Shectman  $\delta$  statistic (Dressler & Shectman, 1988). All of these indices are significantly correlated with morphology and star formation in our sample.

It is currently not known whether the properties of galaxies in clusters are primarily a function of their distance from the cluster center (Whitmore & Gilmore, 1991; Whitmore, Gilmore & Jones, 1993) or of local density (Dressler, 1980). In our sample, local density and radius are strongly correlated with each other, so that using either environmental variable will also detect variations in galaxy properties that are more directly dependent on the other variable. We select projected radius as our environment index, because local density measurements are typically affected by much larger statistical errors, which compromises our aim of finding an environmental variable that is well-correlated with galaxy properties.

More specifically, we use the projected physical distance of a galaxy from the cluster centroid (as defined in Table 5.1), scaled by the inverse of the cluster velocity dispersion  $\sigma$ :

$$R_\sigma = R_{phys} \sigma^{-1} 800 \text{ km s}^{-1}. \quad (5.16)$$

The constant of  $800 \text{ km s}^{-1}$  is typical of rich clusters (Zabludoff, Huchra & Geller, 1990). The scaling by  $\sigma^{-1}$  accounts for the fact that the physical length scales of virialized systems vary as  $\sigma$  (Girardi et al., 1998).

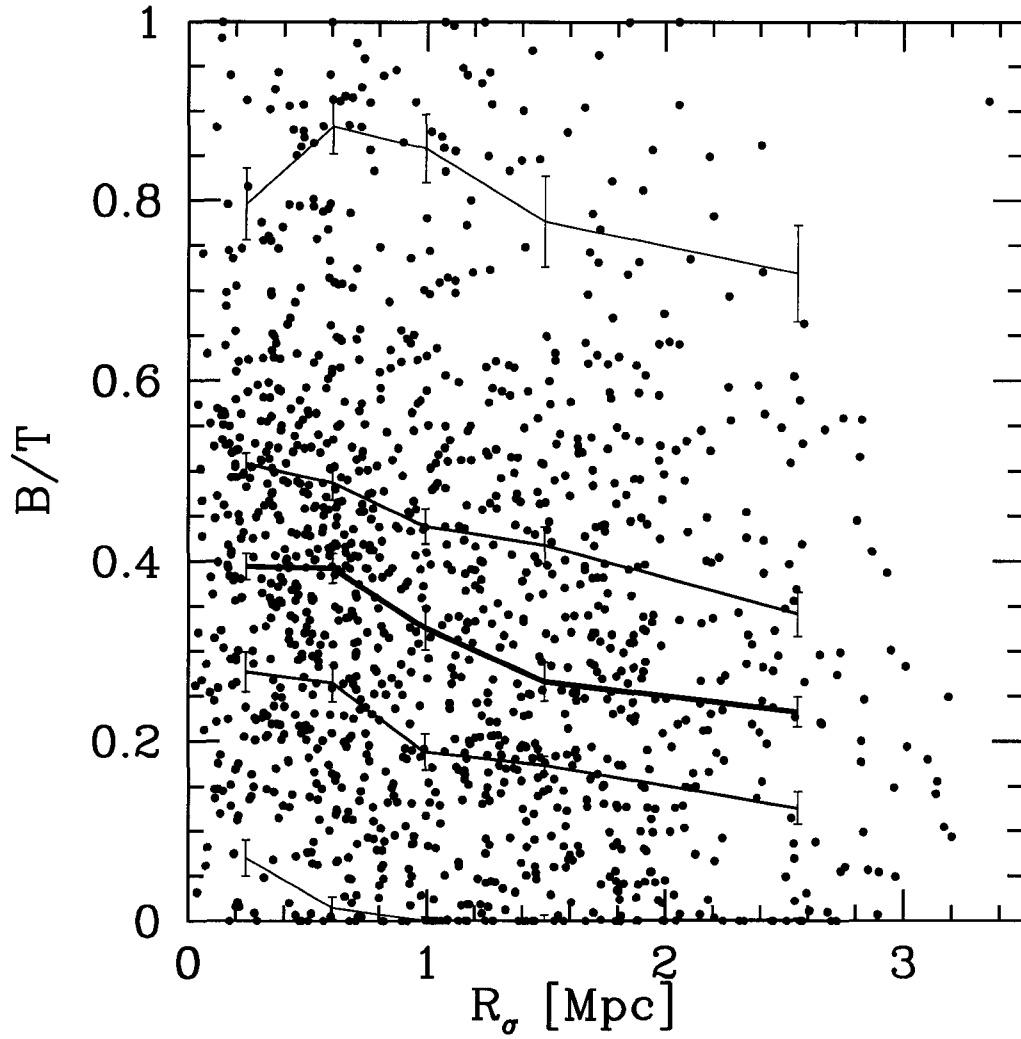


Figure 5.2: Morphology-Environment Relation. The figure shows the (0.0485,0.315,0.5,0.685,0.9515) quantiles of the  $B/T$  distribution as a function of cluster-centric radius,  $R_\sigma$ . The median is represented by the bold line. There is a significant correlation of  $B/T$  with radius. The whole dynamic range of the median  $B/T$  from low to high density environments is  $\sim 0.2$  to  $0.4$ .

Fig. 5.2 shows the morphology-environment relation for our sample of 1281 cluster galaxies as the median (as well as 0.05, 0.315, 0.685, and 0.95) quantiles of the galaxy distribution over  $B/T$  in several bins in  $R_\sigma$ . There is a clear correlation of morphology with environment, even though the shift in the median  $B/T$  is relatively small compared to the typical range of  $B/T$  at each radius. Standard errors on the quantiles are given according to Kendall & Stuart (1977) as

$$SE = \sqrt{\frac{p(1-p)}{nf_1^2}}, \quad (5.17)$$

where  $p$  is the desired quantile,  $n$  is the sample size, and  $f_1$  the frequency of data points per unit interval of the dependent variable near the quantile.

The total rank correlation coefficient for the correlation shown in Fig. 5.2 is

$$r_{B/T, R_\sigma} = -0.215 \ (Z = 7.8\sigma).$$

In analogy to Fig. 5.2, Fig. 5.3 shows star formation (as measured by  $EW([OII])$ ) as a function of  $R_\sigma$ . There is a highly significant gradient in star formation, with the median  $EW([OII])$  in the cluster center being smaller by several Å than at the outskirts. For small radii ( $R_\sigma < 0.5$  Mpc), the median  $EW([OII])$  is almost zero. The correlation coefficient is

$$r_{OII, R_\sigma} = 0.295 \ (Z = 10.9\sigma).$$

Note that our radial sampling limit only extends to  $\sim 1 - 2$  virial radii, and therefore does not cover the outer limit at which the star formation rate is claimed to reach the level of the field galaxy population (Gómez et al., 2003; Lewis et al., 2002).

### 5.6.2 Residual Star Formation Versus Environment

We have verified above that our sample exhibits the morphology-environment relation as well as the star formation-environment relation. Now we determine



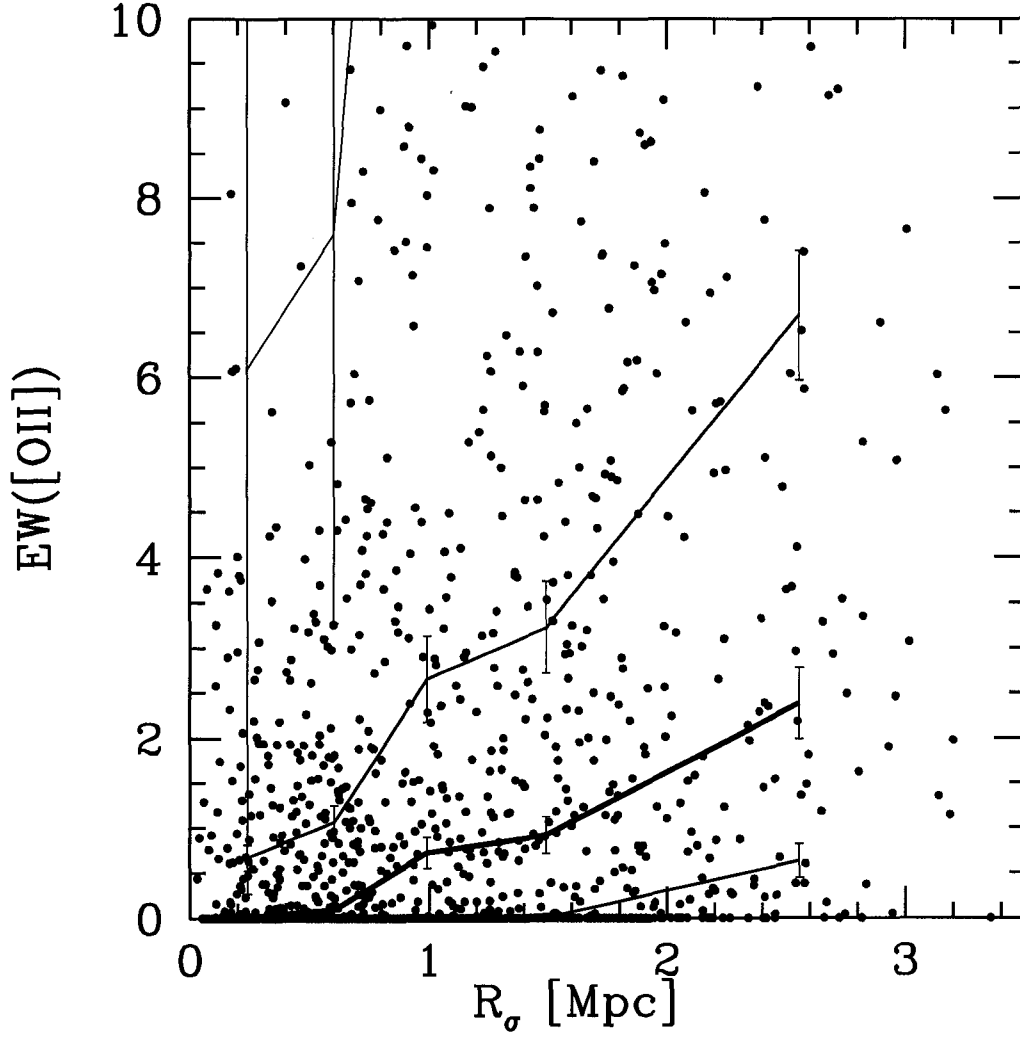


Figure 5.3: Star Formation-Environment Relation. The figure shows the (0.0485,0.315,0.5,0.685,0.9515) quantiles of the  $EW([OII])$  distribution as a function of cluster-centric radius,  $R_\sigma$  (the 0.05-quantile coincides with the  $EW([OII]) = 0$  axis and is therefore not visible in the plot). The median is represented by the bold line. There is a significant correlation of  $EW([OII])$  with radius.

the residual star formation-environment correlation while controlling all other variables (morphology, stellar mass, mean stellar age). This approach will show whether the star formation rate varies as a function of its environment even if morphology, stellar mass, and stellar age are fixed.

#### 5.6.2.1 Residual Correlation of Current Star Formation with Environment

We calculate the partial rank correlation coefficient between  $EW([OII])$  and  $R_\sigma$  while controlling the three variables  $B/T$ ,  $D4000$ , and  $M_*$  for all galaxies with  $M_R \leq -19.2$ . With 1281 galaxies, the correlation coefficient is

$$r_{OII, R_\sigma, B/T, M_*, D4000} = 0.221 \ (Z = 8.0\sigma).$$

This highly significant residual correlation means that, even for galaxies with similar morphology (as expressed by  $B/T$ ), stellar mass, and mean stellar age (as expressed by  $D4000$ ), there remains a star formation gradient with environment. This result indicates that: (1) The star formation gradient is not just another aspect of the morphology-environment relation (or of the stellar mass-environment or mean stellar age-environment relations). Therefore, the properties of galaxies must be the result of more than one formation or subsequent evolutionary mechanism, or the same mechanism affects morphology and star formation differently in different environments. (2) The star formation gradient is not just a result of any environmental variations in the galaxy mass function that result in a higher fraction of (star-forming) dwarf galaxies on the cluster outskirts. (3) The star formation gradient is not due simply to the increasing number of long-dead galaxies towards the cluster center. This last finding, in particular, argues that the star formation gradient is not solely the product of initial conditions, but that an environmental mechanism affects the later evolution of galaxies.

To visualize the residual correlation of star formation with environment, we

calculate a residual  $\Delta EW([OII])$  for each galaxy with the following procedure.

We determine the linear partial regression coefficients

$b_{OII,B/T,M_*,D4000,R_\sigma}$ ,  $b_{OII,D4000,M_*,B/T,R_\sigma}$ , and  $b_{OII,M_*,B/T,D4000,R_\sigma}$

in rank space (Kendall & Stuart, 1977) according to

$$b_{1,j,2,3,\dots,k} = -\frac{\sigma_1}{\sigma_j} \frac{C_{1j}}{C_{11}}. \quad (5.18)$$

The variable  $b_{1,j,2,3,\dots,k}$  is the partial regression coefficient of 1 on  $j$ , and  $\sigma_j$  is the square root of the variance of variable  $j$ . The notation indicates that variables 2, 3, ...,  $k$  are fixed.

The data can then be represented as

$$< X_i > = \overline{X_i} + \sum_j b_{1,j,2,3,\dots,k} (X_j - \overline{X_j}), \quad (5.19)$$

where  $\overline{X_i}$  is the mean of variable  $i$ . In our case, we use this formula to calculate an expectation value for  $rank(EW([OII]))$  for each galaxy  $i$ :

$$\begin{aligned} < rank(EW([OII])) > |_i = \overline{rank(EW([OII]))}_i + \\ & (rank(B/T_i) - \overline{rank(B/T)}) b_{OII,B/T,D4000,M_*,R_\sigma} + \\ & (rank(M_{*,i}) - \overline{rank(M_*)}) b_{OII,M_*,D4000,B/T,R_\sigma} + \\ & (rank(D4000) - \overline{rank(D4000)}) b_{OII,D4000,M_*,B/T,R_\sigma}. \end{aligned} \quad (5.20)$$

Note that the indices still represent the ranked variables, not their absolute values. We choose to calculate linear regression coefficients in rank space because the linearity condition is usually much better fulfilled there.

We use the observed relation between  $EW([OII])$  and  $rank(EW([OII]))$  to convert the result into an expectation value for  $EW([OII])$ . This procedure allows us to perform the linear regression in rank space, but to obtain an expectation value in absolute parameter space. We then calculate  $\Delta EW([OII]) =$

$EW([OII]) - \langle EW([OII]) \rangle$ , plotting it and the (0.05;0.315;0.5;0.685;0.95) quantiles of the  $\Delta EW([OII])$  distribution in Fig. 5.4.

There is a clear residual correlation with environment. Up to a radius of  $R_\sigma \approx 1$  Mpc, the median residual  $EW([OII])$  remains fairly constant, indicating that morphology, stellar mass, and mean stellar age can account for variations of the median  $EW([OII])$  over this range. However, for larger radii, the median varies as a function of environment. The (0.685) quantile varies even more steeply with  $R_\sigma$  than the median, indicating that the properties of strongly star-forming galaxies are particularly correlated with environment.

The sense of the correlation shows that, for a galaxy with a given morphology, stellar age, and mean stellar mass, star formation decreases with decreasing cluster-centric radius.

#### 5.6.2.2 Residual Correlation of Recent Star Formation with Environment

Our analysis above has placed constraints on the impact of environment on current star formation. Since environmental mechanisms can be differentiated by the timescales on which they operate on a galaxy, it is of considerable interest whether this effect acts not only on current, but on recent star formation as well.

As discussed in §5.4.1, we use  $EW([H\delta])$  in absorption as an index of recent star formation. We perform this analysis only on a subset of galaxies with no significant [OII] emission. This constraint removes 198 [OII] emitters from the sample, leaving 1083 galaxies. We proceed in the analysis as for  $EW([OII])$ . There is a very weak, but still detectable correlation; the Spearman correlation coefficient is

$$r_{H\delta, R_\sigma} = 0.084 \ (Z = 2.8\sigma).$$

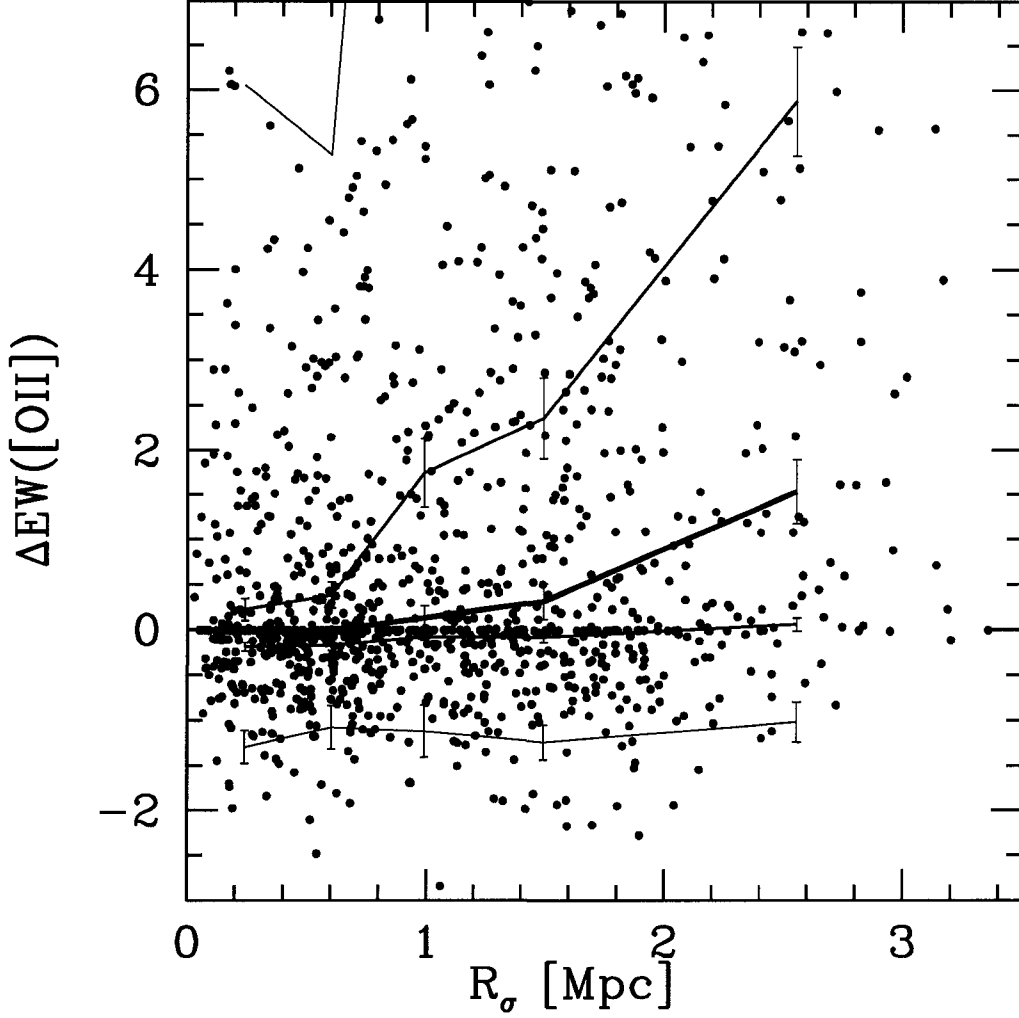


Figure 5.4:  $\Delta EW([OII])$ - $R_\sigma$  Relation. Dots show the *residual*  $EW([OII])$  after subtracting the expectation value for a given  $B/T$ ,  $D4000$ , and  $M_*$ . Lines show the (0.05;0.315;0.5;0.685;0.95) quantiles in each bin, with the median being marked by the bold line. There is a residual correlation, indicating that morphology, stellar mass, and mean stellar age cannot account for the star formation gradient in clusters.

If we remove the effects of  $B/T$ ,  $D4000$ , and  $M_*$ , the residual correlation coefficient is

$$r_{H\delta, R_\sigma, B/T, M_*, D4000} = 0.064 (Z = 2.1\sigma).$$

Fig. 5.5 shows the residual correlation after subtracting the expectation value  $\langle EW(H\delta) \rangle$  for a given  $(B/T, D4000, M_*)$  (calculated analogously to Eq. 5.20).

Although this result indicates a significant residual correlation between environment and recent star formation, observational uncertainties in the variables that we are controlling have not been included in this estimate yet, and we briefly comment on them in §5.6.3.

#### 5.6.2.3 What is the effect of controlling individual variables on the star formation gradient?

In this section, we discuss how strongly the star formation gradient is affected by controlling each of the three individual variables —  $B/T$ ,  $M_*$ , and  $D4000$  — individually. We have seen in §5.6.2 that controlling  $B/T$ ,  $M_*$ , and  $D4000$  simultaneously does indeed reduce the correlation coefficient for the star formation-environment relation, indicating that they do contribute to the observed star formation gradient. Which of these three variables is the most important contributor? We address this question by re-calculating the residual correlation coefficient while holding only one variable fixed at a time.

First of all, controlling stellar mass has very little effect on the star formation gradient. If we control only  $M_*$ , the correlation is reduced from  $r_{OII, R_\sigma} = 0.295$  ( $Z = 10.9\sigma$ ) to  $r_{OII, R_\sigma, M_*} = 0.275$  ( $Z = 10.1\sigma$ ). This is because there is only a weak (although significant) segregation of stellar mass with environment ( $r_{M_*, R_\sigma} = -.125$ ). In this context, we point out that, although there is a segregation of stellar mass with environment, we do not find a significant segregation of

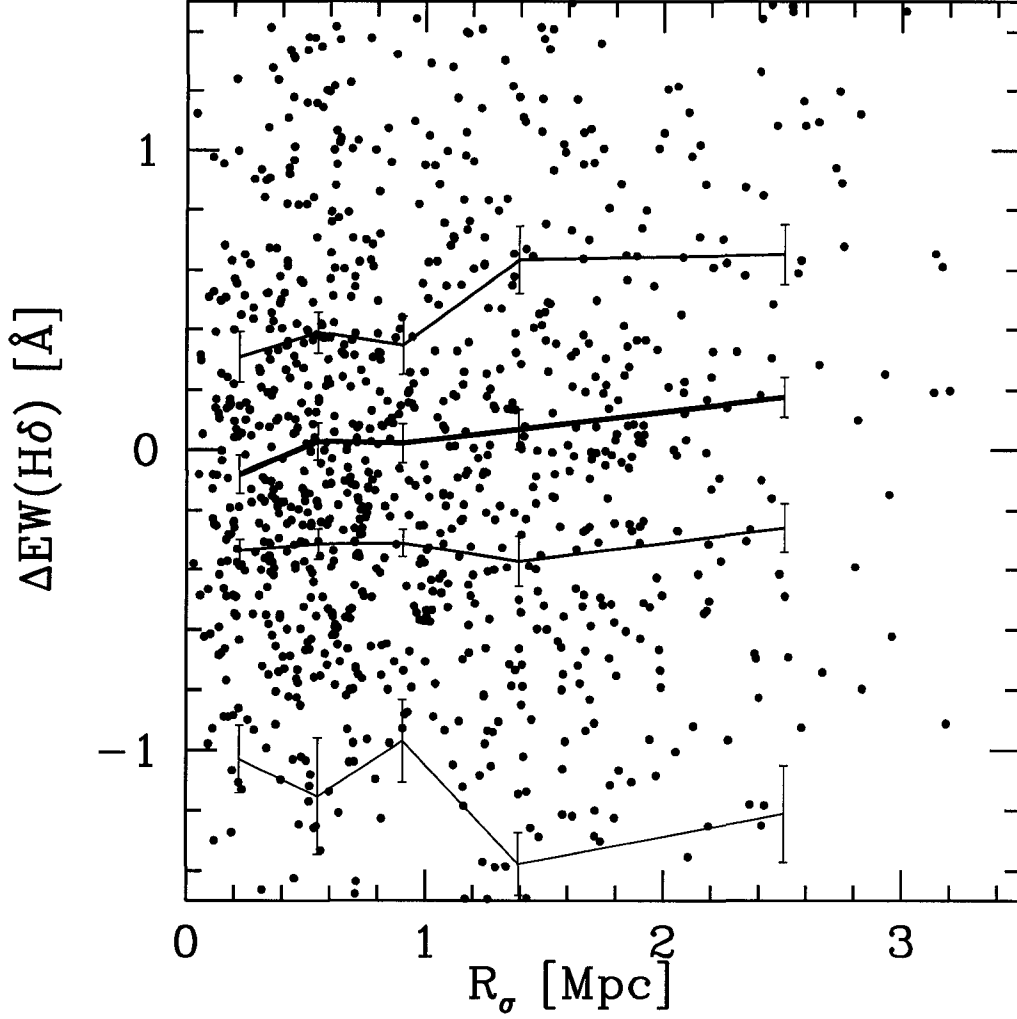


Figure 5.5: *Residual* absorption  $EW[H\delta]$ - $R_\sigma$  relation after removing the effects of  $B/T$ ,  $D4000$ , and  $M_*$  for non-[OII]-emitting galaxies. Solid lines show the (0.05;0.315;0.5;0.685) quantiles of the  $EW([H\delta])$  distribution as a function of cluster-centric radius,  $R_\sigma$ . The 0.95 quantile is off the scale and has very large error bars, and is therefore omitted from this plot. There is a significant residual correlation ( $2.1\sigma$ ) of recent star formation with radius.

*R*-band luminosity with environment. Possibly, the gradient in color is masking the gradient in stellar mass. This result must be kept in mind when searching for luminosity segregation in clusters: even a negative result does not imply an absence of stellar mass segregation.

Second, controlling  $B/T$  has the strongest effect on the star formation gradient, indicating that the morphology-environment relation does contribute significantly to the overall star formation gradient with environment. Controlling only for  $B/T$ , we find  $r_{OII, R_\sigma, B/T} = 0.228$  ( $Z = 8.3\sigma$ ).

Third, controlling  $D4000$  also removes some of the star formation gradient:  $r_{OII, R_\sigma, D4000} = 0.260$  ( $Z = 9.5\sigma$ ). However, the impact is rather weak, compared to that of controlling  $B/T$ , showing that the stellar age gradient is not an important contributor to the star formation gradient. Furthermore, by controlling  $B/T$ , we are already accounting for most of the effect of the  $D4000$  gradient, because the residual  $D4000$  gradient when controlling  $B/T$  is only  $r_{D4000, R_\sigma, B/T} = -0.066$  ( $Z = 2.4\sigma$ ), as opposed to the total correlation coefficient of  $r_{D4000, R_\sigma} = -0.148$  ( $Z = 5.3\sigma$ ). This result indicates that morphology and stellar age are closely related and possibly controlled by the same initial conditions and/or environmental factors.

We have mentioned in §5.4.4 that  $D4000$  is biased low by current star formation. What is the impact of this bias on the  $D4000$  gradient? If we remove all galaxies with current star formation from the sample, the gradient of  $D4000$  with radius will be more representative of the real stellar age gradient in the cluster. After removing all galaxies with significant ( $> 2\sigma$ )  $[OII]$  emission, we find a correlation of  $D4000$  with radius of  $r_{D4000, R_\sigma} = -0.023$  ( $Z = 0.7$ ), statistically indistinguishable from zero. Furthermore, when we control  $EW([OII])$ , the  $D4000$  gradient also vanishes ( $r_{D4000, R_\sigma, OII} = -0.039$   $Z = 1.4\sigma$ ). Therefore, rather than the  $D4000$  gradient accounting for the star formation gradient, the star formation



gradient may be the dominant cause of the observed  $D4000$  gradient, which reassures us that the correlation of star formation with environment cannot be a consequence of stellar age differences.

#### 5.6.2.4 Is the residual star formation gradient a major component of the total star formation gradient?

In this section, we examine whether the detected residual correlation of star formation and environment is a major component of the total star formation gradient. A measure of the actual strength of the residual correlation, compared to the total correlation, is provided by the regression coefficient. The partial regression coefficient in rank space is

$$b_{OII, R_\sigma \cdot B/T, D4000, M_*} = 0.192,$$

compared to the total star formation gradient of

$$b_{OII, R_\sigma} = 0.288.$$

In other words, two thirds of the total star formation gradient, on average, is apparent even within a population of galaxies with the same morphology. Thus, the morphology-independent component of the star formation gradient is a major contributor to the total star formation gradient.

The effect of environment is also appreciable compared to the effect of morphology, as the comparison to the partial regression coefficient with  $B/T$  as the independent variable shows:

$$b_{OII, B/T, R_\sigma, D4000, M_*} = -0.286.$$

We therefore conclude that, although the morphology-environment relation contributes significantly to the star formation gradient in clusters, and morphology itself has a stronger impact on star formation than radius alone, most of the

star formation gradient remains even after accounting for the effect of morphology. This morphology-independent star formation gradient is therefore a major contributor to the total star formation gradient, and the star formation properties in clusters cannot be understood simply as a result of the morphology-environment relation and/or initial conditions.

### 5.6.3 Effect of Errors and Uncertainties on the Residual Correlations

In this section, we examine how strongly measurement uncertainties in our indices of morphology, stellar mass, and mean stellar age affect the residual correlation coefficients. Statistical uncertainties in a variable weaken any total correlations with this variable. This compromises our efficiency at removing the effect of the variable in the calculation of a residual correlation coefficient and will generally cause us to overpredict the significance of a residual correlation coefficient.

We have seen above that the morphology, stellar mass, and stellar age do not vary strongly enough with environment and/or star formation to account for the observed star formation gradient with environment. Is that only because our measurements of these variables are affected by large errors? To investigate this possibility, we have to quantify the impact of measurement uncertainties on the residual correlation coefficients. We focus on  $B/T$  first, because, for faint galaxies, errors in  $B/T$  are of the same order of magnitude as  $B/T$  itself.

How much are total and partial correlation coefficients affected by observational uncertainties in one variable? Let us assume that  $x$  and  $y$  are the “true” variables, unaffected by observational uncertainties (e.g., that  $y$  is the true bulge fraction), and that  $y'$  is the observed variable (i.e., the observed  $B/T$ ). For simplicity, we neglect errors in  $x$ . The partial correlation coefficient when controlling

for only one variable can be written out as (Kendall & Stuart, 1977)

$$r_{xy'.y} = \frac{r_{xy'} - r_{xy}r_{yy'}}{\sqrt{(1 - r_{xy}^2)(1 - r_{yy'}^2)}}. \quad (5.21)$$

Algebraic transformation yields the “ideal” correlation coefficient between the variables that are unaffected by observational uncertainties:

$$r_{xy} = \frac{r_{xy'}}{r_{yy'}} - \frac{r_{xy'.y}}{r_{yy'}} \sqrt{(1 - r_{xy}^2)(1 - r_{yy'}^2)}. \quad (5.22)$$

If  $y'$  is merely a noisier version of  $y$ , then partialing out the “ideal” variable  $y$  will account for any correlations with  $x$  and leave no residual correlation between  $x$  and  $y'$ , i.e.,  $r_{xy'.y} = 0$ . In that case, the “ideal” correlation coefficient is

$$r_{xy} = \frac{r_{xy'}}{r_{yy'}}. \quad (5.23)$$

This relation defines a weakening factor  $\xi_y \equiv r_{yy'}$  that describes how much observational uncertainties in the variable  $y$  have reduced a given total correlation coefficient. We determine  $\xi_y$  empirically by calculating the correlation coefficient between two different realizations — e.g., two choices of  $B/T$  for each galaxy within the GIM2D uncertainty interval — of the variable  $y$ . In our data, we find  $\xi_{B/T} = 0.907 \pm 0.008$ , i.e., all total correlation coefficients that contain the observable  $B/T$  are about 10% weaker than they would be if we had a perfect measure of bulge fraction. By scaling all total correlation coefficients that contain  $B/T$  by  $\xi_{B/T}^{-1}$ , we can thus recover the original strength of the correlation coefficient.

There is a potential source of systematic errors in the total correlation coefficient between  $B/T$  and  $EW([OII])$  as well. Aperture bias could introduce a negative correlation, and thus enhance the existing intrinsic correlation between morphology and star formation. The spectrograph used for the spectroscopic survey has a fiber diameter of  $3''$ , which corresponds to a physical size of 0.7 kpc at the distance of A1060 and 3.4 kpc at the distance of A3266. In galaxies with larger

bulges, the fiber captures less light from the disk, and more from the quiescent bulge component. This bias could artificially lower the observed star formation rate.

Given that the observed  $r_{OII,B/T}$  is too *weak* to explain the star formation gradient, our claim of a residual correlation appears conservative. Nonetheless, we test whether this bias is a serious concern. We calculate the correlation coefficient of  $EW([OII])$  with angular diameter distance,  $D_A$ , while controlling the bulge absolute magnitude,  $M_{bulge}$ . Thus, we are comparing whether  $EW([OII])$  varies as a function of distance if we hold the bulge luminosity (which we assume to be correlated with bulge size) fixed. If aperture bias is a problem, then for a given bulge size,  $EW([OII])$  should increase with increasing distance as the fiber captures more light from the disk. The correlation coefficient is  $r_{OII,D_A.M_{bulge}} = +0.038$ , ( $Z = 1.4\sigma$ ). This effect has almost no impact on the observed correlations: the coefficient for the  $EW([OII]) - B/T$  correlation after removing  $D_A$  is  $r_{OII,B/T.D_A} = -0.442$  at  $17\sigma$ , practically identical to the total correlation coefficient  $r_{OII,B/T} = -0.443$ . Therefore, aperture bias does not significantly affect our results.

Our determination of the stellar mass may also be affected by observational uncertainties. The reliability of our stellar mass values can be gauged by comparing the stellar mass estimates based on different filter bands. In our sample, correlation coefficients between our stellar mass values based on  $R$ -,  $J$ -, and  $K$ -band photometry are typically on the order of  $r = 0.95$ . We therefore adopt this as the weakening factor  $\xi_{M_*} = 0.95$ , i.e., we assume that all total correlation coefficients containing  $M_*$  are  $\sim 5\%$  weaker than they would be if we had a perfect measure of stellar mass.

With the last of the three variables that we are controlling,  $D4000$ , there are

two sources of uncertainty: one is measurement uncertainty, which we treat in the same way as the measurement uncertainty in  $B/T$  and obtain  $\xi_{B/T}^{measure} = 0.98$ . The second source of uncertainty is whether the observable  $D4000$ , even in the absence of measurement errors, is representative of the variable that we really want to control, the mean stellar age. We have already estimated the correlation coefficient between  $D4000$  and the mean stellar age, ( $MSA$ ) to be  $r_{D4000,MSA} = 0.91$  (§5.4.4, and we thus adopt the product of  $\xi_{B/T}^{measure}$  and  $r_{D4000,MSA}$  as our total weakening factor  $\xi_{D4000} = 0.89$ .

However, galaxies with even moderate quantities of recent or on-going star formation have substantially lower  $D4000$  than the correlation  $r_{D4000,MSA}$  above would imply for their stellar age. How strong is this effect, and could it generate an artificial residual star formation gradient with environment? If we regard the mean stellar age as the “ideal” variable, and  $D4000$  as the corresponding observable, then Eq. 5.22 gives us:

$$r_{OII,MSA} = \frac{r_{OII,D4000}}{r_{D4000,MSA}} - \frac{r_{OII,D4000,MSA}}{r_{D4000,MSA}} \sqrt{(1 - r_{OII,MSA}^2)(1 - r_{D4000,MSA}^2)}. \quad (5.24)$$

In our case, we are concerned about  $D4000$  being biased low by current star formation, which means that  $r_{OII,D4000,MSA} < 0$ . Therefore,  $r_{OII,D4000}/r_{D4000,MSA}$  is a (negative) lower limit on the correlation coefficient  $r_{OII,MSA}$ . This result, in turn, means that our claim of a significant residual correlation coefficient is conservative.

What is the numerical effect of all these statistical and systematic uncertainties on our residual correlation coefficient? When we correct the individual total correlation coefficients by the inverse of the weakening factors, we obtain a resulting residual correlation coefficient of

$$r_{OII,R_{\sigma},B/T,D4000,M_{\star}} = 0.192 \pm 0.037.$$

Because the residual correlation coefficient after applying our correction factors does not obey the usual probability distribution implied by Eq. 5.5 anymore, we have propagated the uncertainties in the individual correlation coefficients into the partial correlation coefficient using a Monte Carlo algorithm. For this reason, we explicitly quote uncertainties for all residual correlation coefficients that we calculate using  $\xi$ -corrections, instead of providing a level of significance from Eq. 5.5. Nonetheless, even with all corrections applied, this result is only slightly smaller than the uncorrected one ( $r = 0.221$ ) and is evidence of a significant residual correlation.

For the observed residual correlation between  $EW(H\delta)$  and  $R_\sigma$ , we have repeated this analysis, correcting for the observational uncertainties in  $B/T$ ,  $D4000$ , and  $M_*$ . We find

$$r_{H\delta, R_\sigma, B/T, D4000, M_*} = 0.056 \pm 0.052.$$

This result suggests, conservatively, that the residual correlation of  $EW(H\delta)$  with  $R_\sigma$  observed earlier at  $> 2\sigma$  may not be significant.

One last concern: the partial correlation coefficient is calculated from the total correlation coefficients, and not from the properties of the individual galaxies directly. The total correlation coefficients encompass only a limited range of information about the relation between two variables (i.e., they characterize variations of the median, as compared to the variance). In some variables, such as  $EW([OII])$  (which is non-negative, but has many outliers at large values), the median does not characterize the galaxy distribution well. Therefore, information is lost in the calculation of the total and, consequently, partial correlation coefficients. The partial correlation coefficient between star formation and projected radius does not test whether individual galaxies show an excess or deficit of star formation over an expectation value based on their morphology, mean

stellar age, and luminosity, but only whether the median variations of all these quantities can explain the median variation of star formation with environment.

Is it possible to estimate whether this loss of information affects our conclusions? We do so with regard to the most important of the three variables that we are controlling,  $B/T$ . We define a new variable,  $\delta EW([OII])$ , which we obtain by subtracting from the  $EW([OII])$  of each individual galaxy the median  $EW([OII])$  for galaxies with a similar  $B/T$ . The variable  $\delta EW([OII])$  characterizes the actual excess star formation of a galaxy, relative to the expectation value from its  $B/T$ . We then apply our correlation analysis to  $\delta EW([OII])$  instead of  $EW([OII])$ . This procedure is more stringent than the standard correlation analysis alone, because it can account better for outlying data points (i.e., a strongly star forming early-type galaxy will influence the correlation coefficient differently than a strongly star forming late-type galaxy).

We then find a partial correlation coefficient of

$$r_{\delta OII, R_{\sigma}.B/T, D4000, M_{*}} = 0.198 \ (Z = 7.2\sigma),$$

compared to the original estimate of  $r_{OII, R_{\sigma}.B/T, D4000, M_{*}} = 0.221$ . Therefore, the loss of information due to the assumptions that go into the calculation of the correlation coefficient (i.e., linearity, normality) does not compromise our conclusions.

## 5.7 Conclusions

We have analyzed the photometric and spectroscopic properties of a sample of galaxies in six nearby clusters. Our primary aim has been to investigate whether the star formation-environment relation (Lewis et al., 2002; Gómez et al., 2003) can be explained as an aspect of the morphology-environment relation (Dressler, 1980), of a stellar age gradient, or of mass segregation within the cluster, or

whether there is evidence that the star formation rates of two galaxies with the same morphology, stellar mass, and mean stellar age vary with environment. Our results are based on a partial rank correlation analysis, which avoids the problems usually associated with binning data, namely, that the finiteness of bin widths may leave residual correlations within each bin.

We find that there is a correlation of current star formation with environment even for galaxies with comparable morphologies, stellar masses, and mean stellar ages. We conclude the following: (1) The star formation gradient is not just another aspect of the morphology-environment relation (or of the stellar mass-environment or stellar age-environment relations). This result indicates that the properties of galaxies must arise from more than one formation or subsequent evolutionary mechanism whose efficiencies are dependent on environment, or that the same mechanism affects morphology and star formation differently in different environments. (With respect to the latter case, Christlein & Zabludoff (2004) show that the morphology-environment relation depends primarily on enhancing the bulge luminosity in denser environments, as opposed to fading the disks of galaxies. Their result thus favors galaxy-galaxy interactions and mergers as the most likely transformation mechanism.) (2) The star formation-environment relation is not simply a result of samples in denser environments being biased (by initial conditions or subsequent evolution) towards more massive galaxies. (3) The star formation-environment relation is not solely due to the fact that the centers of clusters are generally dominated by older galaxies. This last result rules out that the star formation gradient was already seeded in initial conditions during the epoch of galaxy formation. The star formation gradient must thus be due in some part to late evolution.

By taking into account variations of morphology, stellar mass, and stellar age



with environment, and nonetheless observing a residual correlation of current star formation with environment, our work constitutes clear statistical evidence for a substantial ongoing effect of environment on galaxy evolution— a true “smoking gun.”

Although our sample consists of cluster galaxies, it is possible that the environmental mechanisms whose signature we have observed are not specific to clusters. Because of the near-degeneracy of local density with radial distance from the cluster center in our sample, we cannot determine whether the residual star formation gradient is primarily controlled by local density, and may therefore also occur in lower-density environments such as poor groups, or by radial distance, which would suggest that the mechanism or mechanisms responsible for this gradient are specific to the cluster environment.

## CHAPTER 6

### CONCLUSIONS

This thesis has presented a statistical investigation of the photometric, spectroscopic, and environmental properties of galaxies in six nearby, rich clusters of galaxies (Christlein & Zabludoff, 2003). With this dataset, we have provided new constraints on how galaxy properties vary with their environment and what factors might be responsible for this variation.

We have introduced a new method for recovering the true galaxy distribution function from an incomplete sample (Christlein, McIntosh & Zabludoff, 2004). This discrete maximum likelihood method (DML) is ideally suited for multivariate statistical investigations of large surveys, because all necessary corrections for sample incompleteness are incorporated in weighting factors associated with individual galaxies. The algorithm is therefore independent of the dimensionality of the galaxy parameter space that we wish to investigate, and unlike other maximum likelihood estimators used in studies of galaxy surveys, does not require us to define the analytical form of the solution a priori.

Based on this survey, we have, for the first time, calculated *R*-band luminosity functions for cluster galaxies from a spectroscopic sample, and performed an internally consistent comparison to the luminosity function of field galaxies and among the luminosity functions of clusters (Christlein & Zabludoff, 2003). We have also compared luminosity functions determined from different filter bands to examine how color gradients due to star formation, dust, or metallicity affect the determination of luminosity functions (Christlein, McIntosh & Zabludoff,

2004).

Furthermore, we have performed bulge-disk decompositions and calculated the luminosity functions for bulges and disks directly (as opposed to inferring them indirectly from the total luminosity function and an estimate of the variation of the bulge fraction with luminosity, as some past studies have done). We have used this information to disentangle how the properties of bulges and disks vary with morphology and environment and thus place constraints on the mechanisms that generate the morphological sequence of galaxies (Christlein & Zabludoff, 2004).

Finally, we have integrated photometric and spectroscopic information into a multi-variate analysis designed to address the shortcomings of past studies of the star formation-environment relation (Christlein & Zabludoff, 2004b). Previous work has not clarified whether this relation is degenerate with the morphology-environment, the stellar mass-environment, or the stellar age-environment relations, and thus, whether it is merely the result of the same processes of formation or evolution that set up any of these gradients. Nor has it clarified whether the observed star formation gradient with environment is a signature of late environmental transformations or a result of initial conditions. By holding morphology, stellar mass, and mean stellar age fixed while looking for a residual correlation between star formation and environment, we can address both questions.

Our results and conclusions are:

- The *R*-band luminosity function in clusters is very similar to that in the field, with the exception of an overabundance of galaxies at the most luminous tip of the luminosity function. There is also no evidence for variations of the luminosity function among different clusters, despite their covering almost an order of magnitude in virial mass. However, the luminosity func-

tion of quiescent galaxies shows striking variations between the field and clusters, and there is evidence that it is also not universal among clusters.

These results indicate that any physical mechanisms underlying the correlations between morphology and environment have no strong impact on the average  $R$ -band luminosity of galaxies. However, the strong variation in the fraction of quiescent galaxies between the field and clusters is most likely due to the impact of environment on the star formation rates of galaxies. An alternative possibility is that the cluster environment may be conducive to the formation of large numbers of faint quiescent galaxies, but such a scenario would have to be fine-tuned to maintain a constant faint end slope of the overall luminosity function.

- The  $U$ -band luminosity function in clusters is slightly steeper than that of the  $R$ -band luminosity function, but the variation of the faint end slope with filter band is not drastic. A simple separation into quiescent and star-forming galaxies cannot explain the slight steepening of the faint end slope. This means that either metallicity, dust, or star formation gradients that are not accounted for by this separation introduce a color gradient with luminosity. Despite this sensitivity to star formation, dust, or metallicity, the dependence on the filter band is so weak that the luminosity function still provides fair constraints on the mass function.
- Our analysis of the variation of bulge and disk luminosities as a function of morphology and environment shows that the morphological sequence of galaxies is generated by variations of the bulge luminosity. Furthermore, bulge luminosities are systematically higher in dense environments, whereas disk luminosities do not show such a dependence.

The effect of any morphological transformations on the luminosity function is not strong: we detect only a weak (but significant) correlation of bulge luminosity with radius, and no significant correlation of total luminosity with radius (which is partly due to the fact that the redder color of galaxies in denser environments masks a weak, but significant trend of stellar mass with environment). This explains why, despite the increase of bulge luminosity in denser environments, the luminosity functions of the field and clusters are very similar.

These results strongly favor bulge-enhancing processes, such as galaxy-galaxy interactions or mergers (Barnes, 1999; Bekki, 1998; Mihos & Hernquist, 1994), rather than disk fading processes, such as ram-pressure stripping (Gunn & Gott, 1972) or strangulation (Larson, Tinsley & Caldwell, 1980; Balogh, Navarro & Morris; Bekki, Couch & Shioya, 2002), as the mechanism that gives rise to intermediate- and early-type galaxies, including S0s, from late-type galaxies.

- There is a residual correlation of star formation with environment even if we hold morphology, stellar mass, and mean stellar age fixed. This allows two conclusions: (1) As we are fixing three variables that could have been set by initial conditions and still find a residual star formation gradient with environment, this gradient must be due at least in part to environmental mechanisms acting during the later evolution of a galaxy. (2) The star formation gradient with environment must be due either to several mechanisms whose relative efficiencies vary with environment, or one mechanism that affects star formation differently in different environments.

These primary purpose of our work is to provide new constraints on the vari-

ations of fundamental statistical properties of galaxies with environment. Although our sample consists of cluster galaxies, the cluster environment need not be the driving mechanism behind the correlations that we have presented here (Whitmore & Gilmore, 1991; Whitmore, Gilmore & Jones, 1993). Because of the near-degeneracy of local density with radial distance from the cluster center in our sample, we cannot determine whether the residual star formation gradient is primarily controlled by local density, and may therefore also occur in lower-density environments such as poor groups, or by radial distance, which would suggest that the mechanism or mechanisms responsible for this gradient are specific to the cluster environment.

Further statistical studies of this or other samples may provide additional valuable constraints that directly address the nature of the transformation mechanisms. What other effects — besides star formation — does the mechanism responsible for the residual star formation correlation with environment have on a galaxy? And in which environments does it act? Ram pressure stripping and strangulation would predict a positive correlation between disk luminosity and *residual* star formation. On the other hand, correlations between the residual star formation and bulge luminosity could indicate that the residual star formation gradient is caused by central star formation being triggered by the galaxy-galaxy interactions (Larson & Tinsley, 1978; Barton Gillespie, Geller & Kenyon, 2003) favored by lower-density environments.

Correlations with environment will also provide important clues as to the nature of the transformation mechanisms involved. Ram-pressure stripping (Gunn & Gott, 1972), unlike strangulation (Larson, Tinsley & Caldwell, 1980; Balogh, Navarro & Morris; Bekki, Couch & Shioya, 2002), is predicted to be efficient only in the densest cores of clusters. In order to distinguish between these two possi-

bilities, a crucial next step is therefore to establish whether the residual correlation of star formation with environment is unique to the cluster environment, or whether it is observable in lower-density environments, such as groups, as well. Furthermore, direct comparisons between the galaxy populations in groups with those in subclumps infalling into clusters will provide constraints on the impact of the cluster environment on galaxy properties.

More insight into this problem will come from a VLA neutral hydrogen survey (van Gorkom et al., 2003) of the clusters in our sample. These data will allow us to examine the impact of environment on the neutral hydrogen content of galaxies while still holding morphology, stellar mass, and mean stellar age fixed. Among the questions that we will be able to address is whether star formation rates vary with environment because the HI content (for a galaxy with a given morphology and stellar mass) is lower in dense environments, or because different environments trigger star formation with different efficiencies. In addition, we will examine whether the HI content varies with environment if we hold morphology fixed, which will help distinguish between different mechanisms that could affect the HI content (e.g., galaxy-galaxy interactions or ram-pressure stripping). The survey may also detect tidal features indicative of ongoing interactions or ram pressure stripping by the intracluster medium.

## APPENDIX A

### THE *R*-BAND PHOTOMETRIC DETECTION CATALOG

An accurate survey of galaxy number counts requires a reliable detection algorithm that does not generate spurious detections but still reaches to faint surface brightness limits, good star/galaxy classification, and accurate photometry. For this reason, we decided on an approach combining automatic object detection and extraction with extensive manual control and corrections.

The detection surface brightness threshold is low, particularly on images with higher backgrounds and thus lower S/N, thus generating many spurious detections. In addition, the clusters considered in this paper are at low redshifts, and many of their member galaxies appear as large and highly structured objects. This complicates automatic detection mechanisms such as those provided by Source-Extractor, which is more suited to higher redshift regimes. There are two major concerns. The first is the appropriate choice of the background mesh size used by Source Extractor to compute the background count level. Too large a mesh size will not be able to detect small-scale background variations, such as might be caused by scattered light from bright galactic stars. Too small a background mesh, on the other hand, tends to overestimate the background level at the positions of very extended objects. The second problem consists of the accidental blending of fainter objects with nearby bright sources – giant galaxies or bright stars. When run with default parameters, Source Extractor sometimes associates detections with bright objects, even in entirely different parts of the frame. On the other hand, lowering the deblending threshold risks splitting up



some of the rather structured galaxy images.

We thus employ a strategy allowing for a maximum of individual control and human intervention. First, we generate four Source Extractor output catalogs. In addition to the default catalog with standard parameters and a background mesh size of 64 pixels, a second catalog uses a larger mesh size of 200 pixels. Photometry for very extended objects (larger than 5000 pixels in isophotal area, consistent with the threshold and mesh sizes used by Tran et al. (2001) and Zabludoff & Mulchaey (2000)) is automatically drawn from this second catalog. We generate a third catalog with a low deblending threshold and small mesh size of 32 to detect small scale fluctuations that may be superimposed on – and consequently confused with – larger extended objects. A fourth catalog is generated from the cosmic-ray-free frame.

We then cross-correlate the four catalogs, associating detections in each of the four catalogs with their counterparts in the other catalogs, if present. We remove any detection that does not have a counterpart on the cosmic-ray-free image and whose combination of magnitude and surface brightness is characteristic of cosmic rays. We earmark detections that meet only one of these criteria for a visual inspection (such detections are typically spurious – plate flaws or scattered starlight – or grazing cosmic rays).

At this point, we identify objects that may have been accidentally blended with brighter sources by the following procedure: We assume that, among the three catalogs, all individual objects in the frame meeting our magnitude and surface brightness limits have been detected. If we find that a detection does not have a direct counterpart in another catalog, but resides in the area occupied by another object in that catalog, the first detection is earmarked as a potential child object, and the other detection as its parent. This could either be a genuine

subdetection within a brighter and more extended object, such as an HII region in a large spiral, or it could be a separate object in need of deblending. To determine these potential parent-child hierarchies, we make use of the segmentation images provided by Source-Extractor. For each of our catalogs, they show which pixels of the input image have been assigned to which object.

We use two automatic methods to classify each object as a star or galaxy. The first of these is the stellarity index provided by Source-Extractor. This is a fractional value between 0 and 1, generated by a neural network algorithm, that is ideally 1 for a star and 0 for a galaxy. The second mechanism is based on the fact that stars and galaxies occupy two distinct regions in the plot of magnitude versus surface brightness. We visually estimate a linear separatrix between the two regions. We consider objects with lower isophotal surface brightness for a given isophotal magnitude than indicated by this separatrix are galaxies. Variations in the seeing and other effects on the image quality (such as imperfect tracking) lead us to separately determine parameters for the seeing, the star/galaxy threshold, and the separatrix for each individual frame to optimize the agreement between both methods.

We generate a catalog of detections that are then subjected to a visual inspection. In each cluster, we perform a full review on a number of frames (several for each night, especially such where star/galaxy separation may be problematic due to bad seeing or tracking), including the central frame. We inspect every detection down to  $m_R = 19.75$  visually (minus the cosmic rays that have been removed automatically). We also perform a limited review on the remaining frames, ignoring the objects that have been classified as stars by both automatic classifiers and reviewing only the earmarked detections plus all galaxies. This approach is justifiable, as the star classifier is usually robust, and the full reviews confirm that

almost none of the stellar-appearing objects would be reclassified as galaxies in a visual inspection.

During the review, we give every inspected object a third, manual classification as a star, galaxy or defect. In addition, we inspect all detected potential child-parent hierarchies and break them up manually if this appears justified. We also have the option of manually associating two separate detections that have not been classified as parent and child.

By comparing the visual appearances of detections of similar magnitudes, we catch blatant systematic errors in the photometry. While the photometry obtained from Source-Extractor is usually reproducible within our quoted errors, galaxies superimposed on bright galaxy envelopes or on scattered light from galactic stars are almost invariably estimated to be too bright. We perform individual photometry on these cases, using a custom-made algorithm that models non-uniform backgrounds over a circular region, and adopt these magnitudes over those from Source-Extractor. In a very small number of cases in which the close proximity of highly structured objects makes a direct determination impossible, we estimate the magnitude visually. Of 25211 reviewed galaxy detections (4251 of which are above our standard analysis threshold of  $m_R = 18$ ), we estimate 34 magnitudes (none of them at  $m_R \leq 18$ ). From visual comparisons to galaxies of similar appearance, we estimate the errors on these magnitudes to be  $\approx \pm 0.25$  mag. In addition, in cases where stars are so closely superimposed on galaxies that the two have not been detected separately, we manually add the galaxy to the catalog (71 occurrences, 20 of them at  $m_R \leq 18$ ). We usually estimate magnitudes for these galaxies as described above. All objects for which we remeasure or estimate magnitudes are marked in our catalog. Altogether, such objects constitute 307 detections, only 10 of them at  $m_R \leq 18$ .

Although we do not inspect and classify all stars individually, we inspect every detection (including stars) in a collection of thumbnail images in a second pass through the catalog, thus enabling us to account for most misclassifications of galaxies as stars and for star/galaxy blendings as well.

If the manual corrections to the parent-child hierarchy require an adjustment of the photometry, it is automatically performed by adding or subtracting isophotal fluxes, depending on whether child objects have been deblended or added and whether this affects the magnitude that enters the final catalog. At this stage, we take all other manual corrections into account as well, adding the manual classification as a third star/galaxy classifier to the two automatically generated flags, removing defects, and generating a final catalog.

Despite our careful approach, we may have missed a small number of galaxies, predominantly because they have been blended with closely superimposed or particularly bright galactic stars. An estimate of this effect is provided by cross-correlating the catalog of spectroscopic targets with our detection catalog at a later stage. Typically, we do not detect  $< 1\%$  of the spectroscopically confirmed galaxies. We trace some of these cases to multiple spectroscopic targetings of the same galaxy, and the rest to the glare of bright foreground stars. We then add these galaxies to the catalog manually, if that has not happened during the normal review process.

The cross-correlation with the spectroscopic catalog also provides a constraint on the star/galaxy misclassifications. We conclude that this effect is less than 1% of all spectroscopic objects. Most of the misclassifications are associated with star/galaxy superpositions, and only an extremely small number (on the order of one or two objects per cluster) represent real misclassifications – usually moderately bright, compact galaxies that had been misclassified as stars.

We have tested the effect of uncertainties in the star/galaxy classification on some of our GLFs by alternatively including either all detections not unambiguously identified as stars or all detections clearly identified as galaxies. The GLFs produced by these two extremes are statistically indistinguishable from the standard GLFs. The first scenario steepens the faint end slope of the field GLF by  $\Delta\alpha \approx -0.02$  and of the cluster GLF by  $\Delta\alpha \approx -0.06$ , the second scenario induces changes of  $\Delta\alpha \approx +0.02$  and  $\Delta\alpha \approx +0.03$ , respectively. The reason that this effect is minor — even under such worst-case scenarios — is that the sample is dominated by galaxies in regions of the  $(m_R; \mu_R)$  plane that are unambiguously non-stellar.

## APPENDIX B

### CORRECTION OF $f_s$ FOR SELECTION IN A DIFFERENT FILTER BAND

We selected the spectroscopic targets for our survey using approximate  $m_B$  magnitudes, while our photometry and analysis use  $m_R$  magnitudes throughout. While target selection is probably representative of cluster membership for a given  $m_B$ , a color difference  $\Delta(B - R)$  between field and cluster galaxies (Pimbblet et al., 2001) would create a bias in  $f_s(m_R; \mu_R)$ . A cluster galaxy of a given  $(m_R; \mu_R)$  would have a different  $m_B$ , and thus a different probability of having been selected as a spectroscopic target, than a field galaxy. Therefore, our sampling fraction, determined as the average ratio of spectroscopically sampled galaxies to photometric detections in one  $(m_R; \mu_R)$  bin, would not be representative of the cluster and field populations at  $(m_R; \mu_R)$ . We quantify this effect by choosing a mean color difference  $\Delta(B - R)$  between field and cluster galaxies and assuming that, for a given  $(m_B; \mu_B)$ , the targetting probability  $p^B$  of a given galaxy is independent of cluster membership. From this assumption, it follows that

$$\begin{aligned} p_f^R(m_R) &= p^B(m_R + (B - R)_f) = p^B(m_R - ((B - R)_{cl} - (B - R)_f) + \\ &\quad + (B - R)_{cl}) = p_{cl}^R(m_R - \Delta(B - R)) \end{aligned} \quad (\text{B.1})$$

where  $p_f^R$  and  $p_{cl}^R$  are the probabilities of a field or cluster galaxy being targetted as a function of R magnitude (and surface brightness, which we do not write out here explicitly), and  $\Delta(B - R) = (B - R)_{cl} - (B - R)_f$ .

We now introduce following notation:  $p_{ave}^- \equiv p_{ave}(m - \Delta(B - R))$  is the mean targetting probability (irrespective of population) at magnitude  $m_R - \Delta(B - R)$ ,

and  $p_{ave}^0$  is the corresponding targetting probability at  $m_R$ .

Our aim is to obtain an expression for  $p_f^0$  and  $p_{cl}^0$ , *i.e.* the field- and cluster-specific targetting probabilities, assuming a fixed mean color difference between these two populations. If the targetting function can be approximated to first order then the field galaxy targetting probability is,

$$p_f^0 = 1/2(p_f^+ + p_f^-) \quad (B.2)$$

To fill in the quantities on the right side, we solve the expressions for  $p_{ave}^0$  and  $p_{ave}^-$  for  $p_f^+$  and  $p_f^-$ , respectively. We start by writing,

$$p_{ave}^0 = \frac{N_{det,cl}^0 p_{cl}^0 + N_{det,f}^0 p_f^0}{N_{det}^0} = \frac{(N_{det}^0 - N_{det,f}^0) p_f^+ + N_{target,f}^0}{N_{det}^0} = \frac{(N_{det}^0 - \frac{N_{target,f}^0}{p_f^0}) p_f^+ + N_{target,f}^0}{N_{det}^0} \quad (B.3)$$

In a similar manner, we find

$$p_{ave}^- = \frac{(N_{det}^- - \frac{N_{target,f}^-}{p_f^-}) p_f^0 + N_{target,f}^-}{N_{det}^-} \quad (B.4)$$

We solve these two expressions for  $p_f^+$  and  $p_f^-$ , respectively, and plug them into (B2) to obtain

$$2p_f^0 = \frac{N_{target,cl}^0}{N_{det}^0 - \frac{N_{target,f}^0}{p_f^0}} + \frac{N_{target,f}^-}{N_{det}^- - \frac{N_{target,cl}^-}{p_f^0}} \quad (B.5)$$

We solve for  $p_f^0$ , using the positive-root solution for the quadratic formula that we obtain in the process.

For  $p_{cl}^0$ , we analogously obtain the equation

$$2p_{cl}^0 = \frac{N_{target,cl}^+}{N_{det}^+ - \frac{N_{target,f}^+}{p_{cl}^0}} + \frac{N_{target,f}^0}{N_{det}^0 - \frac{N_{target,cl}^0}{p_{cl}^0}} \quad (B.6)$$

This expression is undefined if the  $(m; \mu)$  plane does not contain detections in regions that are referenced by expressions such as  $N_{det}^-$  and  $N_{det}^+$ . To avoid

this problem, we ignore the gradient of the sampling fraction along the  $\mu$  axis and determine  $p_f^0$  and  $p_{cl}^0$  as functions of  $m_R$  alone (which is justifiable, as the dependence on  $\mu$  is much weaker than on  $m$ ). We then scale all sampling fractions within the corresponding magnitude bin by the value  $p_f^0/p_{ave}^0$  (for the field) or  $p_{cl}^0/p_{ave}^0$  (for the clusters) to obtain the corrected sampling fractions for field and cluster galaxies.

As this correction is ultimately dependent on the choice of the average  $\Delta(B - R)$  between field and cluster galaxies, we choose not to apply it throughout, but merely use it to estimate the magnitude of the effect. We estimate  $\Delta(B - R)$  as a function of  $m_R$  from a comparison between our R-band photometry and the approximate  $b_J$ -band magnitudes that were used for target selection. For a given  $m_R$ , the average field galaxy is actually slightly redder than the average cluster galaxy, presumably because a field galaxy in our sample is, on average, more distant and thus more massive than a cluster galaxy of the same apparent magnitude. We do not quote the numerical values for  $\Delta(B - R)$  here because they are not accurate photometric colors, but they are on the order of a few tenths of a magnitude for brighter galaxies ( $m_R < 16$ ), and the difference is smaller at the faint end ( $m_R = 18$ ). This is encouraging, because it indicates that color terms are large only where the sampling is uniformly complete (bright apparent magnitudes), and small where the sampling fraction gradient is strongest (faint  $m_R$ ). Therefore, there are no significant biases in the sampling fraction either for bright or faint galaxies. When correcting for the bias, we find a minute change in the slope of the cluster GLF of  $\Delta\alpha < 0.02$  and of the field GLF of  $\Delta\alpha < 0.01$ , much smaller than the statistical errors. The changes in  $M_R^*$  are of the same order of magnitude.



## APPENDIX C

### HIGHER-ORDER CORRECTIONS IN CALCULATING THE SAMPLING FRACTION

By determining the sampling fraction on a discrete grid in bins with finite widths in  $(m, \mu)$  space, we neglect potential higher-order variations of the sampling fraction on scales of the bin width. The mean sampling fraction within one bin is only representative of the sampling fraction at the bin center if the densities of photometric detections and spectroscopically sampled galaxies do not vary more steeply than to first order. To estimate the impact that higher-order variations may have on our determination of the sampling fraction, and consequently on our luminosity function, we attempt to model their distribution within the bins to second order in  $m$ . We restrict ourselves to variations in  $m$ , which greatly simplifies the calculation and still accounts for the strongest variations of the sampling fraction.

We empirically determine the second-order moments of the galaxies relative to the bin centers and use them to constrain the  $0th$  order coefficient of the Taylor-expansion of the detection density. We then use these coefficients, representing the detection densities at the bin center, rather than the total counts of detections across the bin, to calculate the sampling fraction.

We first Taylor-expand the detection density as  $\rho = \sum_n a_n \delta m^n$  and define the  $m - th$  order moment of the galaxy distribution in one bin as

$$p_m = \int_{-\Delta m/2}^{+\Delta m/2} \rho(\delta m) \delta m^m d\delta m, \quad (\text{C.1})$$

where  $\delta m$  is the magnitude relative to the bin center. The coefficient representing

the detection density at the bin center is then given to second order as

$$a_0 = (c_{22}p_0 - c_{20}p_2)/(c_{00}c_{22} - c_{20}c_{02}) \quad (\text{C.2})$$

where  $c_{mn}$  are the coefficients of the  $m - th$  moment of the Taylor expansion, *i.e.*,  $c_{mn} = (\Delta m/2)^{n+m+1}/(n+m+1)$  if  $n+m = \text{even}$  and 0 otherwise.

Empirically, we calculate  $p_m = \sum_{bin} \delta m^m$ .

As in the case of the color selection correction, we do not apply this correction to our calculations (a drawback of using higher-order moments is increased noise), but only use it to estimate the magnitude of this effect. The change in the faint end slope of the field composite GLF is negligible, on the order of  $\Delta\alpha \approx 0.01$ .

## APPENDIX D

### THE DISCRETE MAXIMUM LIKELIHOOD METHOD

Our aim is to recover the galaxy luminosity function, i.e., the parent distribution from which those galaxies with both redshifts and *U*-band photometry in our sample have been drawn. Reconstructing the parent distribution from a set of sampled galaxies requires two corrections. First, a volume correction is needed to account for the fact that bright galaxies are over-represented because they can be observed to greater distances and therefore over a larger volume. This is not a problem in volume-limited surveys (e.g., of only one cluster, where all galaxies are at approximately the same distance), but it is a concern whenever a GLF is derived from a magnitude-limited sample spanning a range in redshift. For example, in our case, a galaxy of a given absolute magnitude may be observable in one cluster, but not in the two other, more distant ones. Second, a completeness correction is necessary because not all galaxies that are photometrically detected in our *R*-band master catalog have been sampled, i.e., have both spectroscopic and *U*-band photometric information. There are various reasons why photometrically detected galaxies are not part of the sample: some faint galaxies were not targeted for spectroscopic observations, a small number were targeted unsuccessfully, and others were not detected on the *U*-band images. The completeness of the sample relative to the *R*-band master catalog varies primarily as a function of apparent magnitude and surface brightness.

GLFs for individual clusters are usually calculated by simply weighting each galaxy by the inverse of its sampling probability, which is either known (in a

strictly magnitude-limited survey) or can be recovered empirically (by comparing the number of spectroscopically sampled galaxies of a given magnitude to the total number of photometric detections). Composites are usually calculated by scaling and averaging the GLFs from several individual clusters. However, this method does not make optimal use of the information in a survey. The maximum likelihood solution for the parent distribution function is not only determined by the fact that a galaxy has been observed in one cluster, but also by the question whether or not it would have been observable in any of the other clusters in the sample.

A method that is typically used for field galaxies, but that could be adapted for a sample of clusters, is the  $V/V_{max}$  method, which weights each individual galaxy by the inverse of the volume over which it could have been observed in the survey. This method has two drawbacks: First, it implicitly assumes that the galaxy space density distribution is homogeneous. This is primarily a problem for the field GLF, but if applied to clusters, it also requires introducing an additional scaling to account for the fact that different clusters have different numbers of member galaxies. The second problem is that the method is impractical if the sample is not constrained by a fixed limiting magnitude, but by a fractional completeness that varies with apparent magnitude, which is the case for our cluster data. The definition of the “volume” over which a galaxy would have been observable in such a survey is not straightforward.

The most popular algorithms for calculating luminosity functions from a sample spanning a range in redshift are based on maximum likelihood (ML) methods. ML algorithms assume a parent distribution characterized by a limited number of parameters. Taking the selection criteria (sample completeness and magnitude limits) of the survey into account, they then calculate the probability that the ob-

served sample has been drawn from this assumed parent, and iteratively adjust the parameters to maximize this probability. ML methods are preferred because they are not dependent on assumptions about the redshift distribution of galaxies in space and are therefore unbiased by density inhomogeneities due to large scale structure. This advantage applies particularly to determinations of the field GLF, but ML algorithms are applicable to every galaxy sample regardless of the redshift distribution, including clusters.

Several variants of the ML method exist that differ in the way they represent the parent distribution. The parametric maximum likelihood method (PML) by Sandage, Tamman & Yahil (1979) characterizes the parent as a Schechter function (Schechter, 1976) with three free parameters: the shape parameters  $M^*$  and  $\alpha$  and the normalization constant  $\phi^*$ . The disadvantages of this method are that assumptions regarding the shape of the parent are required, and more complex, multi-dimensional analyses are impossible because the Schechter function is a function of only one variable, absolute magnitude. Cross & Driver (2002) have expanded the PML to include absolute magnitude and surface brightness.

Efstathiou, Ellis & Peterson (1988), in the stepwise maximum likelihood (SWML) approach, parametrize the galaxy parent distribution with binned distributions, a method that could, in principle, be expanded to an arbitrary number of dimensions (for an example of an adaptation of the SWML method to two-dimensional distributions, see Christlein & Zabludoff (2003)). However, in a sample of moderate size such as ours, with only a few hundred galaxies, the advantages of binning galaxies are negated by shot noise if the number of bins approaches the number of galaxies in the sample (although the effects of shot noise can be reduced by projecting the distribution back onto a subspace of lesser dimensionality).

Both the PML and SWML methods suffer from the fact that a parametrized

form of the GLF has to be assumed *a priori*, and that variables that are not explicitly represented by this form are discarded. They are thus not ideal for multi-dimensional analyses, i.e., for recovering the galaxy distribution function (GDF), which, in analogy to the univariate GLF, describes the abundance of galaxies as a function of multiple variables such as their luminosity in different filter bands, surface brightness, star formation indices, environment, morphology, etc.

In our case, the fact that we are modeling the sample completeness as a function of  $m_R$  and  $\mu_R$ , while plotting the luminosity function as a function of  $M_U$ , requires us to keep track of a minimum of three variables ( $M_U, M_R, \mu_R$ ) and to treat the galaxy distribution function as a function of these (note that in this context  $\mu$  is measured in the galaxy rest frame). The PML method is clearly ruled out for our purposes because there is no generally accepted functional form to describe the distribution of galaxies in  $(M_R, m_U, \mu_R)$  space.

While the SWML could, in principle, be applied to this problem, it is not very efficient, particularly with regard to future work aimed at expanding our analysis of the galaxy distribution function to higher-dimensional parameter spaces (including such variables as star formation, morphology, and local environment). It either requires modifications to the algorithm to change the *ansatz* for the galaxy distribution function every time a different cut through the distribution function is to be made, or the distribution has to be binned *a priori* in as many dimensions as photometric, spectroscopic and morphological parameters are available for each galaxy. Both approaches are computationally cumbersome, because the *ansatz* has to be hardwired into the algorithm and modified every time the dimensionality of the distribution function is changed. Furthermore, there is no advantage to binning a galaxy distribution in so many dimensions that many bins hold only a very small number of galaxies. In such cases, higher-order mo-

ments of the galaxy distribution within a single bin can become important, so that the mean properties of the bin do not accurately represent the properties of the individual galaxies contained therein.

We have therefore developed a new variant of the maximum likelihood algorithm that does not require any assumptions about the functional form of the galaxy distribution prior to the calculation, nor does it require binning the galaxies or sacrificing any dimensions of the available parameter space in the sample.

In our algorithm, the GDF is represented by the sampled galaxies themselves, and not by bins or functional parameters that only model certain aspects of the sampled galaxy population. Our approach can be regarded as an extreme application of the SWML method for infinitesimally small bins in parameter spaces of arbitrary dimensionality. We formally describe the GDF as a sum of weighted delta functions. It is nonzero at all coordinates in parameter space where our sampled galaxies lie. For this reason, we refer to it as the Discrete Maximum Likelihood method (DML) throughout this paper. The *ansatz* for the GDF is:

$$\varphi(\vec{x}) = C \sum_n \omega_n \delta(\vec{x}; \vec{x}_n), \quad (\text{D.1})$$

where  $C$  is a normalization constant and  $\vec{x}_n$  is a parameter vector for galaxy  $n$  of arbitrary dimensionality. In our case, the parameter space will be  $(M_R, M_U, \mu_R)$ . The weighting factors  $\omega_n$  are free parameters to be determined by the DML algorithm. We use the nomenclature  $\varphi$  to indicate that this *ansatz* can be generalized to arguments  $\vec{x}$  of any dimensionality. This *ansatz* is similar to that of the  $C$  method (Lynden-Bell, 1971; Choloniewski, 1986), but the procedure for solving for the free parameters,  $\omega$ , is different, with our method retaining the benefits of ML estimators.

The decisive advantage of this method over the SWML method is that, instead of absorbing the individual galaxies into an *ansatz* for the GDF that repre-

sents only a limited number of galaxy properties, it associates a weighting factor with each individual galaxy. The free parameters remain tied to the individual galaxies, rather than to fixed grid points in a pre-selected subspace of parameter space. Therefore, it provides a completeness and volume correction, but still retains the full range of photometric, spectroscopic and morphological properties of the galaxy for a subsequent analysis. Computationally, the algorithm is independent of the dimensionality of the distribution function that is to be analyzed, and only needs to be adjusted to account for the correct dependencies of the sampling fraction. This combines the advantages of the  $V/V_{max}$  method with those of the ML methods. The algorithm is also very simple to implement computationally.

The ansatz in Eq. D.1 is not of course a realistic physical representation of the parent distribution function. After calculating the weighting factors, the distribution can be smoothed or binned, or treated with other multi-variate analysis techniques. In contrast to conventional techniques, these procedures do not have to be hardwired into the algorithm *a priori*, but can be applied after the completeness and volume corrections have been carried out, allowing us to perform a multitude of statistical analyses on a data set without having to find a new maximum likelihood solution for each analysis.

The derivation of an algorithm to solve for the free parameters  $\omega$  is analogous to the SWML method. We start with the probability that a galaxy  $i$  with parameters  $\vec{x}_i$  would have been observed in the survey:

$$p_i = \frac{\varphi(\vec{x}_i) f(\vec{x}_i | F_i)}{\int \varphi(\vec{x}) f(\vec{x} | F_i) d\vec{x}} = \frac{\left( \sum_g \omega_g \delta(\vec{x}_i; \vec{x}_g) \right) f(\vec{x}_i | F_i)}{\int \left( \sum_g \omega_g \delta(\vec{x}; \vec{x}_g) \right) f(\vec{x} | F_i) d\vec{x}}, \quad (\text{D.2})$$

where  $F_i$  describes the field of galaxy  $i$  and includes redshift, distance, Galactic extinction, position in the sky, and other information.  $f(\vec{x}_i | F_j)$  is therefore the



probability that a galaxy with properties  $\vec{x}_i$  would have been included in the sample if it was in the position of galaxy  $j$ . All indices refer to specific members of the sample of both spectroscopically and  $U$ -band photometrically sampled galaxies, i.e.,  $\omega_h$  is the weighting factor associated with galaxy  $h$ . The  $\delta$  function now allows us to consider the function only at discrete points and thus solve the integral:

$$p_i = \frac{\left(\sum_g \omega_g \delta(\vec{x}_i; \vec{x}_g)\right) f(\vec{x}_i | F_i)}{\sum_g \omega_g f(\vec{x}_g | F_i)}. \quad (\text{D.3})$$

We now form the composite probability to have obtained the observed sample from the assumed distribution and take its logarithm:

$$\ln \mathcal{L} = \sum_i \ln(\sum_g \omega_g \delta(\vec{x}_i; \vec{x}_g)) + \sum_i \ln f(\vec{x}_i | F_i) - \sum_i \ln \sum_g \omega_g f(\vec{x}_g | F_i). \quad (\text{D.4})$$

Taking the derivative by  $\omega_h$ , we find

$$\frac{\partial \ln \mathcal{L}}{\partial \omega_h} = \sum_i \frac{\delta(\vec{x}_i; \vec{x}_h)}{\sum_g \omega_g \delta(\vec{x}_i; \vec{x}_g)} - \sum_i \frac{f(\vec{x}_h | F_i)}{\sum_g \omega_g f(\vec{x}_g | F_i)} = \frac{1}{\omega_h} - \sum_i \frac{f(\vec{x}_h | F_i)}{\sum_g \omega_g f(\vec{x}_g | F_i)}. \quad (\text{D.5})$$

Setting this expression to 0 yields an expression for  $\omega_h$  that is suitable for an iterative solution:

$$\omega_h = \left( \sum_i \frac{f(\vec{x}_h | F_i)}{\sum_g \omega_g f(\vec{x}_g | F_i)} \right)^{-1}. \quad (\text{D.6})$$

This algorithm converges very quickly. To obtain an absolute normalization factor, we integrate the GDF down to  $M_U = -17$ . We repeat this integration for each galaxy in the sample, applying its visibility conditions  $F_i$  to predict how many galaxies of a given  $M_U$  and  $\mu_R$  would have been sampled, had they been at the position of this galaxy. Thus, we derive a prediction for the total number of galaxies with  $M_U < -17$  that we would expect to have sampled, and compare it to the actual number to derive a normalization factor:

$$C = N_{\text{sampled}} \left( \sum_i N_{cl(i)}^{-1} \sum_j \omega_j f(\vec{x}_j | F_i) \right)^{-1}, \quad (\text{D.7})$$

where  $N_{cl(i)}$  is the number of galaxies sampled in the cluster to which galaxy  $i$  belongs,  $N_{sampled}$  is the total number of galaxies in both the spectroscopic and  $U$ -band sample, and  $f(\vec{x}_j | F_i)$  is the probability that galaxy  $j$  would have been sampled in the survey field and at the redshift and coordinates of galaxy  $i$ .

Once the weighting factors have been calculated, it is possible to apply techniques of multi-variate analysis and plot different cuts through the GDF over any desired variable, or to restrict the analysis to subsets of galaxies, without having to modify the algorithm or recalculate the weighting factors. In principle, GLFs in the  $U$ -band and the  $R$ -band can be plotted from the same output, using the same weighting factors and normalization constants. Furthermore, a sample can be split (e.g., by emission line properties), and the two distributions can be plotted and compared separately; the weighting factors do not have to be recalculated. Therefore, correct relative normalization between the two subsamples is automatically guaranteed.

Care has to be applied in cases where, due to survey design, the sampling fraction model does not fully account for all dependences of the sample completeness. In our study, spectroscopic targets have been selected by approximate  $b_J$  magnitudes. This introduces a residual dependence of the sampling fraction  $f$  on  $b_J - R$  colors that is not accounted for in our approach, which only models  $f$  as a function of  $(m_R, \mu_R)$ . If a subset of galaxies is selected by parameters that are correlated with  $b_J - R$  colors, such as cluster membership or emission line properties, biases can result, because the mean empirical sampling fraction calculated for galaxies of a given  $(m_R, \mu_R)$  may be systematically different from the true sample completeness for the selected subset of galaxies. We resolve this problem by applying systematic corrections specific to the individual subsamples (particularly the EL and NEL subsamples) to the sampling fraction, which

requires us to recalculate the weighting factors for these subsets. These corrections are discussed in App. E.

For a similar reason, when calculating the GDF for individual clusters out of a sample of multiple clusters, it is not legitimate to first calculate the weighting factors for the entire sample and then just plot the galaxies associated with one particular cluster, because the weighting factors determined that way would contain volume corrections that are appropriate to a sample spanning a range in redshift, but not to the sample of an individual cluster. For a given absolute magnitude, the sampling fraction is different from cluster to cluster. Therefore, the weighting factors have to be recalculated for each individual cluster in order to analyze individual cluster GDFs. This is not a drawback of the DML method, but common to all ML methods.

When the recovered value of the GDF is exactly zero in any region of parameter space (for example, for very faint  $M_R$ ), it is necessary to understand whether this is because the sampling fraction is so low that none of a potentially large population of galaxies have been sampled, or because there are no galaxies with these properties. Again, this is a problem common to all LF algorithms, but a careful treatment is particularly important in the DML, because it represents the GDF solely with sampled galaxies and therefore, by default, does not probe the sampling fraction outside of the sampled regions of parameter space (whereas other methods like the SWML will return an undefined result if the sampling fraction is zero).

The sampling probability for a galaxy of any given absolute magnitude, surface brightness, etc., can be probed in the DML by introducing mock galaxies. The critical value for the sampling probability can be defined in the following way: Consider the the GDF in two dimensions,  $M_U$  and  $M_R$ , and define a bin in

$M_U$ . The number of galaxies in the well-sampled parts of the galaxy distribution in this bin, which is represented by the reconstructed GLF, is  $N_0$ . Further, assume that there is a region of parameter space within this  $M_U$  bin with a low sampling probability. The number of galaxies in this region is  $N$ , and the sampling probability for such a galaxy is  $f$ . The poorly sampled region is a problem only if a) the expected number of sampled galaxies in this region is less than 1, i.e.  $Nf < 1$ , and b) the number of galaxies that exist in this region is a substantial fraction  $p$  of the total number of galaxies in this  $M_U$  bin, i.e.  $N > p(N + N_0) > pN_0$ . This yields the condition for poor sampling,  $pN_0 < f^{-1}$  (note that this expression is independent of the size of the hypothetical poorly sampled region). The reconstructed luminosity function is therefore not affected by a systematic relative error of more than  $p$  as long as the sampling fraction  $f$  is greater than the critical value of  $(pN_0)^{-1}$  everywhere within this bin. By introducing mock galaxies with the given  $M_U$ , a value of  $\mu_R$  characteristic for galaxies with this  $M_U$ , and arbitrary  $M_R$ , this requirement can be checked easily and the effective limits of the survey in  $M_R$  determined.

Once we have found the effective survey limits (for example, in  $M_R$ ), it is necessary to ascertain whether it is likely that substantial fractions of the galaxy distribution lie outside these limits. For example, for any  $M_U$ , there are extreme values of  $M_R$  for which no galaxies could have been sampled spectroscopically. However, there is a fairly tight correlation between  $M_U$  and  $M_R$ , and for bright  $M_U$ , it is extremely unlikely that any galaxies reside beyond the faint effective limit in  $M_R$ . In §3.4.4, we use the procedure outlined above to define the effective absolute magnitude limits in  $M_R$  and  $M_U$  for our sample.

## APPENDIX E

### DETERMINATION OF THE SAMPLING FRACTION

This section describes the empirical determination of the sampling fraction,  $f(\vec{x}_h | F_i)$ , which is the probability that we have both  $U$ -band photometry and spectroscopic information for a given object with certain physical properties  $\vec{x}_n$  (e.g., absolute magnitude and surface brightness) if it is in a particular field  $F_i$  (characterized by redshift, Galactic foreground extinction, position on the sky). We determine the completeness of the  $U$ -band and spectroscopic catalogs relative to the  $R$ -band photometric catalog. We assume the  $R$ -band catalog to be complete for  $m_R < 18$  and  $\mu_R < 23.61$  (the approximate equivalent of a per-pixel detection threshold of  $1.5\sigma$  on a typical frame). This assumption is conservative; in fact, we detect objects more than five magnitudes fainter, but do not include them in the analysis. We cannot detect objects at lower surface brightness, but at  $m_R < 18$ , the distribution of photometric detections in the  $(m_R, \mu_R)$  plane is well separated from this limit, indicating that few objects are likely to be lost to the surface brightness limit.

Following Bayes’s theorem (Bayes, 1764), we factorize the sampling fraction — the probability that, for a given object, redshift information ( $cz$ ) and  $U$ -band photometry ( $U$ ) are available on condition that the object meets our sample selection criteria. These selection criteria, denoted by  $sel$ , are the criteria that we apply to our galaxy sample to select those galaxies that are to be included in our GLFs. This is primarily cluster membership. For some analyses, we also select galaxies by their emission line properties, and comment on this at the end of this section.

The expression for the conditional sampling probability is then

$$p(cz \wedge U \mid sel) = p(cz \wedge U \wedge sel)/p(sel). \quad (\text{E.1})$$

We now transform Eqn. E.1 into a form that is suitable for numerical evaluation and factorize it into separate terms for the spectroscopic and the  $U$ -band photometric completeness. There are two straightforward factorizations that allow us to separate terms connected to the spectroscopic and the  $U$ -band sampling fractions:

$$p(cz \wedge U \mid sel) = p(U \mid cz \wedge sel) \times p(cz \mid sel). \quad (\text{E.2})$$

or

$$p(cz \wedge U \mid sel) = p(U) \times p(cz \wedge sel \mid U) \times p(cz \mid sel)/p(cz \wedge sel). \quad (\text{E.3})$$

The separation of terms for the spectroscopic and  $U$ -band photometric sampling fractions is motivated by the opportunity to apply specific, systematic corrections to either. It is only these corrections that introduce any difference between the two methods, but the differences are statistically insignificant ( $< 1\%$  even in the faintest magnitude bins). We use the first factorization because of its greater simplicity.

Both the spectroscopic and  $U$ -band sampling fractions cannot be described analytically in our case. We thus have to determine them empirically by comparing the number counts of all detections in the  $R$ -band photometric master catalog with those with redshifts and  $U$ -band photometry. Because the  $R$ -band photometric catalog is complete, we calculate the sampling fraction as a function of  $m_R$  and  $\mu_R$ . These two parameters alone do not allow for an unambiguous description of the sampling fraction; color and/or environment also play a role. We discuss below how we address this problem. The bin widths that we use to calculate the sampling fraction are  $\Delta m = 0.75$  and  $\Delta \mu = 0.25$ . We have experimentally

found that these choices of bin widths reproduce stable results for the resulting GLFs.

$p(cz \mid sel)$  cannot be determined exactly, because information about cluster membership is only available for spectroscopically sampled objects. The assumption  $p(cz \mid sel) \approx p(cz)$  is only justified if the sampling fraction is not strongly correlated with the selection criteria for a given  $(m_R, \mu_R)$ . Unfortunately, the experimental design of our survey — target selection in  $b_J$ , compilation of a master catalog in  $R$  — violates that assumption, because there are systematic differences in  $B - R$  colors between cluster and field galaxies, and therefore systematic differences in the spectroscopic sampling probability for a given  $(m_R, \mu_R)$ . This color selection bias could be remedied if the sampling fraction model explicitly contained color terms, but it is impractical to estimate the sampling fraction as a function of more than two parameters with a sample the size of ours.

Instead, we use an algorithm that systematically compensates for color selection bias and estimates  $p(cz \mid sel)$  under the assumption that the targeting probability is uncorrelated with the selection criteria for a given  $m_{b_J}$ . The probability that a galaxy of a given  $m_R$  obeying the selection criteria has been targeted is

$$\begin{aligned}
 p(target \mid sel) \mid m_R = & (N_{target,nonsel}^0 N^+ + N_{target,sel}^+ N^0 + 2N^0 N_{target,nonsel}^+ + \\
 & 2N_{target,sel}^0 N^+) + ((N_{target,nonsel}^0 N^+ + N_{target,sel}^+ N^0 + 2N^0 N_{target,nonsel}^+ + \\
 & 2N_{target,sel}^0 N^+)^2 - 8N^0 N^+ (2N_{target,sel}^0 N_{target,nonsel}^+ + N_{target,nonsel}^0 N_{target,nonsel}^+ + \\
 & N_{target,sel}^0 N_{target,sel}^+))^{1/2} / 4N^0 N^+. \quad (E.4)
 \end{aligned}$$

Here,  $N$  denotes the number of galaxies that have been photometrically detected and thus included in the  $R$ -band master catalog.  $N_{target}$  is the number of galaxies targeted for spectroscopic observations. The subscript  $sel$  indicates galaxies that are to be included in the GLF (e.g., that are cluster members). The superscript 0

indicates that these quantities are evaluated at  $m_R$ , and the superscript  $+$  means that they are evaluated at  $m_R + \Delta(B - R)$ .  $\Delta(B - R)$  is the difference in B-R colors between galaxies that have been selected (e.g., that are cluster members), and those that have not. We roughly estimate the  $m_R$ -dependent differences in  $\Delta B - R$  between cluster and field galaxies from the sampled galaxies themselves.

In all cases,  $b_J$  is not an accurate photometric magnitude, but rather the approximate magnitude that served as the basis for the target selection for the spectroscopic survey. Furthermore, for this derivation we assume  $N_{target,sel} \approx N_{target}N_{cz,sel}/N_{cz}$ , i.e., that the spectroscopic success rate for a given  $(m_R, \mu_R)$  is not strongly correlated with the subsample selection criteria. Any error in this assumption is likely to be small, because the success rate is very high for the magnitude range analyzed. We then modify the sampling fraction  $f_s(m_R, \mu_R)$  by the factor  $p(target | sel)/p(target) |_{m_R}$ . We generally apply these corrections to our results, except where we explicitly specify otherwise in order to estimate the size of the effect.

Ultimately, we cut our galaxy sample not just by cluster membership, but also by emission line properties. The color selection bias described above applies to this case as well, because emission line (EL) galaxies have systematically different  $(B - R)$  colors than non-emission line (NEL) galaxies. Therefore, the probability that a galaxy has been targeted for spectroscopic observations is systematically higher if it is an EL galaxy. We remedy this problem in the same way as described above, by assuming a  $\Delta(B - R)$  color difference between EL and NEL galaxies and applying our systematic correction to recover an estimate for  $p(cz | sel)$ .

In principle, a similar correction would be required to account for the fact that the completeness of the  $U$ -band catalog has a residual dependence on  $U - R$ . Although Fig. 3.2 shows our compensation for the incompleteness of the  $U$ -band



sample to work very well in reproducing the correct  $R$ -band GLF, systematic errors arising from this problem are more likely to affect the  $U$ -band GLF. To quantify the impact that these errors may have on our results, we pursue the following approach:

We model the completeness of the  $U$ -band photometric sample, as

$$\begin{aligned}
 p(U \mid cz \wedge sel) &\equiv f_U \equiv f_U(m_R, m_U, \mu_R) \\
 &\approx \overline{f_U}(m_R, \mu_R) + \frac{\partial f_U}{\partial m_U}(m_U - \overline{m_U}|_R) \\
 &= \overline{f_U}(m_R, \mu_R) + \frac{\partial f_U}{\partial m_R} \frac{dR}{dU}(m_U - \overline{m_U}|_R).
 \end{aligned} \tag{E.5}$$

We determine  $\frac{\partial f}{\partial m_R}$  from our default sampling fraction model and obtain  $\frac{dR}{dU}$  (which is of order unity) directly from the sample.  $(m_U - \overline{m_U}|_R)$  is the difference between the  $m_U$  for which we determine the sample completeness and the mean  $\overline{m_U}$  appropriate for the given  $m_R$ . Therefore, this approach uses a first-order approximation to recover the sample completeness for a given  $m_U$  and  $\mu_R$  instead of  $m_R$  and  $\mu_R$  (thus removing most of the bias, as the dependence on apparent magnitude is stronger than that on surface brightness).

The resulting  $U$ -band GLF is consistent with the  $U$ -band GLF without these corrections ( $p_{\chi^2} = 0.994$ ). Because the systematic effects are negligible and the above argument is only an approximation to assess the order of magnitude of the effect, we choose not to apply these corrections in our calculations.

## APPENDIX F

### HOW MUCH LIGHT FROM FAINTER GALAXIES?

Our photometric and spectroscopic samples are limited in apparent magnitude and sky coverage. Because of these constraints, we have sampled only a part of the total galaxy population in each of our three clusters. In this and the following section, we will estimate what additional contributions in terms of numbers of galaxies and cumulative luminosity could come from cluster galaxies outside our photometric and spatial survey limits.

Our composite GLF reaches to  $M_U = -17.5$ . For a GLF with a rising faint end slope, such as the one in our sample, there are many more faint than bright galaxies. On the other hand, most of these are dwarfs that contribute only little to the total  $U$ -band light from a cluster. How much of the total  $U$ -band light have we sampled within our survey limit? To quantify this fraction, we fit a functional form to the sampled data and extrapolate it beyond the magnitude limit. As functional forms, we use our Schechter functions for the overall and NEL GLFs, and power laws for the EL GLFs. There are two possible approaches: We can fit a Schechter function to the overall GLF and extrapolate it to a very faint cutoff magnitude, or we can fit functional forms to the EL and NEL GLFs, extrapolate them individually, and add the results.

These two approaches are generally not equivalent, because the sum of the power law EL and Schechter NEL GLFs is not necessarily a Schechter function. While a single Schechter function, as we have fitted to our overall GLF, assumes that the faint end is approximated by a single power law, the superposition of an

EL and NEL GLF — assuming that there is no break or turnover in either GLF — usually shows a strong upturn at the faint end, just beyond our sampling limit, due to the increasing dominance of the power law EL GLF over the Schechter NEL GLF with its flat faint end slope. As we do not know whether the overall, EL, or NEL GLFs can be extrapolated all the way to the lower end of the galaxy magnitude range, or whether there is a break or turnover at a magnitude fainter than our survey limit, it is not clear which of these approaches best represents the real GLF of cluster galaxies.

Claims of an upturn at the faint end of the cluster GLF have been made in the past on the basis of photometric surveys that employed statistical subtraction rather than spectroscopic membership confirmation for background decontamination (Valotto, Moore & Lambas (2001) and references therein). Recent spectroscopic surveys of the GLF in clusters (Christlein & Zabludoff, 2003; de Propris et al., 2003) have not turned up evidence for such an upturn in the  $R$ - and  $b_J$ -bands, suggesting that previous claims may have been affected by biases inherent in statistical background subtraction (Valotto, Moore & Lambas, 2001). Some recent photometric surveys of nearby clusters (Trentham & Hodgkin, 2002) also indicate that the faint end slope is substantially flatter than our estimate for the EL GLF. Nevertheless, the question is not yet settled, because the spectroscopic surveys are generally shallower than purely photometric ones and the results from photometric surveys are still ambiguous. The uncertainty in the slope of the EL GLF in our present study is too large to predict at which magnitude an upturn might become significant.

Because of this uncertainty, we pursue both extrapolation approaches and compare the results below. To estimate our uncertainties, we perform the extrapolations for a range of Schechter parameters within the  $1\sigma$  error contours for

the individual fits to the overall, EL, and NEL GLFs.

Fig. F.1 shows the results for the cumulative  $U$ -band luminosity per cluster (averaged over the three clusters in our sample) as a function of the limit  $M_U$  to which we integrate the GLF. We use the zero-point flux for the Johnson  $U$ -band given by Colina, Bohlin & Castelli (1996) to convert absolute magnitudes to luminosities. The luminosity of a source with  $M_U = -19$  is given by  $L_U^{-19} = 2.065 \times 10^{39} \text{ erg s}^{-1} \text{ \AA}^{-1}$ . The faintest known dwarf galaxies with recent star formation in the Local Group have  $M_V \approx -10$  (Pritchet & van den Bergh, 1999). We adopt this value in the  $U$ -band as the faintest magnitude to which we extrapolate the GLF.

Judging from the overall GLF, galaxies fainter than our survey limit of  $M_U = -17.5$  contribute very little to the total cluster  $U$ -band luminosity. Therefore, we have sampled most of the cluster  $U$ -band light in our survey. Extrapolating the GLF to  $M_U = -10$ , the fraction of light in the bright end ( $M_U < -17.5$ ) is  $0.85_{+0.05}^{-0.06}$  (errors based on the  $1\sigma$  errors in  $\alpha_U$  and  $M_U^*$ ). It is notable that most (83%) of the  $U$ -band light at  $M_U < -17.5$  is contributed by giant NEL galaxies, which make up  $\sim 80\%$  in number. NEL galaxies dominate the  $U$ -band emission from clusters at least down to  $M_U = -14$ .

However, if we extrapolate the EL and NEL GLFs individually, the sum of the extrapolated GLFs, for  $M_U > -17.5$ , is not as flat as the fit to the overall GLF. The sum shows a strong upturn at fainter magnitudes due to the dominance of the power law EL GLF at faint magnitudes. In this case, populations of EL galaxies much fainter than  $M_U = -17.5$  contribute substantially to the total  $U$ -band luminosity of a cluster *if there is no break in the EL GLF*. In fact, for a faint end slope  $\alpha < -2$ , the luminosity integral diverges. If we allow for the existence of a steep faint end upturn that arises from the extrapolation of the EL GLF to fainter magnitudes, our best estimate for the fraction of the light sampled at  $M_U < -17.5$

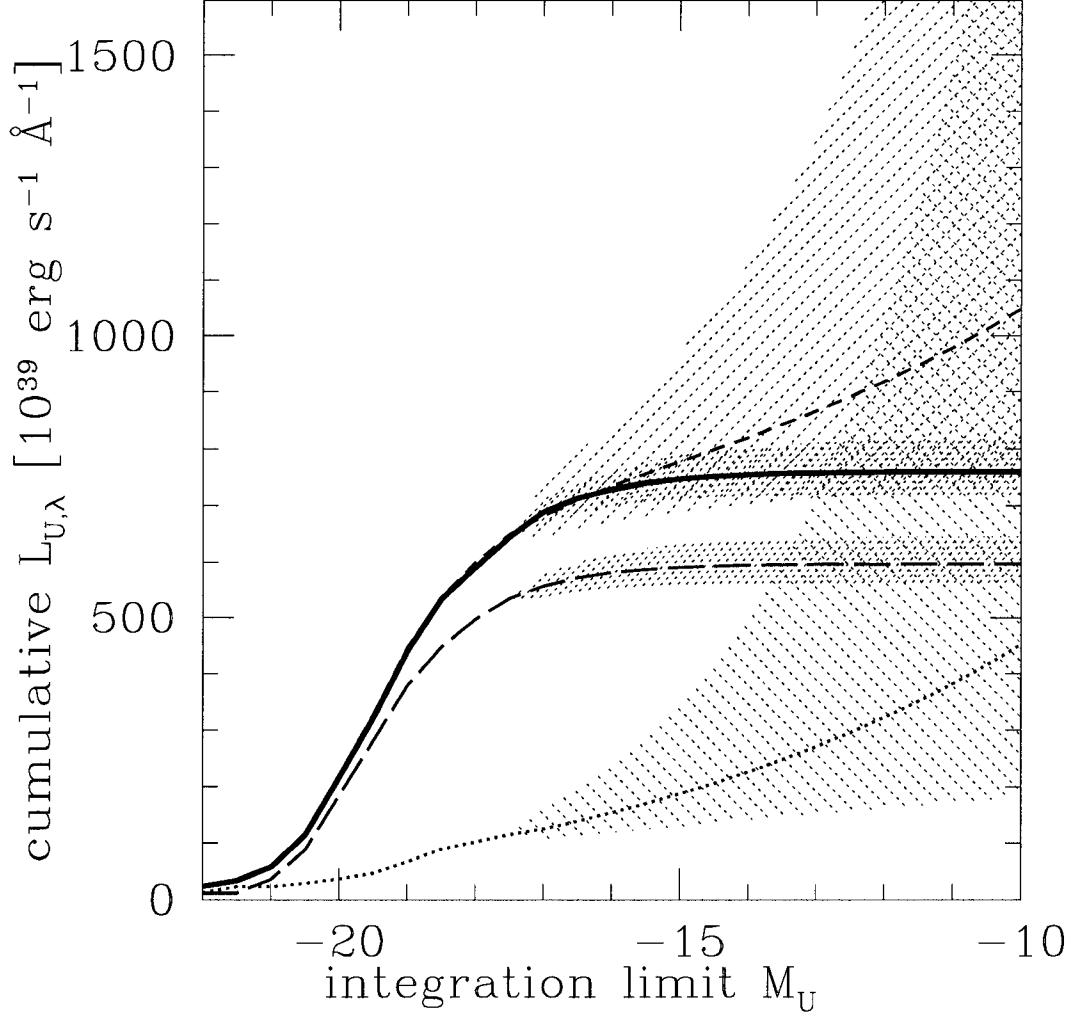


Figure F.1: Integrated  $U$ -band luminosity for emission line (EL), non-emission line (NEL) and overall GLFs per cluster as a function of faint integration limit. The solid bold line shows the cumulative  $U$ -band luminosity of all galaxies, using the Schechter fit for the overall GLF for extrapolation. The long-dashed line shows the same for NEL galaxies. The dotted line shows EL galaxies. The upper, short-dashed line gives the sum of the EL and NEL integrals.

is  $0.62^{+0.25}_{-0.38}$ . In the worst case, we would have sampled only 23% of the total  $U$ -band light from the clusters, and EL galaxies much fainter than our magnitude limit would constitute the dominant source of the EL flux. If the cluster GLF is indeed flatter than our EL GLF power law slope, as suggested by the results of Trentham & Hodgkin (2002), then the real fraction of the light that we have sampled is likely to be near the top end of this range.

Are these results valid for any individual cluster? The numbers above apply to an “average” of the three clusters in our sample. However, we have demonstrated in §3.1 that, even if the GLFs in our clusters are consistent with each other in shape, they differ in normalization. To estimate the cumulative  $L_U$  for any individual cluster, we therefore have to consider them separately. We use the composite GLF to constrain the shape parameters for the extrapolation under the assumption (which is consistent with the data) that the same fit describes all three clusters. We then renormalize the Schechter function to reproduce the total luminosities at  $M_U < -18.5$  in each individual cluster. Table 3.4 shows the luminosities for the composite GLF and the individual clusters. It gives both the total luminosities at the bright end ( $M_U < -18.5$ ) and the extrapolated luminosities to a limit of  $M_U = -10$ . The table shows that the systematic uncertainties about the shape of the faint end GLF, particularly whether there is an upturn or not, are larger than the random uncertainties associated with either extrapolation method.

Note that we have adopted a limit of  $M_U = -18.5$  instead of the  $M_U = -17.5$  that we regard as our confidence limit for the composite GLF. This is because, for the same reasons laid out in §3.4.4, the more distant clusters, A754 and A85, are only sampled to  $M_R \approx -19.0$ . For  $M_U > -18.5$ , significant numbers of galaxies could be lost at fainter  $M_R$ . However, we still use the full extent of the composite

GLF (to  $M_U = -17.5$ ) to constrain the Schechter parameters for the extrapolation.

## APPENDIX G

### HOW MUCH LIGHT FROM LARGER RADII?

What is the effect of the limited spatial coverage of our survey? We calculate galaxy number and luminosity profiles as a function of radius (normalized so that the harmonic radius is exactly 1 in each cluster). We then fit these profiles with isothermal beta models (Cavaliere & Fusco-Femiano, 1976, 1978) projected onto the radial dimension, and extrapolate the models to estimate the total number of galaxies and total  $U$ -band luminosity out to 4.1 harmonic radii ( $\sim 5$  Mpc in the case of A496, roughly the infall radius). This extrapolation is justified by recent results from Gómez et al. (2003), who show that star formation rates typically exhibit a break around 3-4 virial radii, comparable to the outer limit of our extrapolation and to the edge of the cluster.

We estimate that our survey samples  $0.50^{+0.11}_{-0.09}$  of the total  $U$ -band luminosity of a cluster from galaxies with  $M_U < -17.5$  within the survey region of A496. The fraction of galaxies sampled in this region is  $0.80^{+0.5}_{-0.6}$ . This result indicates that the  $U$ -band luminosity profile is more extended than the galaxy number count profile, presumably due to radial color gradients in the cluster. Within 0.9 harmonic radii, our survey samples more than three quarters of the galaxies in our clusters, and about 1/2 of the  $U$ -band luminosity from galaxies with  $M_U = -17.5$ .



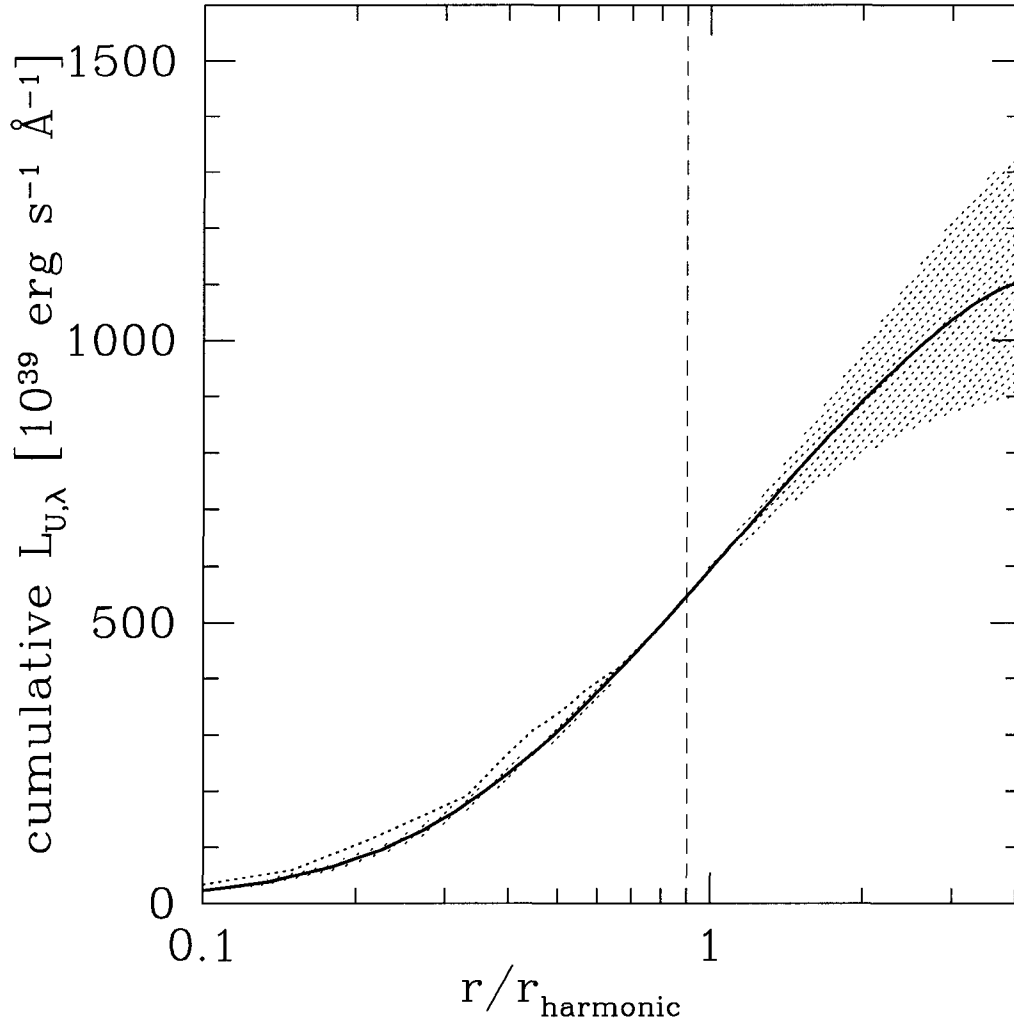


Figure G.1: Integrated  $U$ -band luminosity as a function of the cluster sampling radius in units of the harmonic radius for each cluster. The dotted line shows the measured profile out to a limit of  $r = 0.9r_{\text{vir}}$  (denoted by a thin vertical line). The dark solid line shows the extrapolation from the best fit isothermal- $\beta$  model. Errors are based on the  $1\sigma$  uncertainties in the best fit. The extrapolated profile is normalized to yield the correct sampled luminosity within  $r = 0.9r_{\text{vir}}$ . Within this limit, we sample  $\sim 50\%$  of the  $U$ -band light from galaxies with  $M_U < -17$ .

## REFERENCES

- Avni, Y., 1976 ApJ 210, 642
- Bahcall, N. A., 1973 ApJ 183, 783
- Balogh, M. L., Schade, D., Morris, S. L., Yee, H. K. C., Carlberg, R. G., Ellingson, E., 1998 ApJL 504, 75
- Balogh, M. L., Navarro, J. F., Morris, S. L., 2000 ApJ 540, 113
- Barnes, J. E., 1999, in: The Evolution of Galaxies on Cosmological Timescales, ADP Conference Series, Vol. 187
- Barnes, J. E., Hernquist, L., 1992 ARA&A 30, 705
- Barton Gillespie, E., Geller, M. J., Kenyon, S. J., 2003 ApJ 582, 668
- Bayes, T., 1764, Philosophical Transactions of the Royal Society of London
- Beijersbergen, M., Hoekstra, H., van Dokkum, P. G., van der Hulst, T., 2002, MNRAS 329, 385
- Bekki, K., 1998 ApJL 502, 133
- Bekki, K., Couch, W. J., Shioya, Y., 2002 ApJ 577, 651
- Bell, E. F., de Jong, R. S., 2001 ApJ 550, 212
- Benson, A. J., Frenk, C. S., Sharples, R. M., 2002 ApJ 574, 104
- Bernstein, R. A., Freedman, W. L., Madore, B. F., 2002, ApJ 571, 107
- Bernstein, R. A., Freedman, W. L., Madore, B. F., 2002, ApJ 571, 56

- Bertin, E., Arnouts, S., 1996 A&AS 117,393
- Biviano, A., Girardi, M., Giuricin, G., Mardirossian, F., Mezzetti, M., 1992 ApJ 396, 35
- Biviano, A., 2000, in: "Constructing the Universe with Clusters of Galaxies", IAP 2000 meeting, Paris, eds. F. Durret and D. Gerbal
- Blanton, M., et al., 2001, AJ 121, 2358
- Boroson, T., 1981 ApJS 46, 177
- Bromley, B. C., Press, W. H., Lin, H., Kirshner, R. P., 1998, ApJ 505, 25
- Bruzual, G. A., Charlot, S., 2003 MNRAS 344, 1000
- Cavaliere, A., Fusco-Femiano, R., 1976, A&A 49,137
- Cavaliere, A., Fusco-Femiano, R., 1978, A&A 70,677
- Choloniewski, J., 1986, MNRAS 226, 273
- Christlein, D., 2000 ApJ 545,145
- Christlein, D., Zabludoff, A., 2003, ApJ 591, 764
- Christlein, D., McIntosh, D. H., Zabludoff, A. I., 2004 ApJ, in press
- Christlein, D., Zabludoff, A. I., 2004, ApJ, submitted
- Christlein, D., Zabludoff, A. I., 2004, ApJ, submitted
- Colina, L., Bohlin, R., Castelli, F., Instrument Science Report CAL/SCS-008
- Colless, M., 1989 MNRAS 237, 799

- Colless, M., 1999 RSPTA 357, 105
- Couch, W. J., Barger, A. J., Smail, I., Ellis, R. S., Sharples, R. M., 1998 MjJ 497, 188
- Cross, N., Driver, S.P., 2002 MNRAS 329,579
- de Propriis, R., Stanford, S.A., Eisenhardt, P.R., Dickinson, M., Elston, R., 1999 AJ 118,719
- de Propriis, R., et al., 2003, MNRAS, 342, 725
- de Vaucouleurs, G., 1948 AnAp 11, 247
- Domínguez, M., Muriel, H., Lambas, D. G., 2001 AJ 121, 1266
- Dressler, A., 1978 ApJ 223, 765
- Dressler, A., Oemler, A., Couch, W. J., Smail, I., Ellis, R. S., Barger, A., Butcher, H., Poggianti, B. M., Sharples, R. M., 1996 ApJ 490, 577
- Dressler, A., 1980 ApJ 236, 351
- Dressler, A., Shectman, S. A., 1988 AJ 95, 985
- Dressler, A., Smail, I., Poggianti, B. M., Butcher, H., Couch, W. J., Ellis, R. S., Oemler, A., 1999 ApJS 122, 51
- Driver, S.P., Couch, W.J., Phillipps, S., 1998 MNRAS 301,369
- Efstathiou, G., Ellis, R. S., Peterson, B. A., 1998, MNRAS 232, 431
- Fukugita, M., Ichikawa, T., Gunn, J.E., Doi, M., Shimasaku, K., Schneider, D.P., 1996 AJ 111,1748
- Gaidos, E. J., 1997 AJ 113, 117

- Gibbons, J., *Nonparametric Methods for Quantitative Analysis*, 1976, Columbus, OH: American Sciences Press
- Gisler, G. R., 1978 MNRAS 183, 633
- Girardi, M., Giuricin, G., Mardirossian, F., Mezzetti, M., Boschin, W., 1998, ApJ 505, 74
- Girardi, M., Borgani, S., Giuricin, G., Mardirossian, F., Mezzetti, M., 2000, ApJ 530, 62
- Gómez, P. L., et al., 2003 ApJ 584, 210
- Graham, J.A., 1982 PASP 94, 244
- Gunn, J. E., Gott, J. R. III., 1972 JpJ 176, 1
- Hashimoto, Y., Oemler, A., Lin, H., Tucker, D. L., 1998 ApJ 499, 589
- Helou, G., Madore, G., Schmitz, M., Bica, M., Wu, X., & Bennett, J. 1991, in *Databases and On-Line Data in Astronomy*, ed. D. Egret & M. Albrecht (Dordrecht: Kluwer), 89
- Henry, R. C., 1999, ApJL 516, 49
- Hubble, E., Humason, M. L., 1931 ApJ 74, 43
- Jarvis, J.F., Tyson, J.A., 1981 AJ 86, 476
- Jenkins, A., Frenk, C. S., White, S. D. M., Colberg, J. M., Cole, S., Evrard, A. E., Couchman, H. M. P., Yoshida, N., 2001, MNRAS 321, 372
- Johnson, H. L. 1966, ARA&A, 4, 193
- Kendall, M. G., Stuart, A., *The advanced theory of statistics*, Griffin 1977

- Kennicutt, R. C., Roettiger, K. A., Keel, W. C., van der Hulst, J. M., Hummel, E.,  
1987 AJ 93, 1011
- Kennicutt, R. C. Jr., 1998 ARA&A 36, 189
- Kewley, L. J., Geller, M. J., Jansen, R. A., 2004 AJ 127, 2002
- Kochanek, C.S., Pahre, M.A., Falco, E.E., Huchra, J.P., Mader, J., 2001 ApJ 560,566
- Kochanek, C. S., White, M., Huchra, J., Macri, L., Jarret, T. H., Schneider, S. E.,  
Mader, J., 2003, ApJ 585, 161
- Landolt, A. U. 1992, AJ, 104, 340
- Larson, R. B., Tinslet, B. M., 1978 ApJ 219, 46
- Larson, R. B., Tinsley, B. M., Caldwell, C. N., 1980 ApJ 237, 692
- Lasker, B.M., Sturch, C.R., McLean, B.J., Russel, J.L., Jenkner, H., Shara, M.M.,  
1990 AJ 99,2019
- Leinert, Ch., et al., 1998, A&A Sup, 127 Jan I, 1
- Leitherer, C., Schaerer, D., Goldader, J.D., Delgado, R.M., Carmelle, R., Kune,  
D.F., de Mello, D.F., 1999 ApJS 123,3
- Lewis, I., et al., 2002 MNRAS 334, 673
- Lin, H., 1995, PhD Thesis, Harvard Univ.
- Lin, H., Kirshner, R.P., Sheckman, S.A., Landy, S.D., Oemler, A., Tucker, D.L.,  
Schechter, P.L., 1996 ApJ 464,60
- Liu, C. T., Kennicutt, R. C., 1995 ApJ 450, 547

- Lonsdale, C. J., Persson, S. E., Matthews, K., 1984 ApJ 287, 95
- Lugger, P. M., 1986 ApJ 303, 535
- Lynden-Bell, D., 1971, MNRAS 155, 95
- Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., Fruchter, A., 1996, MNRAS 283, 1388
- Madau, P., Pozzetti, L., Dickinson, M., 1998 ApJ 498, 106
- Madgwick, D.S., et al., 2002 MNRAS 332, 827
- Madgwick, D. S., et al., 2002, MNRAS 333, 133
- Massarotti, M., Busarello, G., La Barbera, F., Merluzzi, P., 2003, A& 404, 75
- McMahon, P.M., 1993 PhD Thesis, Columbia Univ.
- Mihos, J. C., Hernquist, L., 1994 ApJL 425, 13
- Mink, D.J., Wyatt, W.F., 1995 adass 4,496
- Mo, H. J., Yang, X., van den Bosch, F. C., Jinh, Y. P., 2004 MNRAS 349, 205
- Monet, D. et al. 1996, USNO-SA2.0, (U.S. Naval Observatory, Wash. DC)
- Morgan, W.W., Kayser, S., White, R.A., 1975 ApJ 199, 545
- Morrison, J.E., McLean, B., GSC-Catalog Construction Team, II, 2001 DDA,32.0603
- Muriel, H., Valotto, C., & Lambas, D. 1998, ApJ, 506, 540
- Nelson, A.E., Gonzalez, A.H., Zaritsky, D., Dalcanton, J.J., 2002 ApJ 556,103

- Oemler, A., 1974 ApJ 194, 1
- Paolillo, M., Andreon, S., Longo, G., Puddu, e., Gal, R. R., Scaramella, R., Djorgovsky, S. G., de Carvalho, R., 2001 A&A 367, 59
- Pence, W., 1976, ApJ 203, 39
- Peng, C. Y., Ho, L. C., Impey, C. D., Rix, H.-W., 2002 AJ 124, 266
- Peterson, B.A., 1970 ApJ 159,333
- Pimbblet, K.A., Smail, I., Kodama, T., Couch, W.J., Edge, A.C., Zabludoff, A.I., O'Hely, E., 2002 MNRAS 331, 333
- Poggianti, B. M., Smail, I., Dressler, A., Couch, W. J., Barger, A. J., Butcher, H., Ellis, R. S., Oemler, A., 199 ApJ 518, 576
- Postman, M., Lauer, T.R., 1995 ApJ 440,28
- Pozzetti, L., Madau, P., Zamorani, G., Ferguson, H. C., Bruzual, A. G., 1998, MNRAS 298, 1133
- Pritchett, C., van den Bergh, S., 1998, AJ 118, 883
- Reiprich, T. H., Böhringer, H., 2002, ApJ 567, 716
- Rix, H.-W., White, S. D. M., 1990 ApJ 362, 52
- Rood, H. J., Turnrose, B. E., 1968 ApJ 152, 1057
- Rood, H. J., 1969 ApJ 158, 657
- Salpeter, E. E., 1955 ApJ 121, 161
- Sandage, A.R., 1972 ApJ 178,1



- Sandage, A., Tamman, G. A., Yahil, A., 1979, ApJ 232, 352
- Sanromà, M., Salvador-Solé, E., 1990 ApJ 360, 16
- Sarazin, C. L., 1980 ApJ 236, 75
- Schechter, P., 1976 ApJ 203, 297
- Schlegel, D., Finkbeiner, D., & Davis, M., 1998, ApJ 500, 525
- Shectman, S.A., Schechter, P.L., Oemler, A.A., Tucker, D., Kirshner, R., Lin, H., 1992 csg conf 351
- Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Lin, H., Kirshner, R. P., & Schechter, P. L. 1996, ApJ, 470, 172
- Simard, L., et al., 2002 ApJS 142, 1
- Simien, F., de Vaucouleurs, G., 1986 ApJ 302, 564
- Smith, R.M., Driver, S.P., Phillipps, S., 1997 MNRAS 287, 415
- Solanes, J., M., Salvador-Solé, E., Sanromà, M., 1989 AJ 98, 798
- Springel, V., White, S.D.M., Tormen, G., Kauffmann, G., 2001 MNRAS 328, 726
- Stanford, S. A., Eisenhardt, P. R., Dickinson, M., Holden, B. P., de Propris, R., 2002, ApJS 142, 153
- Steidel, C. C., Adelberger, K., Giavalisco, M., Dickinson, M., Pettini, M., 1999 ApJ 519, 1
- Terlevich, A. I., Caldwell, N., & Bower, R. G. 2001, MNRAS, 326, 1547

- Thuan, T. X., Balkowski, C., and Tran Thanh Van, ed. 1993, *Physics of Nearby Galaxies: Nature or Nurture?* (Gif-sur-Yvette, Editions Frontières)
- Tran, K. H., Simard, L., Zabludoff, A., Mulchaey, J.S., 2001 ApJ 549,172
- Tremaine, S., Richstone, D.O., 1977 ApJ 212,311
- Trentham, N., 1998 MNRAS 294,193
- Trentham, N., Hodgkin, S., 2002 MNRAS333, 423
- Tyson, J. A., 1995, in Calzetti D., Livio M., Madau P., eds, *Extragalactic Background Radiation*, Cambridge University Press, p. 103
- van Gorkom, J. H., Bravo-Alfaro, H., Dwarakanath, K. S., Guhathakurta, P., Poggianti, B. M., Schiminovich, D., Valluri, M., Verheijen, M., Wilcots, E., Zabludoff, A., 2003 Ap&SS 285, 219
- Valotto, C. A., Nicotra, M. A., Muriel, H., Lambas, D. G., 1997 ApJ 479, 90
- Valotto, C. A., Moore, B., Lambas, D. G., 2001, ApJ 546, 157
- Vollmer, B., Balkowski, C., Cayatte, V., van Driel, W., Huchtmeier, W., 2004 A&A 419,35
- Vollmer, B., Beck, R., Kenney, J. D. P., van Gorkom, J. H., 2004 AJ 127,3375
- Whitmore, B. C., In *Clusters of Galaxies*, ed. W. R. Oegerle, M. J. Fitchett, and L. Danly (New York, Cambridge University Press), p. 139
- Whitmore, B. C., Gilmore, D. M., 1991 ApJ 367, 64
- Whitmore, B. C., Gilmore, D. M., Jones, C., 1993 ApJ 407, 489
- Yang, Y., Zabludoff, A. I., Zaritsky, D., Lauer, T. R., Mihos, J. C., 2004 ApJ 607, 258

York, D. G., et al., 2000 AJ 120, 1579

Zabludoff, A. I., Huchra, J. P., Geller, M. J., 1990 ApJS 74, 1

Zabludoff, A.I., Mulchaey, J.S., 1998 ApJ 496, 39

Zabludoff, A.I., Mulchaey, J.S., 2000 ApJ 539,136

Zaritsky, D., Zabludoff, A.I., Willick, J.A., 1995 AJ 110, 1602