

Finding Structures in Photometric Redshift Surveys



Christine S. Botzler

FINDING STRUCTURES IN PHOTOMETRIC REDSHIFT SURVEYS

Ph.D. Thesis

at the Faculty of Physics & Astronomy
of the
Ludwig-Maximilian University Munich

submitted by

Christine S. Botzler

from Munich

Munich, February 2004

1. Evaluator: Prof. Dr. Ralf Bender
2. Evaluator: Dr. Hans Böhringer

Date of the oral exam: 28.05.2004

We shall not cease from exploration, and the end of all our exploring will be to arrive where we started and know the place for the first time.

– T. S. Eliot

*to my parents and grandmothers –
for their endless support*

Contents

1	Introduction	1
1.1	Galaxy groups and clusters	2
1.2	Photometric redshifts	8
1.3	The Munich Near-Infrared Cluster Survey (MUNICS)	8
1.4	Outline	10
2	Structure Finding – An Overview	13
2.1	Complications	13
2.2	Choice of the right technique	13
2.3	Differentiating the techniques	14
2.4	General conspectus	14
3	Friends-of-Friends Algorithms	19
3.1	The Huchra & Geller friends-of-friends (FOF) algorithm	19
3.2	The extended friends-of-friends (EXT-FOF) algorithm	21
3.3	The linking parameters	25
3.3.1	Spectroscopic surveys	26
3.3.2	Simulated datasets	27
3.3.3	Photometric surveys	27
4	Testing the Extended Friends-of-Friends	29
4.1	Application to the CFA1 redshift survey	29
4.1.1	The galaxy catalog	29
4.1.2	Creation of the group catalogs	30
4.1.3	Comparison of the group catalogs	32
4.1.4	Analysis of the discrepancies	34
4.2	Application to the LCRS redshift survey	40
4.2.1	The galaxy catalog	40
4.2.2	Creation of the group catalogs	41
4.2.3	Comparison of the group catalogs	44
4.2.4	Analysis of the discrepancies	45
4.3	LCRS with simulated photometric redshifts	47
4.3.1	Creation of the galaxy dataset with simulated photometric redshifts	47
4.3.2	Creation of the group catalog	47

4.3.3	Comparison of the group catalogs	48
4.3.4	Analysis of the discrepancies	49
5	Authenticity of Structures	51
5.1	Spectroscopic redshifts	51
5.2	Voronoi tessellation	53
5.2.1	The definition of Voronoi tessellation	54
5.2.2	Application to structure finding	56
5.3	Likelihood approach	59
5.3.1	Bayesian statistics and the likelihood	60
5.3.2	Application to structure finding	61
5.4	Color-magnitude diagrams	66
5.4.1	The cluster red sequence	67
5.4.2	Application to structure finding	68
6	Structures in the MUNICS Survey	71
6.1	The galaxy catalog	71
6.1.1	The field geometry	72
6.1.2	The spectroscopic sample	72
6.1.3	Classification of all galaxies	76
6.2	Finding the structures	77
6.2.1	The algorithm	78
6.2.2	Optimizing the linking parameters	79
6.3	The structure catalog	81
6.3.1	The field S6F5	84
6.3.2	Examples of clusters in S6F5	87
6.4	Spectroscopic verification	96
7	Summary	101
A	Maximizing the Likelihood	115
B	MUNICS Structures	119
B.1	The field S2F1	119
B.2	The field S2F5	122
B.3	The field S3F5	125
B.4	The field S5F1	128
B.5	The field S5F5	131
B.6	The field S6F1	134
B.7	The field S7F5	137

List of Figures

1.1	Comparison of photometric and spectroscopic redshifts for a subsample of the MUNICS objects	11
1.2	Redshift distribution of the MUNICS galaxies	12
3.1	Flow chart for the Huchra & Geller FOF algorithm	20
3.2	Flow chart for the EXT-FOF algorithm	23
4.1	Variation of the linking criteria D_L and V_L for the CFA1	32
4.2	Comparison of the CFA1 linking distance D_L and separation D_{ij} for a typical galaxy pair	37
4.3	Comparison of the CFA1 linking distance D_L and separation D_{ij} for two special cases of galaxy pairs	38
4.4	Comparison of the CFA1 linking velocity V_L and the velocity separations V_{ij} and V_i	40
4.5	Variation of the linking criteria D_L and V_L for the LCRS	44
4.6	Comparison of the LCRS linking distance D_L and separation D_{ij} for a typical galaxy pair	46
5.1	Point patterns, Voronoi tessellations and Delaunay meshes for two different two-dimensional point distributions	55
5.2	Variation of the characteristic K band magnitude of cluster luminosity functions with redshift	66
5.3	$m_V - m_I$ vs. m_I model red sequences	70
6.1	SED-based classification of objects in the spectroscopic and photometric MUNICS datasets	74
6.2	Statistics of the entire MUNICS structure catalog	83
6.3	Geometrical borders and projected distribution of structures in S6F5	88
6.4	Redshift distribution of galaxies and structures in S6F5	89
6.5	Voronoi and likelihood maps for structure 4 in S6F5	90
6.6	CMD of structure 4 in S6F5	91
6.7	Voronoi and likelihood maps for structure 10 in S6F5	92
6.8	CMD of structure 10 in S6F5	93
6.9	Voronoi and likelihood maps for structure 20 in S6F5	94
6.10	CMD of structure 20 in S6F5	95

6.11 Comparison of the photometric mean structure redshift \bar{z} and its spectroscopic counterpart $\bar{z}_{cluster}$	99
B.1 Geometrical borders and projected distribution of structures in S2F1	120
B.2 Redshift distribution of galaxies and structures in S2F1	121
B.3 Geometrical borders and projected distribution of structures in S2F5	123
B.4 Redshift distribution of galaxies and structures in S2F5	124
B.5 Geometrical borders and projected distribution of structures in S3F5	126
B.6 Redshift distribution of galaxies and structures in S3F5	127
B.7 Geometrical borders and projected distribution of structures in S5F1	129
B.8 Redshift distribution of galaxies and structures in S5F1	130
B.9 Geometrical borders and projected distribution of structures in S5F5	132
B.10 Redshift distribution of galaxies and structures in S5F5	133
B.11 Geometrical borders and projected distribution of structures in S6F1	135
B.12 Redshift distribution of galaxies and structures in S6F1	136
B.13 Geometrical borders and projected distribution of structures in S7F5	138
B.14 Redshift distribution of galaxies and structures in S7F5	139

List of Tables

1.1	Definition of the Abell richness classes R	3
1.2	Compendium of some basic properties of clusters and groups	6
1.3	Cluster classifications and some related properties	7
1.4	Observational properties of the MUNICS fields	9
1.5	Galactic and stellar content in the MUNICS fields	10
2.1	Structure finding algorithms	18
4.1	Comparison of the CFA1 FOF and EXT-FOF structures	33
4.2	Comparison of the LCRS FOF and EXT-FOF structures	45
4.3	Comparison of the LCRS EXT-FOF structure catalogs with and without simulated photometric redshifts	48
6.1	Spectroscopic sample of MUNICS objects	73
6.2	Image-based classification of the spectroscopic sample	73
6.3	SED-based classification of the spectroscopic sample	75
6.4	Combination of image- and SED-based classification of the spectroscopic sample	76
6.5	Image- and SED-based classification of the photometric dataset	77
6.6	Optimized redshift error pre-selection parameters	80
6.7	Redshift error pre-selected galaxies	81
6.8	Overview of the MUNICS structure catalog	82
6.9	Structures in S6F5	85
6.10	Overview of the spectroscopic verification statistics	97
6.11	Spectroscopically verified structures	98
B.1	Structures in S2F1	119
B.2	Structures in S2F5	122
B.3	Structures in S3F5	125
B.4	Structures in S5F1	128
B.5	Structures in S5F5	131
B.6	Structures in S6F1	134
B.7	Structures in S7F5	137

Chapter 1

Introduction

When talking about large-scale structure, astronomers usually think of galaxy groups, clusters, or superclusters, that are arranged in large filaments and walls, separated by huge voids of almost empty space, very much like the soap skin in a bubble bath. The basic tracers of this large-scale structure, the galaxy groups and especially the clusters, are very important for extragalactic astronomy, as well as cosmology.

Clusters constitute excellent laboratories for the study of various important topics of extragalactic astronomy. For example, the interaction of galaxies with each other, or with the intergalactic material can be examined. Furthermore, the distribution of the dark matter, that dominates the cluster mass, can be studied. A large number of publications deal with the correlation between various cluster properties, like X-ray luminosity and temperature (Henry & Tucker, 1979), X-ray temperature and galaxy velocity dispersion (Lubin & Bahcall, 1993), or shape, richness, and velocity dispersion of the galaxies (Struble & Ftaclas, 1994), to name just a few. The relation between galaxy morphology and cluster density has been examined by Dressler (1980), showing that clusters with higher densities have a larger fraction of elliptical and lenticular galaxies. Butcher & Oemler (1984) studied the evolution of colors in cluster galaxies and found that the fraction of blue galaxies increases with redshift¹. It is furthermore of interest to study the evolution of galactic properties, like the luminosity function, in both high- and low-density environments (Trentham, 1998; de Propris et al., 1999; Drory et al., 2001a; Fried et al., 2001; Postman et al., 2001; Feulner et al., 2003), in order to understand the role of environmental effects. These studies yield valuable information on the formation and evolution of galaxies and their conglomerates.

Clusters of galaxies, that are the most massive collapsed objects as of today, have been playing a fundamental role in our understanding of structure formation and cosmology. As has been shown by Press & Schechter (1974), Peebles (1993), and Eke et al. (1996) amongst others, the evolution of the mass-function of groups and clusters of galaxies is highly sensitive to cosmological parameters, the type of dark matter and the biasing of dark against baryonic matter. Thus, comparison of the mass-function

¹the so-called Butcher-Oemler effect

predicted by semi-analytic structure formation models, like the Press-Schechter formalism (Press & Schechter, 1974), or N-body simulations with the observed mass-function of clusters yields strong constraints on structure formation scenarios and the cosmological paradigm (Bahcall & Cen, 1992; Ueda et al., 1993; Jing & Fang, 1994; Bahcall et al., 1997; Botzler, 1999; Bode et al., 2001).

All of these studies depend on a reliable identification of galaxy clusters. In order to find these structures in any type of dataset, it is crucial to know their typical observable properties, like for example the spatial distribution and the colors of the galaxies within the clusters, or the extension and luminosity of the X-ray emission of the hot intracluster gas. Structure finding is always based on the search for either one or more of those characteristic group or cluster properties.

A summary of the composition, classification, and the most important properties of galaxy groups and clusters is given in Sect. 1.1. This should aid the reader in understanding the choice of techniques and criteria used for structure finding, and ease the interpretation of the resulting group and cluster catalogs. Sect. 1.2 gives a short introduction to the method of determining redshifts based on multi-band photometry, the so-called photometric redshifts. While photometric redshift galaxy surveys have increased in popularity over the last years, a paucity of techniques suitable for cluster finding in this type of survey exists. The new structure finding technique introduced in this work, the extended friends-of-friends (hereafter EXT-FOF) algorithm, is designed to deal with photometric redshift surveys. In Sect. 1.3 the Munich Near-Infrared Cluster Survey (hereafter MUNICS), a photometric redshift survey, is introduced. The application of the new EXT-FOF method to the MUNICS survey is one of the main topics in this thesis. Sect. 1.4 gives an overview of the following chapters.

1.1 Galaxy groups and clusters

Galaxy clusters are overdense systems of galaxies. They are bound and virialized, held together by the gravitational force of the cluster constituents. Clusters basically contain three components: A conglomeration of various types of galaxies, hot intracluster gas, and large amounts of dark matter (DM).

The observable cluster properties depend strongly on the number of galaxies contained within the structure, or alternatively on its total mass. A measure for this richness of clusters was defined by Abell (1958). According to this definition, clusters are divided into six richness classes R , ranging from $R = 0$ to 5 , with increasing number N_R of galaxies. The richness N_R is the number of galaxies brighter than $m_3 + 2$ mag, with m_3 being the magnitude of the third brightest cluster member. Only galaxies within a radius of $1.5h^{-1}$ Mpc from the cluster center are counted. h is defined as $H_0 = h \cdot 100 \text{ km s}^{-1}\text{Mpc}^{-1}$ throughout this thesis, where H_0 is the Hubble constant. Table 1.1 shows the partitioning of the clusters into richness classes as a function of N_R . As was proven by studies of nearby clusters with various richnesses, N_R is proportional to the total number of galaxies in the clusters.

R	N_R
0	30 – 49
1	50 – 79
2	80 – 129
3	130 – 199
4	200 – 299
5	≥ 300

Table 1.1: Definition of the Abell richness classes R . N_R is the number of galaxies brighter than $m_3 + 2$ mag, with m_3 being the magnitude of the third brightest cluster member, within a radius of $1.5 h^{-1} \text{Mpc}$ from the cluster center.

One of the most frequently used characteristics for cluster finding, is the density profile of the galaxy distribution in rich clusters. The radial density distribution of galaxies can be approximated by a truncated isothermal sphere (Zwicky, 1957; Bahcall, 1977), or in the central regions by a King profile (King, 1972).

The inner part of the King profile is given by

$$n(r) = n_0 (1 + r^2/r_c^2)^{-3/2}, \quad (1.1)$$

$$s(r) = s_0 (1 + r^2/r_c^2)^{-1}, \quad (1.2)$$

where $n(r)$ and $s(r)$ are the spatial and projected density distributions as functions of distance r from the cluster center. n_0 and s_0 are the respective central densities, that relate like $s_0 = 2r_c n_0$, and r_c is the core radius, that is defined by $s(r_c) = s_0/2$.

The isothermal sphere exhibits a profile slope that can be approximated by

$$s(r \leq r_c/3) \sim \text{constant}, \quad (1.3)$$

$$s(r_c \leq r \leq 10r_c) \propto r^{-1.6}, \quad (1.4)$$

and

$$n(r_c \leq r \leq 10r_c) \propto r^{-2.6}. \quad (1.5)$$

However, substructures in the concentration of galaxies do exist in about 40% of the rich clusters, according to Geller (1990).

Typical values for the central spatial density within the three brightest magnitudes are $n_0(\Delta m \approx 3 \text{ mag}) \sim (10^3 - 10^4) h^3 \text{Mpc}^{-3}$ (Bahcall, 1975, 1977; Dressler, 1978).

Characteristic core radii for rich and compact clusters are in the range of $r_c \approx (0.1 - 0.25) h^{-1} \text{Mpc}$, following Bahcall (1975, 1977), and Dressler (1978).

The cutoff radius r_h , used for the truncation of the isothermal sphere, is defined as the distance from the cluster center, where the projected number density drops to 1.5% of its isothermal sphere central value $s(r_h) \approx 0.015 s_0^{iso}$. A typical value for this quantity for a rich cluster is of the order of $r_h \approx 20 r_c$ (Bahcall, 1977). The range of radial extensions goes from $r_h \approx 0.1 h^{-1} \text{Mpc}$ for small groups to $r_h \approx 2 h^{-1} \text{Mpc}$ for rich clusters (Bahcall in Dekel & Ostriker 1999).

The average density of bright ($\geq L^*$) galaxies in the field is roughly $n(\text{field}) \sim 1.5 \cdot 10^{-2} h^3 \text{Mpc}^{-3}$, while the corresponding average density in a rich cluster can be approximated by $n(\text{cluster}) \sim 3 h^3 \text{Mpc}^{-3}$ within $1.5 h^{-1} \text{Mpc}$ (Longair 1998, and Bahcall in Dekel & Ostriker 1999). Thus, galaxies are overdense in clusters by a factor of up to 200. The galaxy overdensity increases with decreasing radius from the cluster center and reaches typically $10^4 - 10^5$ in the cores of rich compact clusters (Bahcall in Dekel & Ostriker 1999).

The radial velocity dispersion σ of a cluster is defined as the standard deviation of the individual galaxy velocities along the line-of-sight. It is thus a measure for the dynamical mass of the clusters. The velocity dispersion typically ranges from $\sigma \approx (100 - 500) \text{ km s}^{-1}$ for groups and poor clusters to $\sigma \approx (400 - 1400) \text{ km s}^{-1}$ for rich structures (Struble & Rood 1991, and Bahcall in Dekel & Ostriker 1999). The median values are roughly $\sigma \sim 250 \text{ km s}^{-1}$ and $\sigma \sim 750 \text{ km s}^{-1}$, respectively (Bahcall in Dekel & Ostriker 1999). A weak correlation between σ and the richness N_R seems to exist, with richer clusters having higher velocity dispersions (Bahcall, 1981).

Applying the virial theorem to an isothermal distribution yields a dynamical mass for a rich cluster (with $\sigma \approx 750 \text{ km s}^{-1}$) of $M(\leq 1.5 h^{-1} \text{Mpc}) \approx 0.4 \cdot 10^{15} h^{-1} M_\odot$ within $1.5 h^{-1} \text{Mpc}$ from the cluster center. M_\odot is the mass of the sun. The range of characteristic dynamical masses is $(10^{12.5} - 10^{14}) h^{-1} M_\odot$ for poor clusters, and $(10^{14} - 2 \cdot 10^{15}) h^{-1} M_\odot$ for rich structures (Bahcall in Dekel & Ostriker 1999). The mass M_{stars} contained within the stars of the individual galaxies can be approximated by $M_{\text{stars}} \approx 0.05 \cdot M(\leq 1.5 h^{-1} \text{Mpc})$ (Bahcall in Dekel & Ostriker 1999).

Typical blue luminosities $L_B(\leq 1.5 h^{-1} \text{Mpc})$ within $1.5 h^{-1} \text{Mpc}$ from the cluster center range from $(10^{10.5} - 10^{12}) h^{-2} L_\odot$ for groups to $(6 \cdot 10^{11} - 6 \cdot 10^{12}) h^{-2} L_\odot$ for clusters, with a median of $L_B(\leq 1.5 h^{-1} \text{Mpc}) \sim 10^{12} h^{-2} L_\odot$ (Bahcall in Dekel & Ostriker 1999). L_\odot is the solar luminosity. As a result, the typical mass-to-luminosity ratio of a rich cluster can be approximated by $\langle M/L_B \rangle \sim 300 h M_\odot / L_\odot$.

All rich clusters of galaxies contain an intracluster medium, that consists of hot plasma. This plasma can be detected by its emission of X-ray radiation, that is produced by thermal bremsstrahlung. The X-ray emission is spatially extended, going out to the same distances r as the main galaxy concentration $r \approx 1.5 h^{-1} \text{Mpc}$. The typical temperature T_X of the gas is $T_X \sim 5 \text{ keV}$, and ranges from $(2 - 14) \text{ keV}$. A characteristic value for the X-ray luminosity is $L_X \sim 10^{44} h^{-2} \text{erg s}^{-1}$. Both T_X and L_X increase with richness (Bahcall in Dekel & Ostriker 1999).

The hot plasma can be furthermore detected by its influence on the Cosmic Microwave Background (CMB) radiation. This is called the Sunyaev-Zeldovich effect (Sunyaev & Zeldovich, 1972). As the photons of the CMB radiation penetrate the hot intracluster gas, a part of them suffers inverse Compton scattering by the gas electrons. Though the photons have the same probability of gaining and losing energy to first order, there is a net statistical increase of their energy as a second order effect. As a result, the spectrum of the CMB radiation is shifted to marginally higher energies. This perturbation causes a decrement in the low-energy Rayleigh-Jeans region of the spectrum of the CMB radiation in the direction of a cluster, while the high-energetic Wien region shows an excess.

The ratio of the gas mass M_{gas} to the dynamical mass M in the cluster (both within $\sim 1.5 h^{-1} \text{Mpc}$) is found to be in the range of $M_{\text{gas}}/M \approx (0.03 - 0.15) h^{-1.5}$, with a median value of $\langle M_{\text{gas}}/M \rangle \approx 0.07 h^{-1.5}$ (Jones & Forman, 1992; White et al., 1993; White & Fabian, 1995; Lubin et al., 1996). Thus, the total gas mass in a rich cluster is typically $M_{\text{gas}} \sim (10^{13} - 10^{14}) h^{-2.5} M_{\odot}$ (Bahcall in Dekel & Ostriker 1999). M_{gas} is usually larger than the total luminous mass inside the galaxies M_{stars} . This suggests that a significant fraction of the intracluster medium is of cosmological origin, as explained by Bahcall in Dekel & Ostriker (1999).

The total baryon fraction in typical rich clusters is at least $(M_{\text{stars}} + M_{\text{gas}})/M \approx 0.05 + 0.07 h^{-1.5}$. With a Hubble constant of $h = 0.65$, this yields a baryon fraction of $\sim 20\%$, with 5% contained within the luminous parts of the individual galaxies, and about 15% in the form of hot intergalactic gas, while roughly 80% of the cluster mass is made up of dark matter. In general, the spatial distribution of galaxies, gas, and dark matter seems to be very similar.

According to Bahcall et al. (1995), the high mass-to-light ratios for galaxy clusters of $M/L_B \approx (300 \pm 100) h M_{\odot}/L_{\odot}$ can be explained by the dominance of early type galaxies, the ellipticals and lenticulars, in many clusters. While the late type spiral galaxies usually have M/L_B ratios of $\sim 100 h M_{\odot}/L_{\odot}$, ellipticals exhibit ratios of up to $\sim 400 h M_{\odot}/L_{\odot}$, within $0.2 h^{-1} \text{Mpc}$ from the galaxy core. This implies that clusters do not necessarily contain additional amounts of dark matter. The dark matter associated with or stripped of the galaxy halos seems to be sufficient (Bahcall in Dekel & Ostriker 1999).

The number density of galaxy groups and clusters is highly dependent on their richness. The richer the structures, the rarer they are. Characteristic number densities n_s for groups and poor clusters have a range of $(10^{-3} - 10^{-5}) h^3 \text{Mpc}^{-3}$, whereas rich structures have densities of $(10^{-5} - 10^{-6}) h^3 \text{Mpc}^{-3}$. The associated cluster correlation scale $r_s := n_s^{-1/3}$ ranges from $(22 \pm 4) h^{-1} \text{Mpc}$ for rich clusters to $(13 \pm 2) h^{-1} \text{Mpc}$ for groups (Bahcall in Dekel & Ostriker 1999).

Table 1.2 shows a compendium of some of the above mentioned basic properties of clusters and groups. As can be seen once more in the table, groups and poor clusters are an extension of typical cluster properties to lower richness, mass, size, and luminosity.

Galaxy clusters can be arranged into a one-parameter sequence, very much like the galaxies themselves are arranged into the Hubble sequence. The sequence goes from

Property	Rich clusters	Groups and poor clusters
Richness N_R	30 – 300+	3 – 30
Richness class R	0 – 5	–
Core radius r_c	(0.1 – 0.25) h^{-1} Mpc	–
Radial extension r_h	(1 – 2) h^{-1} Mpc	(0.1 – 1) h^{-1} Mpc
Radial velocity dispersion σ		
Median	$\sim 750 \text{ km s}^{-1}$	$\sim 250 \text{ km s}^{-1}$
Range	(400 – 1400) km s^{-1}	(100 – 500) km s^{-1}
Dynamical mass		
$M(\leq 1.5h^{-1}\text{Mpc})$	$(10^{14} - 2 \cdot 10^{15}) h^{-1}M_\odot$	$(10^{12.5} - 10^{14}) h^{-1}M_\odot$
Luminosity		
$L_B(\leq 1.5h^{-1}\text{Mpc})$	$(6 \cdot 10^{11} - 6 \cdot 10^{12}) h^{-2}L_\odot$	$(10^{10.5} - 10^{12}) h^{-2}L_\odot$
Mass-to-light ratio $\langle M/L_B \rangle$	$\sim 300 h M_\odot/L_\odot$	$\sim 200 h M_\odot/L_\odot$
X-ray temperature T_X	(2 – 14) keV	$\leq 2 \text{ keV}$
X-ray luminosity L_X	$(10^{42.5} - 10^{45}) h^{-2} \text{ erg s}^{-1}$	$\leq 10^{43} h^{-2} \text{ erg s}^{-1}$
Structure number density n_s	$(10^{-5} - 10^{-6}) h^3 \text{Mpc}^{-3}$	$(10^{-3} - 10^{-5}) h^3 \text{Mpc}^{-3}$
Structure correlation scale r_s	$(22 \pm 4) h^{-1} \text{Mpc}$	$(13 \pm 2) h^{-1} \text{Mpc}$

Table 1.2: Compendium of some basic properties of clusters and groups (after Bahcall in Dekel & Ostriker 1999).

regular, or early type, to irregular, or late-type, clusters and contains one intermediate category. There is a strong correlation between many cluster properties, like shape, central concentration, galactic content, or X-ray emissivity with the position in this sequence.

Regular clusters are circularly symmetric, have a high central concentration, and are made up of mostly elliptical (E) and lenticular (S0) galaxies ($\sim 35\%$ and $\sim 45\%$, respectively). Irregular clusters have little or no circular symmetry, only show signs of minor central concentration, and have a high fraction ($\sim 50\%$) of spiral (S) galaxies. For comparison, the average galaxy composition in the field is 10% ellipticals, 20% lenticulars, and 70% spirals. While the radial distribution of E, S0, and S is constant in irregular clusters, the relative space density of S decreases significantly towards the center of intermediate and regular type structures. This effect is called segregation by galaxy type. As a result, the spiral galaxies in a regular type cluster are arranged in a halo around a core of elliptical and lenticular galaxies. There is also some evidence for segregation by mass in regular and intermediate type clusters, i.e. more massive galaxies are located closer to the cluster center than less massive ones. No mass segregation can be found in irregular type clusters (Bahcall, 1977; Longair, 1998).

Table 1.3 shows a summary of three of the most common classification schemes and the corresponding properties of the clusters.

Zwicky et al. (1968) classified galaxy clusters according to their morphological appearance. The clusters are divided into three categories: compact, medium-compact, and open. A compact cluster is defined as a structure with a single outstanding concentration of its bright members, while an open cluster shows no obvious signs for

Property	Regular (early) type clusters	Intermediate type clusters	Irregular (late) type clusters
Zwicky type	Compact	Medium-compact	Open
Bautz-Morgan type	I, I – II, II	(II), II – III	(II – III), III
Rood-Sastry type	cD, B, (L), (C)	(L), (F), (C)	(F), I
Shape symmetry	Symmetrical	Intermediate	Irregular
Central concentration	High	Moderate	Low
Central profile	Steep gradient	Intermediate	Flat gradient
Segregation by Galaxy type	Segregation	Segregation	No segregation
Mass	Marginal evidence	Marginal evidence	No segregation
Galactic content	Elliptical-rich	Spiral-poor	Spiral-rich
E fraction	35%	20%	15%
S0 fraction	45%	50%	35%
S fraction	20%	30%	50%
X-ray luminosity	High	Intermediate	Low
Fraction of clusters	$\sim 1/3$	$\sim 1/3$	$\sim 1/3$

Table 1.3: Overview of cluster classifications and some of their related properties (after Bahcall 1977, Longair 1998, and Bahcall in Dekel & Ostriker 1999).

condensations. Medium-compact structures have either one central concentration that is less dense than in the case of a compact cluster, or more than one central concentration.

Bautz & Morgan (1970) created a classification scheme, that describes the relative contrast between the ten brightest cluster members. Type I clusters are dominated by one central cD galaxy. In type II clusters the brightest members have an appearance intermediate to cDs and normal giant ellipticals. Type III structures contain no dominant galaxies. Furthermore, there are two intermediate types, type I-II and II-III.

Rood & Sastry (1971) based their classification on the spatial distribution of the ten brightest cluster members. The Rood-Sastry scheme divides clusters into six different categories: cD, B, L, C, F, and I. The cD or supergiant type is dominated by a central cD galaxy. B or binary type clusters are dominated by a system of two bright galaxies. L or line type structures contain at least three members, that are positioned along a line. The C or core type clusters have at least four members, that lie within comparable separations of each other in the cluster center. In the F or flat type clusters several of the brightest members are assembled in a flattened arrangement, and in the I or irregular type no regular configuration can be detected.

A more comprehensive investigation of galaxy clusters and their properties is given in Bahcall (1977), Longair (1998), and Dekel & Ostriker (1999).

In principle, there are many possible approaches to look for galaxy clusters. All of them have one concept in common: They scan for typical cluster properties, like the

spatial distribution of the galaxies, the extended X-ray emission of the hot intracluster gas, the Sunyaev-Zeldovich effect, or lensing signals caused by the gravitational influence of the dark matter in the cluster. However, in this thesis, structure finding is limited to galaxy datasets, thus no information about the hot intracluster material, or the underlying dark matter distribution is available and can be used.

1.2 Photometric redshifts

With the development of the photometric redshift determination techniques (Baum, 1962; Koo, 1985; Brunner et al., 1999; Fernández-Soto et al., 1999; Benítez, 2000; Bender et al., 2001), approximate redshifts became available for all galaxies in a photometric multi-band survey, without having to do time-consuming spectroscopic follow-up observations.

The photometric redshift technique is a template matching algorithm, that follows Bayesian statistics. In principle, it compares an object's observed fluxes in the different passbands with a set of template spectral energy distributions (SED), that are redshifted and convolved with the respective filter transmission curves. These template SEDs can either be fully-, or semi-empirical, or can be derived from models. The algorithm itself yields a probability distribution as a function of template spectrum and redshift.

Photometric redshift galaxy surveys are increasing in popularity. The list of surveys using photometric redshifts includes the Hubble Deep Fields (HDF) North and South (Williams et al., 1996, 2000), the Las Campanas Infrared Survey (LCIR; Marzke et al. 1999, McCarthy et al. 2001a, b), the Calar Alto Deep Imaging Survey (CADIS; Wolf et al. 2001a), the Classifying Objects by Medium-Band Observations Survey (COMBO-17; Wolf et al. 2001b), the Munich Near-Infrared Cluster Survey (MUNICS; Drory et al. 2001b), and the Fors Deep Field (FDF; Heidt et al. 2003).

However, photometrically determined redshifts possess larger errors than spectroscopically determined ones. The accuracy of the photometric redshift determination is typically worse than the spectroscopic one by more than two orders of magnitude. Thus, the photometric redshift errors correspond to roughly 50 times the typical velocity dispersion of bound structures, like galaxy groups or clusters.

1.3 The Munich Near-Infrared Cluster Survey

The Munich Near-Infrared Cluster Survey is a wide-area medium-deep near-infrared selected galaxy survey, that is described in detail in Drory et al. (2001b). The survey consists of a K' -selected catalog down to an average of $m_{K'} \approx 19.1$ mag (50% completeness limit for point sources; Snigula et al. 2002). Near-infrared imaging was done in the K' and J passbands, covering an area of approximately 1 deg^2 . Additionally, observations have been done in the optical passbands I , R , V , and B , resulting in a total

Field	α (J2000)	δ (J2000)	Area [deg ²]	Limit $m_{K'}$ [mag]
S2F1	03:06:41	+00:00:47	0.0350	18.82
S2F5	03:06:41	-00:13:30	0.0363	19.12
S3F5	09:03:44	+30:01:27	0.0317	19.25
S5F1	10:24:00	+39:47:02	0.0341	19.09
S5F5	10:25:13	+39:46:53	0.0310	19.11
S6F1	11:55:47	+65:35:28	0.0344	18.75
S6F5	11:57:43	+65:35:24	0.0344	19.25
S7F5	13:34:39	+16:51:20	0.0350	19.06
Mean	–	–	–	19.06
Total	–	–	0.2719	–

Table 1.4: Observational properties of the eight MUNICS fields with the best photometric homogeneity and seeing. Their central right ascension α and declination δ for equinox 2000 are shown in the columns two and three. Column four contains the effective overlap area of all six filters. The K' band 50% completeness limit for point sources is given in column five (Snigula et al., 2002).

area of 0.27 deg² covered by all six filters in the eight fields with the best photometric homogeneity and seeing. The dataset used in this thesis was created in August 2003.

The prime motivation for constructing this survey was the detection of early type galaxies at redshifts $0 \leq z \leq 1.5$ for evolutionary studies, and the identification of groups or clusters of galaxies up to redshifts around unity.

Table 1.4 lists the central K' band right ascension and declination of the individual fields, the overlap area common to the observations in all six passbands, and the limiting K' band magnitude for 50% completeness for point sources (Snigula et al., 2002).

In Table 1.5 an overview of the number of detected objects, their separation into galaxies and stars, and the amount of available spectra are given for each field. The star-galaxy separation shown here is based on a combination of the morphological information given by the object detection algorithm YODA (Drory, 2003), and the best-fitting SED template resulting from the photometric redshift determination. In the following work, an object is defined as a galaxy, if it has an extended profile and the best-fitting template SED is galactic. An object is classified as a star, if its profile is determined only by the point-spread function, and its best-fitting SED is stellar. As can be seen from the table, about 77% of the objects are defined as galaxies and 12% are determined to be stars, that leaves only 11% of the objects with mixed classifications.

Photometric redshifts were determined for the MUNICS galaxies (Drory et al., 2003) using semi-empirical templates that match the photometric properties of the MUNICS survey. A spectroscopic follow-up was done for a subset of roughly 11% of the MUNICS galaxies by Feulner et al. (2003). This was used for the verification of the photometrically determined redshifts and showed a typical scatter in the photomet-

Field	Objects	Galaxies		Stars		Spectra	
		#	%	#	%	#	%
S2F1	665	479	72.0	98	14.7	183	27.5
S2F5	864	683	79.1	90	10.4	83	9.6
S3F5	789	603	76.4	107	13.6	10	1.3
S5F1	644	516	80.1	62	9.6	104	16.1
S5F5	485	360	74.2	58	12.0	15	3.1
S6F1	552	435	78.8	43	7.8	9	1.6
S6F5	867	677	78.1	97	11.2	145	16.7
S7F5	788	609	77.3	99	12.6	82	10.4
Total	5654	4362	77.1	654	11.6	631	11.2

Table 1.5: Galactic and stellar content in the MUNICS fields. The star-galaxy separation is based on a combination of the morphological information given by the object detection algorithm YODA (Drory, 2003), and the best-fitting SED template resulting from the photometric redshift determination. Furthermore the total number of detected objects, as well as the number of available stellar and galactic spectra are given.

ric redshift errors of $\sigma(\delta z) = 0.055(1+z)$ and a negligible mean bias of $\langle \delta z \rangle = 0.005$ (Drory et al., 2003). Fig. 1.1 shows the comparison of photometric and spectroscopic redshifts for a subsample of 512 objects (Drory et al., 2003).

Fig. 1.2 illustrates the redshift distribution of all the galaxies within the eight selected fields. The solid line denotes the distribution with respect to photometric redshifts. The dashed line shows the distribution of the spectroscopic galaxy subsample with respect to their spectroscopic redshifts.

1.4 Outline

Having introduced basic informations about galaxy groups and clusters in Sect. 1.1, Chapter 2 gives an overview of the fundamental principles of finding these structures in galaxy datasets. The difficulties of structure finding are discussed, as well as the survey features that influence the choice of the cluster finding technique for a given galaxy dataset. Finally, the most common or recent algorithms are listed and their working principles explained.

Chapter 3 gives a review of the original Huchra & Geller (1982) friends-of-friends (hereafter FOF) algorithm, one of the most frequently used cluster finding techniques. It was designed to search for number overdensities in spectroscopic galaxy surveys and has also been modified to look for structures in simulated galaxy datasets. Due to the relatively large errors, that are inherent to the photometric redshift determination, this algorithm cannot be applied to photometric redshift surveys. Thus, the newly designed extended friends-of-friends (EXT-FOF) algorithm is introduced. This new technique

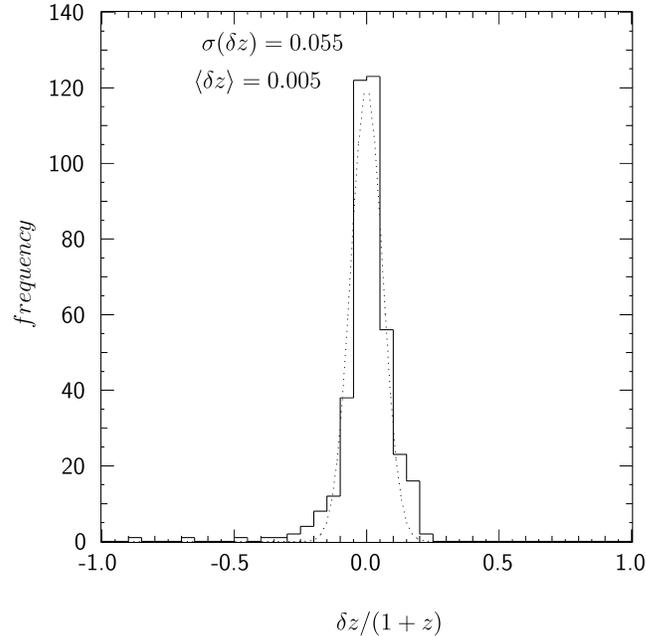


Figure 1.1: Comparison of photometric and spectroscopic redshifts for a subsample of the MUNICS objects after [Drory et al. \(2003\)](#). The solid curve shows the histogram of the redshift errors. The dotted line illustrates the best-fit Gaussian approximation to the error distribution. The typical scatter in the photometric redshift errors is $\sigma(\delta z) = 0.055(1+z)$, and the mean bias is $\langle \delta z \rangle = 0.005$.

is capable of dealing with the large uncertainties of the photometric redshifts, and is the major structure finding tool used in this thesis. Finally, the intricacies of choosing suitable linking parameters, a set of cut-off criteria intrinsic to all types of friends-of-friends algorithms are explained.

The EXT-FOF technique is tested in Chapter 4 to prove its usefulness as a structure finder. In order to do this, the algorithm is applied to two spectroscopic redshift surveys, and one survey with simulated photometric redshifts. It is shown, that for the spectroscopic surveys, the EXT-FOF and the original FOF yield almost identical results, and that all of the discrepancies in the group composition are well-understood, and bring about the applicability of the EXT-FOF for the photometric redshift surveys. Furthermore, it is proven that in case of photometric surveys, the EXT-FOF yields a rather conservative structure catalog, i.e. the recovery rate for finding clusters is very high.

Chapter 5 provides a number of tools, that allow the optimization of the above mentioned linking parameters. These tools include spectroscopic follow-ups of some of the galaxies, or the creation of Voronoi tessellations to measure local galaxy densities. The calculation of probability densities for having a cluster at a given coordinate in the dataset is also useful. Another method to determine the authenticity of the struc-

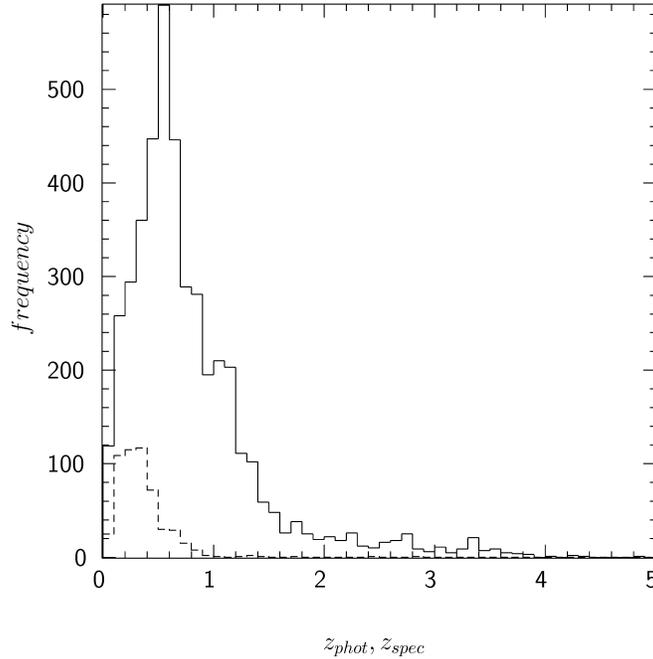


Figure 1.2: Redshift distribution of the MUNICS galaxies. The solid line shows the distribution with respect to photometric redshifts. The dashed line illustrates the distribution of the spectroscopic galaxy subsample with respect to their spectroscopic redshifts.

ture catalog is the creation of color-magnitude diagrams, that illustrate the presence of early type galaxies in clusters. These tools are furthermore useful for separating the likely clusters from any spurious detections.

Chapter 6 sees the application of the EXT-FOF algorithm to the MUNICS survey. First, an overview is given of the input galaxy catalog, discussing the geometric borders of the individual fields, as well as the details of the identification of detected objects as galaxies. The application of the new EXT-FOF algorithm to structure finding in MUNICS is described, presenting the utilized equations and the optimization of the linking parameters. An overview is given of the entire structure catalog and the field S6F5 is used as an example to demonstrate the quality assessment of the individual clusters. Furthermore, three of the structures are chosen for a more detailed examination, and their Voronoi tessellation and probability maps, as well as their color-magnitude diagrams are presented. Finally, the already obtained spectra are used for the spectroscopic verification of the EXT-FOF clusters.

As a conclusion, the results of this work are discussed in Chapter 7.

Chapter 2

Structure Finding – An Overview

In the following chapter, the basics of structure finding in galaxy datasets are explained. Section 2.1 deals with the fundamental problems of constructing group or cluster catalogs. Sect. 2.2 gives a list of the galaxy survey features, that influence the choice of the cluster finding technique. A scheme to distinguish the individual structure finding methods is shown in Sect. 2.3. The concluding Section 2.4 lists the most common and recent algorithms and their fundamental working principles.

2.1 Complications

The actual identification of groups or clusters of galaxies, either in an observed or simulated catalog, is a non-trivial enterprise. To begin with, there is no well-established definition for galaxy groups or clusters, leading to a certain level of arbitrariness in the choice of methods used for cluster finding. Every type of cluster finding technique looks for one or more typical cluster properties. It thus depends strongly on what is commonly perceived to be a useful cluster model, risking the danger of biasing the cluster catalog against atypical structures.

If the resulting cluster catalog is to be used for statistical analyses, it is not only necessary to have a catalog spanning a sufficiently large interval of redshift and reasonable sky area, but also to have well-defined selection criteria. The latter are hard to determine and are usually also strongly dependent on current cluster models.

2.2 Choice of the right technique

The structure finding technique used in a specific case depends on the specifications of the underlying galaxy dataset, like the wavelength of the observed radiation, that probes different cluster components. It also matters whether the galaxy survey was conducted in one wave-band, or is a multi-band survey, yielding valuable information

about galaxy colors in the latter case. Contiguity of the dataset plays a role, too, either allowing for, or denying the use of techniques that are sensitive to border effects. In the case of simulated datasets the exact phase-space information is available for all galaxies, while observed datasets always suffer from uncertainties. A very important feature of a galaxy survey is the precision of the distance information, i.e. whether the redshifts were determined by spectroscopy, or multi-band photometry, or whether there is no distance information at all. A reliable detection of structures in an observed galaxy catalog is very complicated if the galaxies have only photometric redshifts, that are less precise than spectroscopic ones by more than two orders of magnitude, or even worse, if they have none at all.

Of course, the used structure finding technique also depends on the purpose for which the groups or clusters are to be studied, in order to avoid biasing the resulting structure sample against the effects that are to be measured.

2.3 Differentiating the techniques

Multiple techniques have been created for the purpose of structure finding since Abell's pioneering work almost 50 years ago (Abell, 1958). Some of them have their own subtypes for various types of surveys or evolutionary stages. They may differ in their exact mathematical formulation, e.g. the type of distribution assumed to be followed by a parameter. They can also differ in their pre-selection of the input galaxy dataset, e.g. including only certain color or magnitude bins. Some subtypes are created to deal exclusively with three-dimensional information, while others can work with purely two-dimensional or photometric redshift datasets. Some subtypes even differ in the physical property used for distinguishing field from cluster galaxies.

Of course for every galaxy survey, the explicit values of the structure finding algorithm's cut-off or linking parameters have to be determined individually, taking into account the properties of the input datasets. The same is true for the completeness corrections.

All of the different methods have their advantages and drawbacks. It is essential to view them as complimentary.

2.4 General conspectus over structure finding algorithms

Giving a complete list of all types of structure finding techniques applied to galaxy datasets and their individual implementations would exceed the scope of this chapter. Only a brief overview over the more popular ones and their working principles will be given here, to better place the following work in the context of structure finding.

The pioneering work in this field was achieved by Abell (1958), who created the first galaxy cluster catalog by counting galaxies on photographic plates by eye. This

approach had the disadvantage of being very time-consuming and subjective. Since then, galaxy surveys have grown in depth, which led to the necessity of automated structure finding techniques, thus yielding much faster and objective methods. Due to their objectivity, the resulting cluster catalogs are better suited for statistical analyses.

Abell's work was followed by [Abell et al. \(1989\)](#), [Lidman & Peterson \(1996\)](#), and [Dalton et al. \(1997\)](#), amongst others. Their counts-in-cells method looks for surface density enhancements in circular apertures of fixed physical radius. A modification to this approach was used by [Zwicky et al. \(1968\)](#), [Turner & Gott \(1976\)](#), [Couch et al. \(1991\)](#), and [Plionis et al. \(1991\)](#), who searched for isopleth-contours¹ in surface density enhancements. These surface overdensity techniques are used solely on two-dimensional datasets and are prone to suffer from projection effects.

Another rather famous method is the hierarchical clustering technique. It identifies structures on the basis of optimizing typical cluster or group properties by combining individual galaxies into groups. So far, two different features have been used for optimization. One is the spatial separation of the galaxies, containing information about local number densities. The resulting centroid method was used by [Materne \(1978\)](#) and [Tully \(1980\)](#). The other feature is the local luminosity density, that includes information about the force between galaxy pairs. This so called force optimization was used by [Tully \(1980, 1987\)](#) and [Gourgoulhon et al. \(1992\)](#). The hierarchical clustering technique has the advantage of delivering a list of local concentrations within the groups, and superstructures at given values of the property that is to be optimized. It is used on spectroscopic redshift datasets.

One of the more recent approaches is the wavelet decomposition ([Slezak et al., 1990](#); [Paredes et al., 1995](#)). In this method the galaxy distribution is convolved with a so called analyzing wavelet. This transformation acts like a mathematical microscope, amplifying signals on given scale lengths, e.g. typical cluster scales. Up to now, the wavelet decomposition was used on spectroscopic redshift and two-dimensional datasets.

A technique looking for spherical density enhancements has been used by [Lacey & Cole \(1994\)](#) and [Cole & Lacey \(1996\)](#) on N-body simulations. Their algorithm searches iteratively for isopleth-contours in real-space.

Another more recently developed method is the matched or generalized likelihood filter ([Postman et al., 1996](#); [Schuecker & Boehringer, 1998](#); [Kepner et al., 1999](#); [Lobo et al., 2000](#)). On the basis of Bayesian probability theory, the galaxy catalog is convolved with a set of models, the filter, describing typical cluster properties, like the radial profile or the luminosity function, thus determining the likelihood of having clusters with given positions and richnesses in the input dataset. At least three subtypes have been created so far. The original matched filter algorithm ([Postman et al., 1996](#)) assumes a Gaussian distributed high background, whereas the generalized likelihood approach ([Schuecker & Boehringer, 1998](#)) is also valid for Poisson distributed low backgrounds. The adaptive matched filter ([Kepner et al., 1999](#)) is a combination of

¹contours of constant number density

the above, which uses a two-step approach, making it faster in the actual identification of the cluster positions and yielding more precise estimates of cluster redshifts and richnesses. The matched filter technique can be applied to every level of precision in the distance information, ranging from purely two-dimensional imaging surveys to N-body simulations. Unlike most other structure finding algorithms, this approach does not yield a list of galaxies belonging to each cluster. However, the resulting data, a list of three-dimensional cluster positions and richness classes, is a statistically important information.

The HOP algorithm (Eisenstein & Hut, 1998; Governato et al., 1999) was created exclusively for cluster finding in N-body simulations. Here galaxy chains are created by associating every galaxy with its densest neighbor in real space. All galaxies that are connected to the same maximum density object are called a group.

The cluster red sequence approach by Gladders & Yee (2000) works on multi-band imaging data. It was designed to deal with optical/near-infrared color information in addition to projected positions. It is based on the perception that all rich structures contain a population of passively evolving early type galaxies, that are lying on a linear sequence in color-magnitude space. By using a color and magnitude cut, fore- and background galaxies are excluded. The algorithm then looks for peaks in the weighted projected number density distribution of the resulting galaxy subset.

Voronoi tessellation has also been used as a means of detecting large-scale structures. In this case, a unique plane or volume partition of the galaxy distribution is done, yielding a measure for the local densities. This enables the identification of groups or clusters as significantly overdense contiguous regions. At least three subtypes have been devised. The Voronoi galaxy cluster finder by Ramella et al. (2001) works on purely two-dimensional imaging data. In order to minimize projection effects, this method includes a magnitude-bin pre-selection of the galaxies. The Voronoi tessellation technique (Kim et al., 2002) is applied to multi-band imaging surveys. Like in the case of the above mentioned cluster red sequence method, the input galaxy dataset is first run through a color-magnitude filter, eliminating most of the fore- or background objects. The Voronoi-Delaunay method (Marinoni et al., 2002) was designed for spectroscopic redshift surveys. It combines three-dimensional Voronoi diagrams and their dual, the Delaunay mesh, that contains informations about nearest neighbors, with each other. Voronoi tessellation cluster finders have the advantages of being scale-free, insensitive to the shape of clusters and very fast.

The cut-and-enhance method of Goto et al. (2002) is applied to multi-color imaging surveys. It utilizes color-magnitude and color-color cuts to pre-select early type galaxies in various redshift bins. A density enhancement algorithm up-weights galaxy pairs that are close both in projected position and in color. Finally, peaks in the enhanced density maps are identified as clusters.

Probably the most extensively used structure finding algorithm is the friends-of-friends technique (FOF) or percolation algorithm. This method looks for number density enhancements in three dimensions by searching for galaxy pairs that are closer to one another than a given cut-off separation. Up to now, the FOF approach has

only been applied to spectroscopic redshift surveys (Huchra & Geller, 1982; Geller & Huchra, 1983; Ramella et al., 1989, 1997; Trasarti-Battistoni, 1998; Giuricin et al., 2000; Merchán et al., 2000; Tucker et al., 2000; Ramella et al., 2002), or N-body simulations (Davis et al., 1985; Efstathiou et al., 1988; Lacey & Cole, 1994; Cole & Lacey, 1996; Valageas et al., 2000). In the first case, the spectroscopic redshift is a combination of the Hubble expansion and the peculiar line-of-sight velocity of the galaxy relative to the comoving space grid. So the spectroscopic redshift is a very good estimate of the true distance to the galaxy. In the second case, the exact position in three-dimensional space is known. The FOF approach is very straightforward, looking for neighboring galaxies in a basic and obvious way, that has no need for complicated cluster models. One of the resulting advantages is its insensitivity to cluster shape. However, like most of the other structure finding techniques, it has the drawback of being rather slow in its application. Changing the cut-off parameters requires a new run of the procedure, which makes the optimization of the cut-off values time-consuming.

Applying the original FOF algorithm to a galaxy catalog with photometric redshifts constitutes a problem, due to the relatively large uncertainty in the distance information. If the redshift errors are not taken into account, the resulting structure catalog will contain highly unphysical groups and clusters. If the errors are taken into account, yet the algorithm itself is not changed, the procedure yields structures, that are extremely elongated in redshift ($\Delta z \sim 1$). Thus it proves necessary to develop a new type of FOF algorithm, the extended friends-of-friends (EXT-FOF), for photometric redshift datasets, that includes the redshift error information and at the same time cuts down on unrealistically elongated structures.

Table 2.1 lists the above mentioned, most common or recent structure finding techniques, their subtypes, basic working principle and the level of precision in the distance measurement that the given algorithms have been applied to, so far. As can be seen clearly from this table, there exists a paucity of structure finding methods, that are designed to deal with photometric redshift surveys. One of the efforts of this work is to help fill this gap.

Technique	Subtypes	Basic working principle	Applicable to survey type			
			no z	z_{phot}	z_{spec}	N-body
Surface overdensity	<ul style="list-style-type: none"> • Counts-in-cells • Isopleth boundaries 	Looks for enhancements in projected density	+	-	-	-
Hierarchical	<ul style="list-style-type: none"> • Centroid optimization • Force optimization 	Orders galaxies hierarchically by optimizing typical cluster properties	-	-	+	-
Friends-of-friends	-	Searches for neighbors fulfilling minimal distance requirements	-	-	+	+
Wavelet decomposition	-	Analyzes signal on a given scale of the wavelet-transformed galaxy distribution	+	-	+	-
Spherical overdensity	-	Looks for spherical regions exceeding overdensity limit	-	-	-	+
Matched filter	<ul style="list-style-type: none"> • Matched filter • Adaptive matched filter • Generalized likelihood 	Determines probability for having cluster at a given position	+	+	+	+
HOP	-	Connects neighbors in densest regions with each other	-	-	-	+
Cluster red sequence	-	Looks for overdensities in projection and the color-magnitude plane	+	-	-	-
Voronoi tessellation	<ul style="list-style-type: none"> • Voronoi tessellation technique • Voronoi galaxy cluster finder • Voronoi-Delaunay method 	Uses Voronoi plane- or space-partition to determine overdense regions	+	-	+	-
Cut-and-enhance	-	Searches for projected and color overdensities in color-pre-selected distributions	+	-	-	-

Table 2.1: Structure finding algorithms. A rough overview over the most common and latest algorithms used for cluster finding. Their various subtypes and fundamental working principle are listed. The table also shows the level of precision in the galaxy distance information, that the algorithm can deal with. A “+” (“-”) sign denotes, that the technique has (not) been applied to this type of survey.

Chapter 3

Friends-of-Friends Algorithms

The friends-of-friends algorithm is one of the most frequently used cluster finding techniques. It was designed to find number overdensities in spectroscopic galaxy surveys and has also been modified to look for structures in simulated galaxy datasets. This algorithm is very straightforward and has no need for complicated cluster models. It looks for galaxy pairs, that are closer to one another than a given cut-off separation.

The following chapter introduces the newly designed, extended friends-of-friends algorithm (EXT-FOF). This new technique was created to look for groups and clusters of galaxies in photometric redshift datasets and is the major structure finding tool used in this thesis. A brief review of the original [Huchra & Geller \(1982\)](#) friends-of-friends (FOF) method is given in Sect. 3.1. In Section 3.2, the EXT-FOF technique is presented, and the major differences between the two methods are addressed. Sect. 3.3 gives a review of the basics of choosing the linking parameters, that are used in both friends-of-friends algorithms.

3.1 The Huchra & Geller friends-of-friends algorithm

The original FOF algorithm was created to look for groups and clusters of galaxies in the local magnitude-limited Center for Astrophysics Redshift Survey (CFA1; [Huchra & Geller 1982](#); [Huchra et al. 1983](#); [Geller & Huchra 1983](#)). It makes use of three basic pieces of information, the positions of the galaxies in right ascension and declination, and their spectroscopic redshifts.

A slightly modified version of the algorithm's flow chart given by [Huchra & Geller \(1982\)](#) is shown in Fig. 3.1. First, an object i from the galaxy catalog, that has not yet been assigned to any group, is chosen. Then friends j of that galaxy are searched for. They have to fulfill two criteria. Their projected separation D_{ij} from the first galaxy has to be less than a critical value D_L :

$$D_{ij} := 2 \sin \frac{\theta_{ij}}{2} \frac{\bar{V}}{H_0} \leq D_L, \quad (3.1)$$

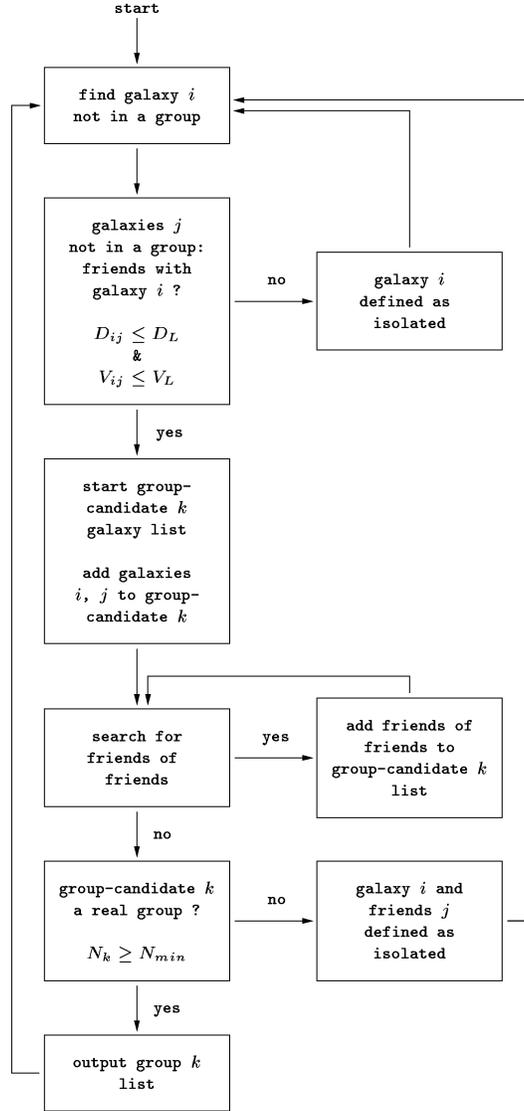


Figure 3.1: Flow chart for the Huchra & Geller friends-of-friends algorithm

where θ_{ij} is the angular separation between galaxies i and j , H_0 is the Hubble constant, and \bar{V} is the mean velocity of the galaxy pair:

$$\bar{V} = \frac{v_i + v_j}{2}, \quad (3.2)$$

v_i and v_j being the velocities of the individual galaxies. Furthermore, their separation V_{ij} in velocity- or redshift-space has to be less than a second critical value V_L :

$$V_{ij} := |v_i - v_j| \leq V_L. \quad (3.3)$$

If there are no friends for galaxy i , it is called an isolated object and removed from the catalog of possible cluster members. If friends are found for galaxy i , a list for group-candidate k is initiated, containing galaxy i and its friends j . The search for friends is extended to the surroundings of the galaxies j . All friends found are once more added to the group-candidate k . This is done until no further friends of friends can be detected. The group-candidate k is called a real group, if the number N_k of its members exceeds the limit N_{min} :

$$N_k \geq N_{min}. \quad (3.4)$$

If this is not the case, all members of said group-candidate are removed from the catalog. The next galaxy from the catalog is then taken and its surroundings are searched for friends.

The resulting structure catalog consists of a list of galaxies, that are either isolated, or assigned to one and only one structure. In contrast to the matched filter technique, the FOF approach does not yield a calibrated richness estimation.

The minimum group size limit N_{min} is usually set to three. Thus only galaxy pairs are excluded from the final group catalog.

As can be seen easily, this algorithm is commutative, i.e. the sequence of the input galaxies does not play a role, and yields reproducible results. Both are very important features for any structure finding technique.

3.2 The extended friends-of-friends algorithm

Due to the relatively large errors of the photometric redshifts, the FOF cannot be used for this type of dataset without some modifications. If the photometric redshift errors are not taken into account, a large fraction of the resulting groups and clusters are not physically bound objects. While the linking velocity V_L of the original FOF usually corresponds roughly to typical cluster velocity dispersions, photometric redshift errors are normally larger by a factor of 50. This implies that the deviation between the measured and the true distance to a galaxy might be off by about 50 times the linking velocity. Identification of these galaxies as group members is impossible with the original FOF technique. On the other hand, galaxies, whose true positions on the velocity axis are deviant by about $100 V_L$, could be combined into groups. Simply including the redshift errors of the individual galaxies into the linking criteria does not solve this dilemma either. The resulting structures would get unreasonably extended in redshift. Allowing for the necessary large value in the linking velocity would enable huge chains of galaxies to be linked together by the FOF algorithm. Thus, a modified version of the FOF algorithm, the extended friends-of-friends, or EXT-FOF, becomes necessary for photometric redshift galaxy surveys.

The EXT-FOF technique utilizes the same informations as the original FOF, i.e. right ascensions, declinations and, in this case, photometric redshifts of the galaxies. Furthermore the individual redshift error of each galaxy is taken into account.

The algorithm itself consists of three parts: In its inner loop, the technique is almost identical to the Huchra & Geller method, with slightly altered linking conditions. The search for groups is done in various a priori redshift-slices, meaning only galaxies that are compatible with a given value of redshift are taken into account for cluster finding. This redshift-slicing constitutes the outer loop. As a result, there is a catalog of structures for every redshift-slice. Every galaxy can only be a member of one structure of the catalog belonging to a given z -slice. Yet it can also belong to other structures in the other z -slices. The third part of the algorithm is the unification of all structures that have at least one member in common. This guarantees that in the final catalog every galaxy can only belong to one group or cluster.

The underlying idea for the modification of the linking conditions is the following: The a priori redshift slices z_{ini} approximate roughly the mean redshifts \bar{V} of the original FOF-technique. So eq. (3.1) changes into

$$D_{ij} := 2 \sin \frac{\theta_{ij}}{2} D(z_{ini}) \leq D_L. \quad (3.5)$$

$D(z_{ini})$ is the distance to z_{ini} in Mpc, being either

$$D(z_{ini}) := \frac{c z_{ini}}{H_0} \quad (3.6)$$

in the case of a low-redshift approximation, or more correctly

$$D(z_{ini}) := d_A(H_0, \Omega_M, \Omega_\Lambda, z_{ini}) \quad (3.7)$$

in the case of a cosmologically exact distance measurement. $d_A(H_0, \Omega_M, \Omega_\Lambda, z_{ini})$ is the angular distance to z_{ini} for a cosmology with given H_0 , matter term Ω_M and cosmological constant term Ω_Λ (Carroll et al., 1992). Eq. (3.3) translates into two equations, one for each of the galaxies i and j :

$$V_i := |v_i - c z_{ini}| \leq V_L/2, \quad (3.8)$$

$$V_j := |v_j - c z_{ini}| \leq V_L/2. \quad (3.9)$$

If the individual redshift errors δz_i and δz_j of the galaxies i and j are taken into account, the left sides of relations (3.8) and (3.9) change:

$$V_i \leq \left[(V_L/2)^2 + (c \delta z_i)^2 \right]^{1/2} =: V_{L,i} \quad (3.10)$$

$$V_j \leq \left[(V_L/2)^2 + (c \delta z_j)^2 \right]^{1/2} =: V_{L,j}. \quad (3.11)$$

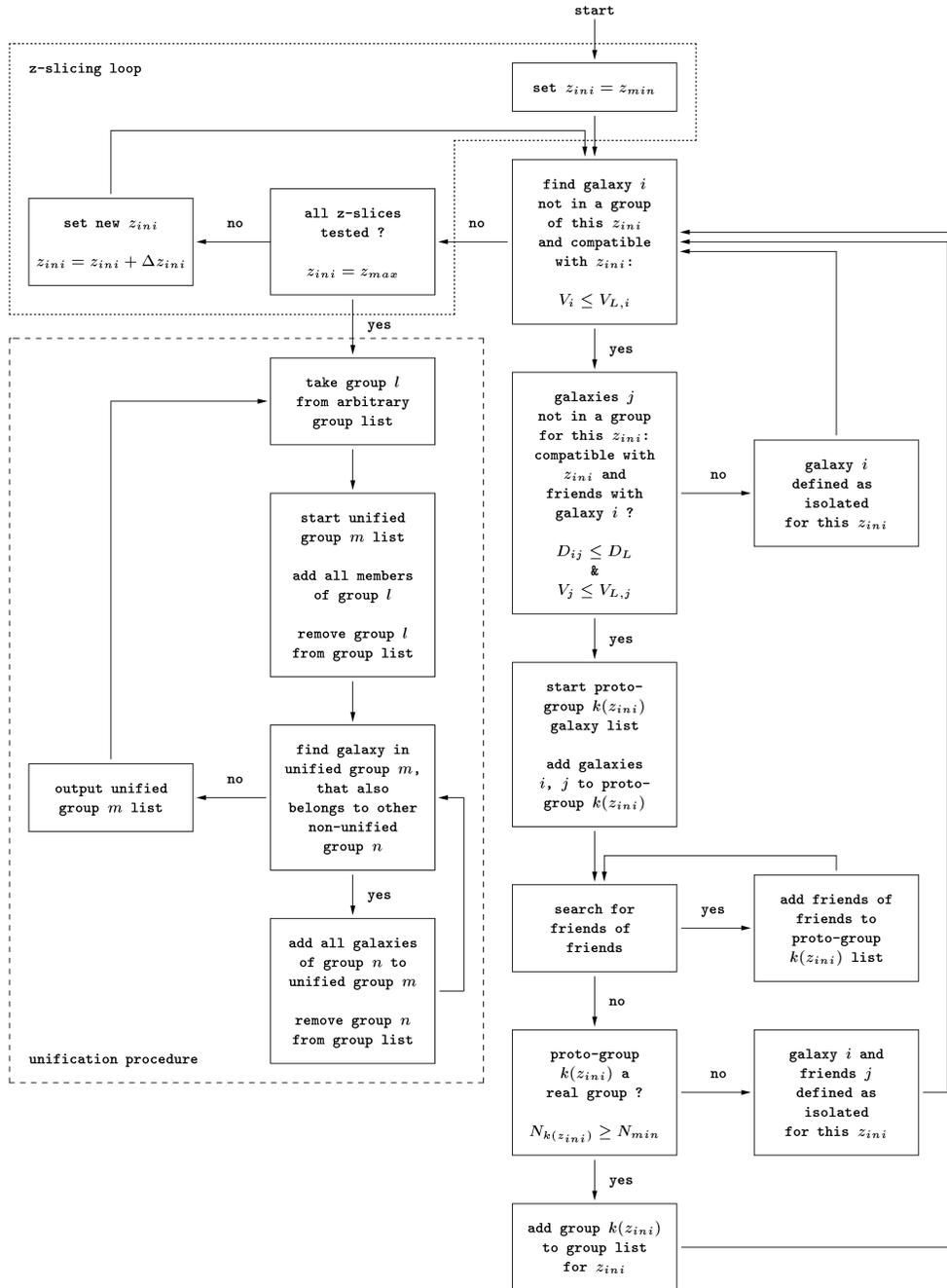


Figure 3.2: Flow chart for the extended friends-of-friends algorithm

Obviously, in the case of very small redshift errors, eqs. (3.8) and (3.9) are good approximations for eqs. (3.10) and (3.11).

A flow chart for the EXT-FOF algorithm is shown in Fig. 3.2. First, the minimal redshift $z_{ini} = z_{min}$ is chosen for structure finding. A catalog of groups, that belong to this z_{ini} -slice, is created as follows: An object i , that has not yet been assigned to any group belonging to this redshift-slice, is chosen from the catalog. Unlike the original FOF, this galaxy also has to be compatible with the chosen z_{ini} , meaning it has to fulfill eq. (3.10) (or eq. (3.8)). Then, friends j of that galaxy are searched for. They also have to be compatible with the chosen z_{ini} , i.e. they have to satisfy eq. (3.11) (or eq. (3.9)) and have to be closer to galaxy i than the cut-off distance D_L (eq. (3.5)). If no friends can be found, object i is moved to a list of isolated galaxies. If friends are found, a group-candidate $k(z_{ini})$ is initiated and the galaxies i and j are added to it. The surroundings of the galaxies j are searched for companions fulfilling eqs. (3.5) and (3.11) (or (3.9)), and the loop is repeated until no further friends can be found. A group-candidate is called a real group, if the number $N_{k(z_{ini})}$ of galaxies belonging to it satisfies the relation

$$N_{k(z_{ini})} \geq N_{min} \quad (3.12)$$

and the group is then added to the catalog of structures for the given redshift-slice. If this is not the case, the galaxy i and its friends are moved to the list of isolated objects. The next galaxy from the catalog is then taken and its surroundings are searched for friends. If all galaxies are assigned either to a group or to the isolated list and no galaxies are left in the input catalog, the search is continued in the next redshift-slice. For this, z_{ini} is increased by a value Δz_{ini} and the original input catalog of galaxies is restored. The search for friends is then repeated as described above. Once the value of z_{ini} has reached the limit z_{max} , the search is stopped. So far, this technique yields $(z_{max} - z_{min}) / \Delta z_{ini}$ individual catalogs of clusters. Since one galaxy can belong to different clusters in different redshift-slices, a unification of these groups is necessary to remove this ambiguity. To do this, a group l of any arbitrary z_{ini} -slice is taken. A, so far empty, unified group m is created. All galaxies of the above mentioned group l are added to the unified group m and group l is removed from the group catalog. Next, the unification procedure looks for a galaxy in unified group m that also belongs to another, non-unified, group n , irrespective of the redshift-slice that this group belongs to. If no such galaxy can be found, unified group m is complete, and the process is repeated for another group l . If there is a galaxy in m that belongs to a group n , as well, all galaxies of group n are added to the unified group m and n gets removed from the catalog. This loop is repeated until unified group m is complete and none of its members are contained in any of the non-unified groups any more. The result is a catalog of disjoint structures.

Like Huchra & Geller's version, this new algorithm is commutative and yields reproducible results.

One of this technique's objectives is to reproduce group catalogs found with the Huchra & Geller FOF in the case of spectroscopic datasets. To do this, the values of

z_{ini} have to approximate every possible value of \bar{V} for every galaxy pair. Theoretically, this can be reached by using a continuum of z_{ini} -values, ranging from the minimum \bar{V} of the two galaxies in the dataset that have the lowest redshifts to the maximum \bar{V} of the galaxy pair with highest redshifts. Of course this is not feasible in practice. Instead, a discrete set of z_{ini} values is used with a very small spacing Δz_{ini} . The boundaries z_{min} and z_{max} are simply set to the minimum and maximum redshifts of the galaxies used for structure finding.

There are two basic differences between the two FOF techniques: While the Huchra & Geller technique tests, whether the projected linking criterion, eq. (3.1), is fulfilled for a galaxy pair at only one redshift, the new algorithm tests, whether eq. (3.5) is satisfied for said galaxy pair at a multitude of redshifts. Furthermore, with the original algorithm, galaxies that are close to one another in projection can be linked together, even though not all of them are close to one another in redshift. Thus, chains of $N \geq N_{min}$ galaxies can be linked together into groups, that are very elongated in redshift. In the case of the EXT-FOF, all $N \geq N_{min}$ galaxies have to be compatible with a given redshift z_{ini} in order to be called a group. This makes the above mentioned outcome of elongated galaxy chains highly unlikely and is the reason why the new algorithm is well-suited for cluster finding in photometric redshift surveys.

The next chapter (Chapter 4) shows all of the possible effects resulting from these discrepancies and proves the validity of the EXT-FOF for photometric redshift datasets.

3.3 The linking parameters

A typical structure finding technique consists of two components. One is the algorithm itself, that has to be well-understood and suited for the specific galaxy dataset. The other is the set of key parameters, in case of the friends-of-friends the linking parameters, that have to be selected very carefully, in order to produce a realistic group or cluster catalog.

These linking parameters determine uniquely the resulting structure catalogs. The exact choice of these parameters depends on the properties of the input galaxy dataset, as can be seen in the Sects. 4.1.2 and 4.2.2. In principle, both parameters can be chosen to remain either fixed or to vary with distance. Using a set of fixed linking parameters is very convenient. However, except for taking into account the variation of D_{ij} with distance, all other distance-dependent selection effects are simply ignored. Varying the linking parameters with distance can for example compensate for the variation of the sampling of the galaxy luminosity function with redshift. This is only sensible for a relatively shallow galaxy survey, though.

In order to shed some light on the fine art of picking suitable linking parameters, this section gives an overview over the typical dimensions selected for the linking parameters, and explains where they originate from.

As was already mentioned, the underlying idea of the friends-of-friends technique is to search for peaks in the spatial density distribution. Thus, in principle, the algorithm has to utilize each galaxy's position information (x_1, x_2, x_3) in all three dimensions. However, not all of those coordinates are easily accessible.

3.3.1 Spectroscopic surveys

In case of a galaxy survey with spectroscopically determined redshifts, like the CFA1, the relevant observable quantities are the right ascension, the declination, and, of course, the spectroscopic redshift. In the used wavebands, the measurement of the right ascension and declination is extremely accurate, being of the order of less than $1''$, and as a result the two angular coordinates x_1 and x_2 are directly and precisely available from observations. The uncertainty in the spectroscopic redshift determination is relatively small, lying at roughly 10^{-4} , but the distance information x_3 is not uniquely extractable from the redshift. Neglecting the small measuring error, the observed redshift is composed of an expression describing the expansion of space at the distance x_3 of the galaxy, and the line-of-sight proper motion of the galaxy relative to the co-moving space grid. The latter component can be relatively large, up to $\delta z \approx 10^{-2}$ or equivalently $\delta v \approx 1000 \text{ km s}^{-1}$. That makes the assessment of the third coordinate x_3 less accurate than the determination of x_1 and x_2 by a factor of 10^5 . This explains why the original FOF algorithm uses two different linking criteria. The projected linking criterion, eq. (3.1), treats the angular coordinates x_1 and x_2 identically, while the less precise distance information x_3 is treated separately in the velocity linking criterion eq. (3.3).

Since structure finding rests upon a search for typical group or cluster properties, the order of magnitude of the linking parameters is defined by those typical properties, and if necessary by the measured errors and systematic uncertainties. Thus it is crucial to understand the composition of and the dominant effects in the linking criteria.

As the measurement of the angular coordinates x_1 and x_2 is very precise, the angular separation θ_{ij} between the galaxies i and j is well-determined. The mean velocity \bar{V} of the galaxy pair is an approximation for the mean distance to those galaxies and contains furthermore the above mentioned line-of-sight proper motion of both objects. This effect leads to a relative error in the determination of D_{ij} of the order of only 10%. As a result, the projected distance D_{ij} can be compared directly to typical cluster dimensions. The projected linking distance D_L is usually set to values ranging roughly from $1r_c \leq D_L \leq 1r_h$. r_c is the core radius of the galaxy distribution in a regular cluster, r_h is the cutoff radius, representing the distance from the center, where the projected density reaches zero. According to Bahcall (1977), r_c can be set to $r_c = 0.125 h^{-1} \text{ Mpc}$ for rich clusters, and r_h can be approximated by $r_h \approx 20r_c = 2.5 h^{-1} \text{ Mpc}$ (see also Sect. 1.1 for more details).

The velocity separation V_{ij} is a combination of the differences of the real distances x_3 to the galaxies i and j and their line-of-sight proper motions. If both galaxies are members of the same bound structure, the difference between their distances x_3 has to

be of the order of r_h , assuming spherical symmetry for rich clusters. Transforming the typical cluster extension r_h into the corresponding velocity v by using $v = H_0 r_h$ ¹ yields $v = 250 \text{ km s}^{-1}$. A comparison of v with the proper motion velocities of rich structures of up to approximately 1000 km s^{-1} shows that V_{ij} is dominated by the latter. The velocity dispersion of a cluster is defined as the standard deviation of the individual galaxy velocities relative to their mean velocity. Thus the velocity separation V_{ij} of two galaxies belonging to the same cluster should be comparable directly to cluster velocity dispersions. The linking velocity V_L is normally set to values typical for the velocity dispersion of groups and clusters, between roughly $100 \text{ km s}^{-1} \leq V_L \leq 1400 \text{ km s}^{-1}$ (see Bahcall in [DeKel & Ostriker 1999](#), or Sect. 1.1).

3.3.2 Simulated datasets

Although the following case is not part of this work, a brief overview is given of the linking parameter treatment in simulated datasets, for completeness reasons.

In galaxy datasets resulting from N-body simulations, the dichotomy in the treatment of the linking criteria disappears. When simulating galaxies with N-body codes, the entire phase-space information is available to the theorist. Thus, the coordinates x_1 , x_2 , and x_3 of every galaxy are accessible at any time and with equal precision. In this case, the two linking criteria eqs. (3.1) and (3.3) of the original FOF reduce to

$$R_{ij} \leq R_L, \quad (3.13)$$

with R_{ij} being the real space distance between galaxies i and j . R_L is the real space linking distance.

Due to the accuracy of the position determination, R_{ij} can be compared directly to interparticle distances that correspond to typical bound structures. Thus R_L is set to $R_L = b\bar{R}$, where \bar{R} is the mean interparticle separation and b is the fraction thereof. Naturally, in order to find high density environments b has to be set to $b < 1$. As was shown by [Davis et al. \(1985\)](#), [Frenk et al. \(1988\)](#), [Croft & Efstathiou \(1994\)](#), and [Lacey & Cole \(1994\)](#) among others, a choice of $0.1 \leq b \leq 0.3$ leads to reasonable structures.

3.3.3 Photometric surveys

Galaxy surveys with photometrically determined redshifts have the same precision in measuring right ascension and declination, and thus the coordinates x_1 and x_2 , as spectroscopic surveys. The accuracy of the redshift measurement is worse by roughly two to three orders of magnitude, though. As a result, the determination of x_3 is much more uncertain than in the case of spectroscopic redshift datasets. While the angular positions x_1 and x_2 can be treated equally in the new projected linking criterion eq. (3.5), the distance information x_3 has to be treated separately. Since the a priori redshift

¹the so-called Hubble law

z_{ini} was made to represent the mean redshift \bar{V} and thus in principle the mean distance to the galaxy pair, the FOF velocity linking criterion is split up into two criteria, i.e. one for each galaxy, eqs. (3.8) and (3.9), or (3.10) and (3.11) respectively, if the errors are taken into account. The velocity separations V_i and V_j are defined as the deviations of the measured galaxy velocities from the a priori redshift. In order to be compatible with the original FOF approach in the case of small photometric redshift errors, these values have to be compared to $V_L/2$. By augmenting the original FOF linking velocity V_L with the individual photometric redshift errors δz_i and δz_j , the new linking velocities $V_{L,i}$ and $V_{L,j}$ are created.

The calculation of D_{ij} is very precise, since θ_{ij} is well-determined and $D(z_{ini})$ was actually designed to represent the mean real-space distance to the galaxy pair. Thus it is obvious, that the projected linking distance D_L can be taken from the same range of values as in the case of the spectroscopic datasets.

The velocity separations V_i and V_j are composed of three terms. The first is the difference between the real distances x_3 of the galaxies and the mean real-space distance of i and j , represented by z_{ini} . As was shown in Sect. 3.3.1, this difference is very small, if the two galaxies are members of the same bound structure. The second term is the proper motion of the galaxies, that is typically bigger than the first by less than one order of magnitude. The third term is the photometric redshift error, dominating over the first by roughly two orders of magnitude. Thus, the linking velocities $V_{L,i}$ and $V_{L,j}$ are created, explicitly taking the individual redshift errors into account. Since these errors dominate over the typical proper motions, they also dominate over the typical V_L values of the original FOF. Thus, in principle $V_{L,i}$ and $V_{L,j}$ are determined by the photometric redshift errors δz_i and δz_j . By removing objects with very large redshift errors from the input galaxy sample, an upper limit can be effectively set to the linking velocities. This error-cutoff is set to the order of typical standard deviations of photometric redshift techniques, so not to remove too many galaxies from the input catalog.

Chapter 4

Testing the Extended Friends-of-Friends Algorithm

In order to prove the validity of the EXT-FOF technique for finding structures in photometric redshift datasets, the new algorithm is tested. This is done in two steps. First, the FOF and EXT-FOF are applied to spectroscopic redshift datasets and the resulting structures are compared to one another. It is proven, that in this case both algorithms yield almost identical results. Furthermore, it is shown that all of the discrepancies in the group composition are well-understood, and bring about the applicability of the EXT-FOF for the photometric redshift surveys. Second, the EXT-FOF is applied to a spectroscopic galaxy dataset with simulated photometric redshifts. The cluster catalog is then compared to the structures resulting from the application of the EXT-FOF to the unmodified spectroscopic dataset. It is shown that in the case of photometric surveys, the EXT-FOF yields a rather conservative structure catalog, i.e. the recovery rate for finding clusters is close to 100%.

In Section 4.1 the EXT-FOF is tested on the first Center for Astrophysics Redshift Survey (CFA1), while Sect. 4.2 uses the Las Campanas Redshift Survey (LCRS). Both of which are spectroscopic surveys. In Sect. 4.3 the EXT-FOF is tested on the LCRS with simulated photometric redshifts. A brief overview of the projects is given, followed by the recipes used for the creation of the cluster catalogs. Finally, the resulting structure catalogs are compared on an object-to-object basis and the discrepancies are analyzed.

4.1 Application to the CFA1 redshift survey

4.1.1 The galaxy catalog

The CFA1 survey (Huchra et al., 1983) is a magnitude limited, shallow, spectroscopic galaxy survey, containing all the galaxies of the Zwicky (Zwicky et al., 1968; Zwicky

& Zwicky, 1971) or Nilson (1973) catalogs, that satisfy the following selection criteria:

$$m_{pg} \leq 14.5 \text{ mag} , \quad (4.1)$$

and

$$b \geq 40^\circ, \delta \geq 0^\circ, \quad \text{or} \quad b \leq -30^\circ, \delta \geq -2.5^\circ. \quad (4.2)$$

The magnitudes are given in the B(0)-Zwicky system. b and δ are the galactic latitude and declination of the objects. The original catalog covers 2401 galaxies.

In this thesis, an electronic version of the CFA1 galaxy catalog, created in May 1997, is used. This new catalog contains 2396 galaxies with corrected magnitudes and redshifts.

4.1.2 Creation of the group catalogs

Following the recipe of Geller & Huchra (1983), two group catalogs are created, one by using the original FOF technique, the other by applying the new EXT-FOF algorithm. A summary of the Geller & Huchra (1983) treatment of the galaxy data and choice of the linking parameters is given below:

First, the given heliocentric galaxy velocities are corrected for a dipole virgocentric flow

$$V_V = V_{in} [\sin \delta_i \sin \delta_V + \cos \delta_i \cos \delta_V \cos (\alpha_i - \alpha_V)], \quad (4.3)$$

where V_{in} is the infall velocity, which is set to 300 km s^{-1} . α_V and δ_V are the right ascension and declination of the Virgo cluster ($\alpha_V = 12:28:42$; $\delta_V = 12:19:06$; J1950; Ramella et al. 1997), and α_i , and δ_i are the right ascension and declination of galaxy i . Furthermore, a correction for the solar motion with respect to the local group is made (Ramella et al., 1997):

$$V_G = 300 \text{ km s}^{-1} \sin l_i \cos b_i. \quad (4.4)$$

l_i and b_i are the galactic longitude and latitude of galaxy i . All galaxies with corrected velocities less than 300 km s^{-1} are given an indicative velocity of 300 km s^{-1} . This is done to avoid the singularity at 0 km s^{-1} .

Only galaxies with velocities less than 12000 km s^{-1} are accepted for the cluster search, while the mean velocity of the groups is limited to less than 8000 km s^{-1} .

A Hubble constant of $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ is used, to be in accordance with the approach used by Geller & Huchra (1983).

The linking parameters are chosen to vary with redshift, in order to compensate for the variation in the sampling of the luminosity function:

$$D_L = D_0 R, \quad (4.5)$$

and

$$V_L = V_0 R, \quad (4.6)$$

with

$$R = \left[\int_{-\infty}^{M_{ij}} \Phi(M) dM \left(\int_{-\infty}^{M_{lim}} \Phi(M) dM \right)^{-1} \right]^{-1/3}. \quad (4.7)$$

$\Phi(M)$ is the Schechter luminosity function with $\alpha = -1.30$, $M^* = -19.40$ mag, and $\Phi^* = 0.0143 \text{ Mpc}^{-3}$, derived directly from the CFA1 catalog.

$$M_{lim} = m_{lim} - 25 - 5 \log(V_{fid}/H_0), \quad (4.8)$$

and

$$M_{ij} = m_{lim} - 25 - 5 \log(\bar{V}/H_0), \quad (4.9)$$

in case of the original FOF technique. For cluster finding with the EXT-FOF algorithm, eq. (4.9) changes to

$$M_{ij} = m_{lim} - 25 - 5 \log(c z_{ini}/H_0). \quad (4.10)$$

The fiducial velocity V_{fid} is set to $V_{fid} = 1000 \text{ km s}^{-1}$, the projected separation D_0 and the velocity difference V_0 at the fiducial velocity are chosen as $D_0 = 0.52 \text{ Mpc}$ and $V_0 = 600 \text{ km s}^{-1}$. This parameter set corresponds to a search for structures having a number overdensity of 20, as described in Geller & Huchra (1983). The resulting variation of the linking criteria D_L and V_L as a function of \bar{V} , or z_{ini} respectively, is shown in Fig. 4.1.

The CFA1 group catalog published by Geller & Huchra (1983) contained only groups with more than two members, so $N_{min} = 3$ is used.

Since the CFA1 is a local galaxy survey, the low-redshift approximation, eq. (3.6), is used for the distance calculation in the EXT-FOF. Because of the very small redshift errors of this spectroscopic survey, eqs. (3.8) and (3.9) are utilized as the velocity linking criteria. The redshift-spacing is set to $\Delta z_{ini} = 10^{-5}$, to make sure that every possible value of \bar{V} can be approximated.

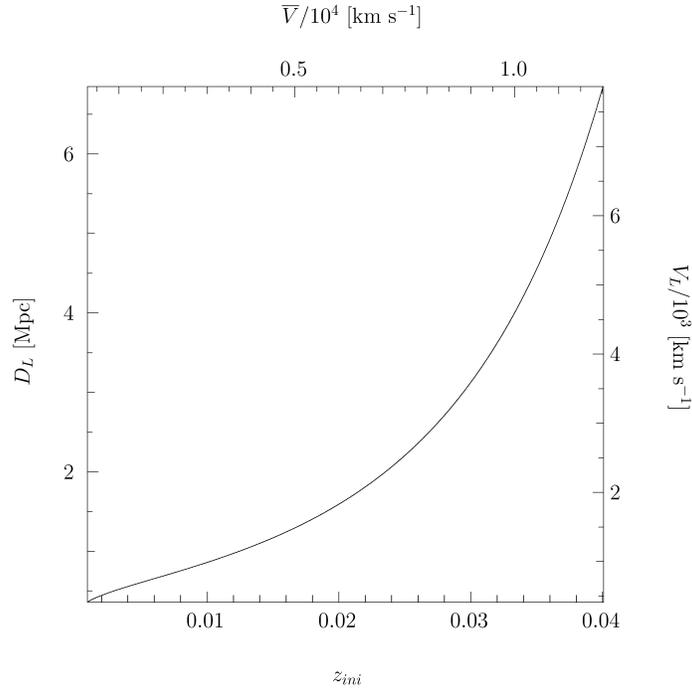


Figure 4.1: Variation of the linking criteria D_L and V_L for the CFA1 group catalogs, as a function of the FOF mean separation \bar{V} , or the EXT-FOF redshift-slicing z_{ini}

4.1.3 Comparison of the group catalogs

The catalog of structures, resulting from the application of the original FOF algorithm, contains 176 groups and clusters, comprising a total of 1480 galaxies. The composition of these FOF structures is slightly deviant from the ones published by [Geller & Huchra \(1983\)](#). This minor discrepancy is due to the different input galaxy catalogs.

Application of the EXT-FOF algorithm yields 165 groups, containing 1518 galaxies in total. Thus, in the case of the CFA1, the EXT-FOF shows a tendency to identify somewhat larger structures.

In order to compare the algorithms, an object-to-object comparison of the group members is done. Eight categories are defined for the level of agreement in the group composition:

- (1). The FOF and EXT-FOF group have identical composition
- (2). The group is only found with FOF, i.e. no galaxy in the group is a member of an EXT-FOF group
- (3). The group is only found with EXT-FOF, i.e. no galaxy in the group is a member of a FOF group

	Cat.	FOF	EXT-FOF	% FOF	% EXT-FOF
Total		176	165	100	100
Identical	1	122	122	69.3	73.9
Only FOF struct.	2	4	–	2.3	–
Only EXT-FOF struct.	3	–	8	–	4.8
FOF struct. larger than EXT-FOF	4	5	5	2.8	3.0
EXT-FOF struct. larger than FOF	5	12	12	6.8	7.3
FOF struct. is combination of EXT-FOF structs.	6	0	0	0.0	0.0
EXT-FOF struct. is combination of FOF structs.	7	28	13	15.9	7.9
FOF and EXT-FOF structs. have some elements in common	8	5	5	2.8	3.0
FOF structs. found with EXT-FOF algorithm		172	–	97.7	–

Table 4.1: Comparison of the CFA1 FOF and EXT-FOF structures. The first column provides a short description of the categories used. In the second column, the corresponding categories are listed for easy reference. The third and fourth column yield the number of FOF and EXT-FOF groups respectively, that fall under the specified category, while the fifth and sixth column show the percentage of these groups.

- (4). The EXT-FOF group is a subset of the FOF group, i.e. the FOF group has more members than the EXT-FOF, and none of those surplus members belong to any other EXT-FOF group
- (5). The FOF group is a subset of the EXT-FOF group, i.e. the EXT-FOF group has more members than the FOF, and none of those surplus members belong to any other FOF group
- (6). The FOF group is a combination of multiple EXT-FOF groups, and can also contain further galaxies that are not part of any other EXT-FOF group
- (7). The EXT-FOF group is a combination of multiple FOF groups, and can also contain further galaxies that are not part of any other FOF group
- (8). The FOF and EXT-FOF group have some members in common, but do not fall under any of the above mentioned criteria

This classification scheme facilitates the examination of the intrinsic characteristics of the two friends-of-friends algorithms, as can be seen in Section 4.1.4.

Table 4.1 shows the statistics of the comparison between the FOF and EXT-FOF structures. The first column provides a short description of the categories used. In the second column, the corresponding category numbers are listed for easy reference. The third and fourth column contain the number of FOF and EXT-FOF groups respectively, that fall under the specified category, while the fifth and sixth column show the percentage of these groups.

122 groups or clusters are recovered identically (category 1) by both algorithms, corresponding roughly to 72% of the FOF or EXT-FOF structures. 17 of the FOF groups are found by the EXT-FOF algorithm either as slightly enlarged (category 5), or reduced (category 4) structures. 28 FOF structures are recovered by the EXT-FOF as a combination of multiple FOF groups (category 7). Further five FOF structures are found by the EXT-FOF, suffering from combinations of the categories 4, 5, 6, and 7 (i.e. category 8). Only four of the 176 FOF groups are not found with the EXT-FOF algorithm (category 2). Thus, a total of 172 FOF structures are recovered by the EXT-FOF algorithm, corresponding to a recovery rate for the EXT-FOF algorithm of almost 98%. On the other hand, the EXT-FOF technique finds eight additional groups (category 3), that are not part of the FOF group catalog, leading to a spurious detection rate for the EXT-FOF technique of less than 5%, provided that the FOF technique delivers a complete structure catalog. A closer look at the category 2 groups shows, that they are extremely small. Three of them have only three members and one of them contains four galaxies. Furthermore, they are very elongated in the redshift direction, an attribute that might speak against them as possible galaxy groups, anyway. The category 3 EXT-FOF groups are also very small. Four of them have three members, the other four contain four galaxies. All things considered, in the case of the CFA1 group determination, the EXT-FOF algorithm yields results very similar to the FOF technique. A slight tendency to find larger groups with the EXT-FOF can be seen here. The cause for the discrepancies between the two friends-of-friends techniques is explained in Section 4.1.4.

4.1.4 Analysis of the discrepancies

The deviations between the two group catalogs can be ascribed to two effects: Either the EXT-FOF algorithm is able to link two galaxies together, that the original FOF cannot; Or the EXT-FOF algorithm can not link N_{min} (here: $N_{min} = 3$) objects together within one redshift-slice.

If the EXT-FOF algorithm finds a link between two galaxies, yet the FOF does not, then one of the following deviations can result:

- (i). The group can only be found with EXT-FOF (i.e. category 3)
- (ii). The group found with EXT-FOF contains more members than the corresponding FOF group (i.e. category 5)
- (iii). The EXT-FOF group is a combination of multiple FOF groups (i.e. category 7)

In the first case, the original FOF might find two linked galaxies, but cannot find a third object, that fulfills the linking criteria. Thus the number of objects within the original FOF group-candidate is too small, i.e. eq. (3.4) is violated, and the group is rejected. If the EXT-FOF method is able to find a third object in the current redshift-slice, that is linked to one of the galaxies in the pair, the number of group members satisfies

eq. (3.12) and the group is accepted. It is also possible, that the original FOF finds two galaxy pairs, but is not capable of connecting them, while the EXT-FOF algorithm is able to link two galaxies of each pair, thus creating a group-candidate containing four galaxies. Eq. (3.12) is then satisfied and the EXT-FOF group is accepted.

In the second case, the EXT-FOF is able to find one or more galaxies, that are defined as isolated by the FOF algorithm, linked to group members, resulting in a larger group in the EXT-FOF catalog.

In the third case, the original FOF finds two or more separate groups. If the EXT-FOF can find links between members of the different FOF groups, those groups are combined into one big EXT-FOF structure. Of course, it is also possible, that further galaxies, defined as isolated by the original FOF, get attached to such a big structure, following the reasoning of the second case. This leads to EXT-FOF groups, that consist of a number of FOF groups and some additional galaxies.

If the EXT-FOF algorithm cannot find N_{min} galaxies linked together within at least one of the redshift-slices, one of the following deviations can result:

- (i). The group can only be found with the original FOF (i.e. category 2)
- (ii). The group found with FOF contains more members than the corresponding EXT-FOF group (i.e. category 4)
- (iii). The FOF group is a combination of multiple EXT-FOF groups (i.e. category 6)

In the first case, the original FOF finds at least $N_{min} = 3$ objects that are linked together, satisfying eq. (3.4). The EXT-FOF, on the other hand, can only find pairs of galaxies in various redshift slices. Since every group-candidate has to fulfill eq. (3.12) in at least one redshift slice, these galaxy pairs are not taken into account in the unification process, and these galaxies are defined as isolated by the EXT-FOF algorithm.

In the second case, one galaxy, defined as group member by the FOF, cannot be attached to an EXT-FOF group. This is due to the fact, that there is no redshift slice, where this galaxy can be linked to at least two (in general: $N_{min} - 1$) members of this EXT-FOF structure. Thus the object is not entered into the pre-unification catalog and is not included in the EXT-FOF group. It is also possible, that more than one galaxy of the FOF group is missing in the corresponding EXT-FOF group. This can happen, if the above mentioned situation is true for all of those galaxies and, additionally, there is no redshift slice where two or more of those missing galaxies are linked with one of the EXT-FOF group members. However, in the case of the CFA1 group catalogs, none of the five FOF groups, that fall under category 4, are larger than their EXT-FOF counterpart by more than one group member.

In the third case, the EXT-FOF finds two or more separate groups, while the FOF is able to link them together into one structure. This can happen, if there are no redshift slices, where at least two members of one EXT-FOF group can be linked with at

least one member of another group. Following the reasoning of the second case, it is furthermore possible, that the large FOF group can contain additional members, that are defined as isolated by the EXT-FOF. As far as the CFA1 catalogs are concerned, this effect never shows up, though.

Category 8 groups are generated by a combination of the two effects: Members get lost in the EXT-FOF groups, because of too few linked objects at a given redshift slice, and at the same time new links can be found to galaxies, that are defined as isolated by the original FOF, while some subset of the FOF and EXT-FOF group is identical.

In one of the cases examined for category 3 of the CFA1 comparison, a group found with the original FOF, is removed from the group list because its mean velocity exceeds the limit. The EXT-FOF is able to find another galaxy linked to this group, pushing the group mean velocity below the limit. Thus, the group is included in the EXT-FOF, yet not in the FOF group catalog. However, this is a border effect and could in principle also work in the other direction, removing groups from the EXT-FOF catalog (i.e. category 2).

The reason, why the EXT-FOF algorithm is sometimes able to find links between objects, while the original FOF is not, can be explained as follows:

The equations used here for the calculation of the projected distance D_{ij} , eqs. (3.1), (3.5), and (3.6), show that $D_{ij} \propto \bar{V}$ and $D_{ij} \propto z_{ini}$, respectively. However, being varied with the luminosity function, the linking distance D_L is not proportional to the redshift. D_L is growing monotonically with z , switching from a convex curvature at low redshifts, to concave at higher redshifts. The slope of D_{ij} is set by the angular separation θ_{ij} of the two galaxies. There are three possible scenarios for the behavior of D_{ij} and D_L :

- Either θ_{ij} is so small, that $D_{ij} \leq D_L$ over the entire redshift range of $\bar{V} \in [300 \text{ km s}^{-1}, 12000 \text{ km s}^{-1}]$, or $z_{ini} \in [0.001, 0.04]$ respectively, i.e. the projected linking criterion, eq. (3.1) or (3.5), is fulfilled at \bar{V} and in all other redshift slices z_{ini} .
- Or θ_{ij} is so large, that $D_{ij} > D_L$ over the entire redshift range. This galaxy pair can neither be linked together by the FOF, nor by the EXT-FOF algorithm.
- Or θ_{ij} lies between the above mentioned values. In this case, D_{ij} intersects D_L twice.

The third scenario can lead to the links, that can only be found with the EXT-FOF algorithm. In this case, the redshift range can be divided into three areas, which is shown in Fig. 4.2. The figure shows a comparison of the projected linking distance D_L (solid line) and the projected separation D_{ij} (dashed line) of the galaxies i and j , as a function of the the mean distance \bar{V} , or the redshift slice z_{ini} , respectively. In the areas I and III, the projected linking criterion $D_{ij} \leq D_L$ is fulfilled. Whereas in area II it is never fulfilled. It becomes obvious, that it is critical in which of those areas the linking

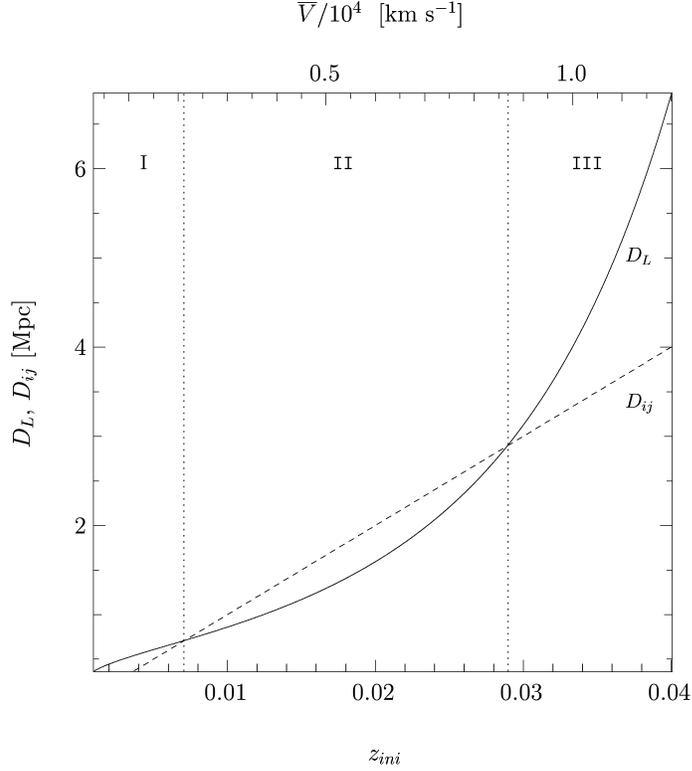


Figure 4.2: Comparison of the CFA1 projected linking distance D_L (solid line) and the projected separation D_{ij} (dashed line) of the galaxies i and j , as a function of the the mean distance \bar{V} , or the redshift slice z_{ini} , respectively. The angular separation θ_{ij} of the chosen galaxy pair leads to two intersections between D_L and D_{ij} , dividing the graph into three areas (dotted lines). The projected linking criterion $D_{ij} \leq D_L$ is satisfied in the areas I and III. It is not satisfied in area II.

criteria are tested. The original FOF can only test the projected linking criterion at a fixed redshift \bar{V} , set by the galaxy pair. The EXT-FOF, on the other hand, tests the linking criteria at all redshifts. If the EXT-FOF redshift-space linking criteria, eqs. (3.8) and (3.9), can be fulfilled for a given galaxy pair, then they are generally satisfied for an interval of z_{ini} values. If eq. (3.5) can also be satisfied for at least one of those redshifts z_{ini} , then all criteria are fulfilled and the galaxy pair is found as linked by the EXT-FOF. Thus, only the EXT-FOF and not the original FOF is able to link a galaxy pair in one of the two following situations: The mean velocity \bar{V} is lying in the area II, and the redshift interval of z_{ini} values, that satisfy eqs. (3.8) and (3.9), reaches either into area I (Fig. 4.3, upper panel), or into area III (Fig. 4.3, lower panel). As a result, the projected linking criterion is not satisfied for the original FOF, while both linking criteria can be fulfilled in at least one z_{ini} slice, in the case of the EXT-FOF.

It can be easily seen, that this effect does not exist, if D_L and D_{ij} follow the same variation with redshift, $D_L(z) \propto D_{ij}(z)$.

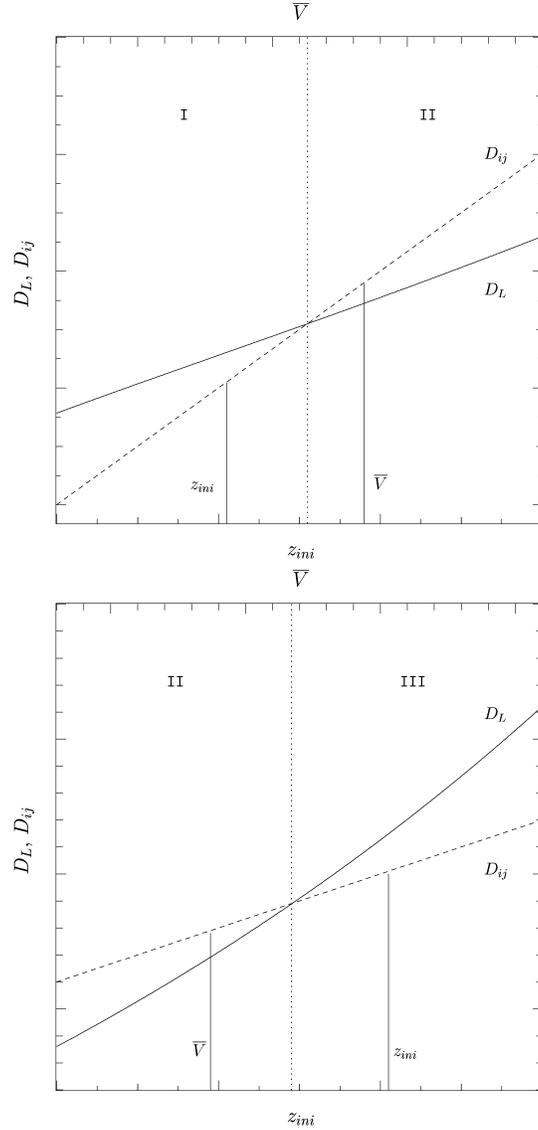


Figure 4.3: Comparison of the CFA1 projected linking distance D_L (solid line) and projected separation D_{ij} (dashed line), as a function of the the mean distance \bar{V} , or the redshift slice z_{ini} , for two example galaxy pairs. These schematic drawings show magnified excerpts of the transition regions (dotted lines) between the areas I and II (upper panel), and the areas II and III (lower panel), that are described in Fig. 4.2. **Upper panel:** The projected separation D_{ij} at the mean redshift \bar{V} (solid vertical line in area II) of the galaxies i and j , does not fulfill the projected linking criterion $D_{ij} \leq D_L$. In the case shown here, the range of z_{ini} values, satisfying the EXT-FOF redshift-space linking criteria, eqs. (3.8) and (3.9), reaches into area I. Thus, $D_{ij} \leq D_L$ is fulfilled for this redshift, and the galaxies i and j are found as linked by the EXT-FOF algorithm. **Lower panel:** \bar{V} is again lying in area II. The range of z_{ini} values, that satisfy the EXT-FOF redshift-space linking criteria, reach into area III in this case.

In the following, it is explained why the EXT-FOF algorithm sometimes can not link together $N_{min} = 3$ objects within at least one z_{ini} slice, and why this is the reason for some of the above mentioned deviations from the original FOF.

The original FOF algorithm can move relatively freely along the redshift axis, when looking for friends. For example, a set of galaxies a , b , and c is considered, with given velocities v_a , v_b , and v_c . For matters of simplicity, $v_a < v_b < v_c$ is assumed. The FOF calls those three galaxies a group, if there is at least one galaxy among them, that can be linked to the other two. If, for example, the galaxies a and b , and b and c fulfill the linking criteria, i.e. eqs. (3.1) and (3.3), then a , b and c are linked together. It is not necessary, that a and c also satisfy the linking criteria, and $V_{ac} > V_L$ is possible. Thus, in principle, the FOF allows for very elongated chains of galaxies along the redshift axis, as long as these galaxies have a small projected separation and the velocity difference V_{ij} between next neighbors satisfies eq. (3.3). This effect is what makes the original FOF unqualified for cluster finding in datasets with rather uncertain redshifts, like photometric galaxy catalogs.

The EXT-FOF algorithm cannot move so freely along the redshift axis. By only looking for friends among galaxies, that are compatible with a given redshift slice, the probability of finding very elongated structures is efficiently reduced.

To illustrate the problem, Fig. 4.4 shows the case of a set of galaxies, a , b , and c , that are found as a group by the FOF, yet not by the EXT-FOF. For simplification, a galaxy set is chosen, that has a very small spread in the projected distance, so that eqs. (3.1) and (3.5) are fulfilled for all three galaxies, and it only has to be tested, whether the velocity linking criteria are fulfilled. The figure shows the variation of the linking velocity V_L (solid line) with redshift z_{ini} , or mean velocity \bar{V} , respectively. The individual velocities v_a , v_b , and v_c of the three galaxies are shown on the lower scale. The filled squares are the original FOF velocity separations V_{ab} , V_{bc} , and V_{ac} , resulting from eq. (3.3), and plotted at the corresponding mean velocities. It becomes obvious, that V_{ab} and V_{bc} satisfy the velocity linking criterion, and together with the already fulfilled projected linking criterion, the original FOF combines the three galaxies into a group. The dashed lines denote the EXT-FOF velocity separations. They are plotted as $2V_a$, $2V_b$, and $2V_c$ to make an easy comparison with V_L possible. The dotted lines show the areas where eqs. (3.8) and (3.9) are satisfied for two galaxies: In area I, the requirement is fulfilled for the galaxies a and b , and in area II it is fulfilled for the galaxies b and c . There is no redshift slice, where all three galaxies satisfy the EXT-FOF velocity linking requirement $V_i \leq V_L/2$. As a result, the EXT-FOF can only find the galaxies a and b , or b and c linked together, in the z_{ini} ranges given by the areas I, or II. Those galaxy pairs do not satisfy eq. (3.12) with $N_{min} = 3$, and thus are not added to the group list.

Theoretically, another effect, that could remove a single link between two galaxies, is conceivable. It would also lead to the observed phenomena of category 2, 4, 6, or 8. If the step size Δz_{ini} of the EXT-FOF redshift slicing is chosen too large, not every possible value of \bar{V} can be approximated anymore, and a link found with the original FOF could not be recovered with the EXT-FOF algorithm. This is never the case in

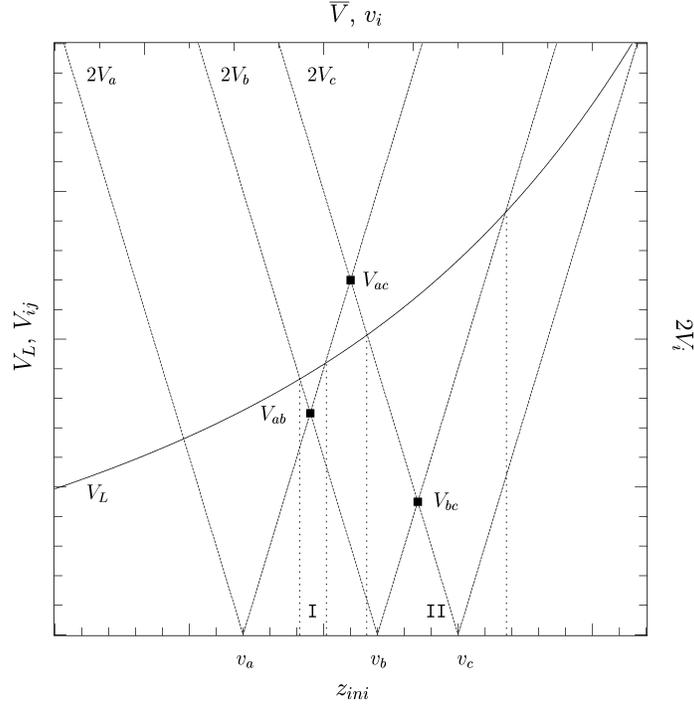


Figure 4.4: Comparison of the CFA1 linking velocity V_L (solid line) with the original FOF (filled squares) and EXT-FOF (dashed lines) velocity separations, V_{ij} and V_i , for a set of three galaxies, a , b , and c , that can only be linked with the original FOF, yet not with the EXT-FOF. The lower scale shows the redshift z_{ini} , the upper scale shows the corresponding mean velocity \bar{V} . The individual velocities v_a , v_b , and v_c of the three galaxies are also shown on the lower scale. The original FOF velocity separations V_{ab} , V_{bc} , and V_{ac} , resulting from eq. (3.3), are plotted at the corresponding mean velocities. The EXT-FOF velocity separations are plotted as $2V_a$, $2V_b$, and $2V_c$ to make an easy comparison with V_L possible. The dotted lines show the areas, where eqs. (3.8) and (3.9) are satisfied for two galaxies: In area I, the requirement is fulfilled for the galaxies a and b , and in area II it is fulfilled for the galaxies b and c . In no redshift slice can all $N_{min} = 3$ galaxies satisfy the EXT-FOF velocity linking requirement $V_i \leq V_L/2$.

one of the examined catalogs, though, proving that the chosen step size is sufficiently small and does not lead to unwanted numerical effects.

4.2 Application to the LCRS redshift survey

4.2.1 The galaxy catalog

The LCRS (Shectman et al., 1996) is an R band selected, spectroscopic galaxy survey. It consists of six $1.5^\circ \times 80^\circ$ strips, three of which are lying in the north, the other three

in the south Galactic cap. The strips located in the north Galactic cap are centered at declinations $\delta = -3^\circ, -6^\circ, \text{ and } -12^\circ$ and range in right ascension from $\alpha = 10^{\text{h}}$ to 15.5^{h} . The southern strips are centered at $\delta = -39^\circ, -42^\circ, \text{ and } -45^\circ$ and range in right ascension from $\alpha = 21^{\text{h}}$ to 4.5^{h} . The survey covers a total area of over 700 deg^2 , divided into 327 individual fields. For 120 of those fields spectroscopy is done with a 50 fiber multiobject spectrograph (MOS), and they have nominal apparent magnitude limits of

$$16.0 \text{ mag} \leq m_R \leq 17.3 \text{ mag}. \quad (4.11)$$

For the remaining 207 fields spectroscopy is done with a 112 fiber MOS, with nominal apparent magnitude limits of

$$15.0 \text{ mag} \leq m_R \leq 17.7 \text{ mag}. \quad (4.12)$$

For every setup, all of the fibers are used, but since each of the fields is observed only once, there is a field-to-field variation in the selection criteria. The protective tubing around the individual fibers of the MOS makes it impossible to observe spectra of objects, that are closer to each other than $55''$, leading to $55''$ “orphans”, that have no spectroscopic redshifts. The surveyed galaxies are sampled randomly within each field. The sampling is rather dense, with an average of roughly 70% of the magnitude-limited object list.

The catalog used here is a combination of the 23695 galaxies that have spectroscopic redshifts and the 1694 $55''$ “orphans”, resulting in a total of 25389 galaxies. Both datasets only include objects, that are lying within the geometric and photometric boundaries of the survey. Artificial redshifts are assigned to the $55''$ “orphans”, by giving each of them the redshift of its nearest neighbor, convolved with a Gaussian of width $\sigma = 200 \text{ km s}^{-1}$ (Tucker et al. 2000, Tucker, priv. comm.).

4.2.2 Creation of the group catalogs

Following the recipe of Tucker et al. (2000), two structure catalogs are once more created, one with the original FOF algorithm, the other with the EXT-FOF. In the following, an overview of the Tucker et al. (2000) treatment of the data is shown. Furthermore, the application of this recipe to the EXT-FOF is explained.

First, all galaxy velocities are corrected for motion relative to the dipole moment of the cosmic microwave background. The equation used by Tucker et al. (2000) is given in galactic coordinates

$$V_V = V_\odot [\sin b_i \sin b_\odot + \cos b_i \cos b_\odot \cos (l_i - l_\odot)], \quad (4.13)$$

where $V_\odot = 368.9 \text{ km s}^{-1}$, $b_\odot = 48.05^\circ$, and $l_\odot = 264.33^\circ$ (Lineweaver et al. 1996, Tucker, priv. comm.), and b_i, l_i are the galactic latitude and longitude of galaxy i . This equation corresponds to eq. (4.3), which is given in equatorial coordinates.

The set of galaxies used for cluster finding is limited to objects having corrected velocities v_i such that

$$7500 \text{ km s}^{-1} \leq v_i < 50000 \text{ km s}^{-1}, \quad (4.14)$$

and absolute magnitudes M_i between

$$-22.5 \text{ mag} + 5 \log h \leq M_i < -17.5 \text{ mag} + 5 \log h. \quad (4.15)$$

Not taking any color corrections into account, M_i is simply set to

$$M_i = m_i - 5 \log (d_L (H_0, \Omega_M, \Omega_\Lambda, v_i/c)) + 5, \quad (4.16)$$

where m_i is the apparent magnitude of galaxy i , and $d_L (H_0, \Omega_M, \Omega_\Lambda, v_i/c)$ is the luminosity distance for the given cosmology to galaxy i (Carroll et al., 1992). Since the LCRS is relatively deep, going out to redshifts of roughly 0.2, it becomes necessary to use cosmologically correct expressions.

Only groups with mean velocities \bar{V}_g such that

$$10000 \text{ km s}^{-1} \leq \bar{V}_g < 45000 \text{ km s}^{-1} \quad (4.17)$$

are accepted in the final group catalog.

A flat cosmology with $H_0 = 65 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ is used.

The distance equations for the original FOF have to be modified to take the cosmologically correct expressions into account. Eq. (3.1) becomes

$$D_{ij} = 2 \sin \frac{\theta_{ij}}{2} D_{ave} \leq D_L, \quad (4.18)$$

with the mean comoving angular distance D_{ave} between the galaxies i and j

$$D_{ave} := \frac{d_A (H_0, \Omega_M, \Omega_\Lambda, v_i/c) + d_A (H_0, \Omega_M, \Omega_\Lambda, v_j/c)}{2} \quad (4.19)$$

replacing the \bar{V} of eq. (3.2).

In the case of the EXT-FOF, $D(z_{mi})$ of eq. (3.7) is used for the calculation of the projected distance D_{ij} . The individual spectroscopic redshift errors of the galaxies are not yet taken into account, thus eqs. (3.8) and (3.9) still apply. A stepsize of $\Delta z_{mi} = 10^{-5}$ is used.

The linking parameters are varied with redshift. Yet, due to the field-to-field variations in the sampling fraction, the photometric limits, and the number of MOS fibers,

this dependency is more complicated than in the case of the CFA1. Eqs. (4.5) and (4.6) are used for the variation of D_L and V_L , and R is given by

$$R = \left[\frac{n^{exp}(f, D)}{n_{fid}^{exp}} \right]^{-1/3}. \quad (4.20)$$

$n^{exp}(f, D)$ is the expected galaxy number density in the field f at a comoving distance D . D is either D_{ave} in the case of the original FOF, or $D(z_{ini})$ in the case of the EXT-FOF. n_{fid}^{exp} is the galaxy number density $n^{exp}(f, D)$ in a fiducial field at a given fiducial distance. $n^{exp}(f, D)$ is calculated with the help of the Schechter luminosity function $\Phi(M)$:

$$n^{exp}(f, D) = F \int_{M_{min}}^{M_{max}} \Phi(M) dM, \quad (4.21)$$

where F is the sampling fraction of field f . M_{min} and M_{max} are the absolute photometric limits of this field, at the given comoving distance D . When comparing two galaxies from two different fields f_a , and f_b , the average of the expected galaxy densities is taken:

$$n^{exp}(f, D) = \frac{n^{exp}(f_a, D) + n^{exp}(f_b, D)}{2}. \quad (4.22)$$

Two Schechter luminosity functions are used (Lin et al., 1996; Tucker et al., 1997), following the approach of Tucker et al. (1997). One is valid for the 50 fiber fields in the southern Galactic cap, and has $\alpha = -0.74$, $M^* = -20.55 \text{ mag} + 5 \log h$, and $\Phi^* = 0.016 h^3 \text{ Mpc}^{-3}$. The other is used for the 50 fiber fields in the northern Galactic cap, as well as for all 112 fiber fields. Its parameters are $\alpha = -0.70$, $M^* = -20.29 \text{ mag} + 5 \log h$, and $\Phi^* = 0.019 h^3 \text{ Mpc}^{-3}$.

For the fiducial field the latter luminosity function is used. Furthermore, this field has a sampling fraction of $F = 1$, and apparent magnitude limits of

$$15.0 \text{ mag} \leq m_R \leq 17.7 \text{ mag}. \quad (4.23)$$

The fiducial redshift is set to $cz_{fid} = 30000 \text{ km s}^{-1}$.

Only groups having at least $N_{min} = 3$ members are searched for. The projected separation D_0 and the velocity difference V_0 are chosen as $D_0 = 0.715 h^{-1} \text{ Mpc}$ and $V_0 = 500 \text{ km s}^{-1}$, respectively, again following Tucker et al. (1997). This choice of linking parameters corresponds to a search for structures with a number overdensity of 80, a much higher overdensity than in the case of the CFA1 structure catalog. Consequently, the fraction of cluster galaxies will be lower in this catalog than in the CFA1 catalog.

The resulting variation of the linking criteria D_L and V_L as a function of redshift z is shown in Fig. 4.5. This graphic is created for an example field having a MOS setup with 112 fibers and photometric limits $m_{min} = 15.0 \text{ mag}$, $m_{max} = 17.7 \text{ mag}$. A typical sampling fraction of $F = 0.7$ is used.

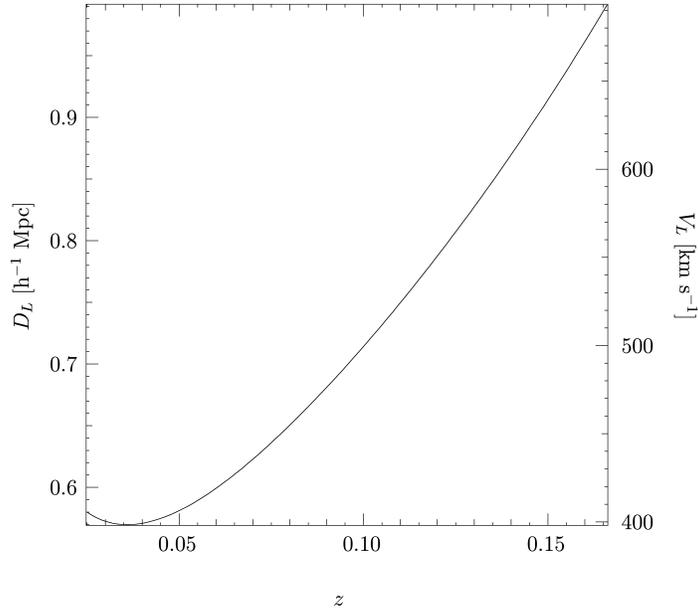


Figure 4.5: Variation of the linking criteria D_L and V_L for the LCRS group catalogs, as a function of redshift z . As a reminder, in this case R is not only a function of distance, but also depends on the field. A typical example field is shown, having a MOS setup with 112 fibers, a sampling fraction of $F = 0.7$, and photometric limits of $m_{min} = 15.0$ mag, $m_{max} = 17.7$ mag.

4.2.3 Comparison of the group catalogs

Application of the original FOF yields a total of 1367 groups. 6747 galaxies are contained in these groups. The structures are slightly deviant from the ones published by Tucker et al. (2000). These discrepancies in the group composition arise from three differences in the treatment of the data: First, no color corrections are included in eq. (4.16), unlike Tucker et al. (2000). Second, for the determination of D_{ave} and $D(z_{ini})$ the comoving angular distance $d_A(H_0, \Omega_M, \Omega_\Lambda, z)$ is used, while Tucker et al. (2000) utilized the proper motion distance $d_M(H_0, \Omega_M, \Omega_\Lambda, z)$. Third, group memberships are determined in a flat universe with a low matter content, whereas Tucker et al. (2000) used an Einstein-de Sitter universe.

The catalog resulting from the EXT-FOF application contains 1285 groups, with a total of 6337 galaxies.

Table 4.2 shows the statistics of the comparison between the FOF and EXT-FOF structures in the case of the LCRS. The same object-to-object comparison is used as described in Section (4.1.3).

1280 of the FOF groups are recovered with the EXT-FOF algorithm, corresponding to a recovery rate of almost 94%. The EXT-FOF, on the other hand, finds only

	Cat.	FOF	EXT-FOF	% FOF	% EXT-FOF
Total		1367	1285	100	100
Identical	1	1145	1145	83.8	89.1
Only FOF struct.	2	87	–	6.4	–
Only EXT-FOF struct.	3	–	1	–	0.1
FOF struct. larger than EXT-FOF	4	120	120	8.8	9.3
EXT-FOF struct. larger than FOF	5	4	4	0.3	0.3
FOF struct. is combination of EXT-FOF structs.	6	5	10	0.4	0.8
EXT-FOF struct. is combination of FOF structs.	7	2	1	0.2	0.1
FOF and EXT-FOF structs. have some elements in common	8	4	4	0.3	0.3
FOF structs. found with EXT-FOF algorithm		1280	–	93.6	–

Table 4.2: Comparison of the LCRS FOF and EXT-FOF structures. The first column provides a short description of the categories used. In the second column, the corresponding categories are listed for easy reference. The third and fourth column yield the number of FOF and EXT-FOF groups respectively, that fall under the specified category, while the fifth and sixth column show the percentage of these groups.

one group, that is not a member of the FOF group catalog (category 3), leading to a spurious detection rate of less than 1%. 1145 structures are found identical by both algorithms, so the rate of identical recoveries lies at roughly 86%. Of the 87 FOF groups, that have no EXT-FOF counterparts (category 2), 77 have only three members, and the remaining ten contain four objects each. Those groups are all very elongated in redshift, making their group status questionable, anyway. The one group falling under category 3 is also extremely small, having only three members. The above comparison shows, that both algorithms yield very similar results.

4.2.4 Analysis of the discrepancies

Like in the case of the CFA1, all of the deviations mentioned above can be explained by the two effects, that are described in Section 4.1.4.

The reason for finding an additional link with the EXT-FOF in the LCRS catalog is very similar to the case of the CFA1. The equations used for the determination of the projected distance D_{ij} , eqs. (4.18), (4.19), and (3.5), (3.7), show that D_{ij} grows with $d_A(H_0, \Omega_M, \Omega_\Lambda, z)$, while the linking distance D_L is varied with the luminosity function. D_{ij} has a convex curvature, D_L is concave. The slope of D_{ij} is once more set by the angular separation θ_{ij} , and there are again three possible scenarios for the behavior of D_{ij} and D_L . Very small angular separations lead to $D_{ij} \leq D_L$, and very large values of θ_{ij} lead to $D_{ij} > D_L$ for the entire redshift range, that is used for cluster finding. Fig. 4.6 shows the case of an example galaxy pair with a medium sized angular

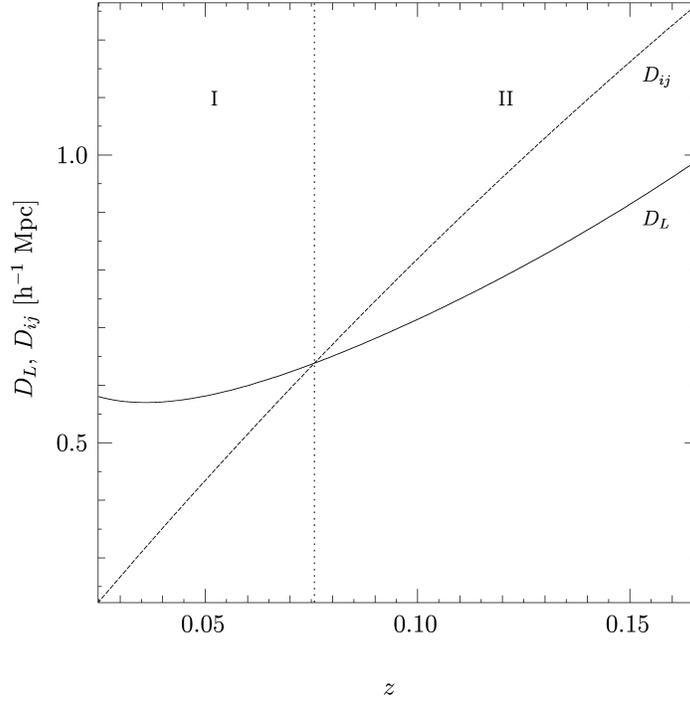


Figure 4.6: Comparison of the LCRS projected linking distance D_L (solid line) and the projected separation D_{ij} (dashed line) of the galaxies i and j , as a function of the redshift z . The angular separation θ_{ij} of the chosen galaxy pair leads to one intersection between D_L and D_{ij} , dividing the graph into two areas (dotted line). The projected linking criterion $D_{ij} \leq D_L$ is satisfied in area I, yet not in area II. This example galaxy pair is lying in a field, with a MOS setup of 112 fibers, a sampling fraction of $F = 0.7$, and photometric limits of $m_{min} = 15.0$ mag, $m_{max} = 17.7$ mag.

separation, leading to one intersection between D_{ij} (dashed line) and D_L (solid line). The intersection divides the graph into two areas (dotted line). In area I the projected linking criterion $D_{ij} \leq D_L$ is satisfied, in area II it is not. Thus, whenever the original FOF is limited to comparing D_{ij} with D_L at a value of D_{ave} , that corresponds to a redshift in area II, and the redshift-space linking criteria, eqs. (3.8) and (3.9), allow the EXT-FOF to compare D_{ij} with D_L in area I, only the EXT-FOF can find a link between the galaxies i and j .

The explanation, why the EXT-FOF cannot link together certain objects in the LCRS catalog, while the original FOF can, is completely analogous to the case of the CFA1: The original FOF can move relatively freely through redshift space in order to link galaxies to one another. The EXT-FOF, on the other hand, is always limited to fulfilling both linking criteria for at least N_{min} objects within one redshift slice.

4.3 LCRS with simulated photometric redshifts

In order to prove the validity of the EXT-FOF algorithm for structure finding in photometric redshift datasets, artificial photometric redshifts are created for all galaxies contained in the LCRS galaxy catalog. A group catalog is determined, based on this new dataset, with the help of the EXT-FOF algorithm, taking into account the photometric redshifts and their errors. A comparison between this structure catalog and the EXT-FOF catalog described in Section 4.2.2 shows the applicability of the EXT-FOF technique in case of galaxy catalogs with photometric redshifts.

4.3.1 Creation of the galaxy dataset with simulated photometric redshifts

For every LCRS galaxy a random redshift offset is created. The offsets follow a Gaussian distribution with a σ of roughly 5% of the survey depth. That percentage resembles the typical proportions between the redshift errors and the depth of a photometric redshift survey. In the case of the LCRS, the width of the Gaussian is set to $c\sigma = 2500 \text{ km s}^{-1}$. The photometric redshifts are then created by adding the individual offsets to the spectroscopic redshifts of the galaxies. Furthermore, the redshift errors δz_i are all set to σ . Except for the redshifts of the galaxies and their errors, the galaxy dataset remains unchanged.

4.3.2 Creation of the group catalog

To construct the EXT-FOF structure catalog of this pseudo-photometric redshift dataset, the recipe described in Section 4.2.2 is followed.

The new galaxy dataset is given the same treatment as the spectroscopic one, as far as the velocity correction, and the culling of galaxies on the basis of the velocity and absolute magnitude limits is concerned. Due to this selection process, the subset of galaxies that enters the structure finding procedure can be deviant from the one used in the spectroscopic redshift case. A total of 22739 galaxies are used for cluster finding in the spectroscopic redshift case, 22617 galaxies are used in the photometric redshift case. 20255 galaxies are common to both culled datasets.

For the calculation of the projected distance D_{ij} , $D(z_{ini})$ of eq. (3.7) is used. The relatively large redshift errors are taken into account, and relations (3.10) and (3.11) apply for the linking velocities $V_{L,i}$ and $V_{L,j}$. The stepsize Δz_{ini} is again set to 10^{-5} .

The same low matter density cosmology is used as described in Section 4.2.2. Furthermore, the same linking parameters and group mean velocity limits are applied. The minimum number of objects is once more set to $N_{min} = 3$.

	Cat.	EXT-FOF w/o δz	EXT-FOF w. δz	% EXT-FOF w/o δz	% EXT-FOF w. δz
Total		1285	1598	100	100
Identical	1	230	230	17.9	14.4
Only EXT-FOF w/o δz	2	94	–	7.3	–
Only EXT-FOF w. δz	3	–	646	–	40.4
EXT-FOF w/o δz struct. larger than EXT-FOF w. δz struct.	4	34	34	2.7	2.1
EXT-FOF w. δz struct. larger than EXT-FOF w/o δz struct.	5	382	382	29.7	23.9
EXT-FOF w/o δz struct. is comb. of EXT-FOF w. δz structs.	6	0	0	0.0	0.0
EXT-FOF w. δz struct. is comb. of EXT-FOF w/o δz structs.	7	229	92	17.8	5.8
EXT-FOF w. and w/o δz structs. have some elements in common	8	316	214	24.6	13.4
EXT-FOF w/o δz structs. found with EXT-FOF w. δz alg.		1191	–	92.7	–

Table 4.3: Comparison of the LCRS EXT-FOF structure catalogs with and without simulated photometric redshifts. The first column provides a short description of the categories used. In the second column, the corresponding categories are listed for easy reference. The third and fourth column yield the number of EXT-FOF without δz and EXT-FOF with δz groups respectively, that fall under the specified category, while the fifth and sixth column show the percentage of these groups.

4.3.3 Comparison of the group catalogs

Section 4.2.3 shows that in case of the spectroscopic dataset the EXT-FOF algorithm yields a group catalog that is very similar to the FOF group catalog. Thus, in principle, the groups resulting from the photometric redshift dataset could be compared to either the FOF, or the EXT-FOF group catalog of the spectroscopic dataset. For the following examination the latter is chosen.

The EXT-FOF algorithm finds a total of 1598 structures in the simulated photometric redshift galaxy catalog. The groups and clusters contain 10831 galaxies. This corresponds to a mean of almost seven objects per structure. In the spectroscopic redshift case, structures contain a mean of almost five objects. Obviously, the EXT-FOF algorithm tends to find more and larger structures in photometric redshift datasets.

The comparison between the EXT-FOF structures resulting from the spectroscopic redshift dataset and the structures resulting from the photometric one is shown in Ta-

ble 4.3. Again, the “category 1 - 8” classification scheme is used, that was already described in Sect. 4.1.3, replacing “FOF” with “EXT-FOF with spectroscopic redshift dataset” and “EXT-FOF” with “EXT-FOF with photometric redshift dataset”.

Only 94 of the 1285 EXT-FOF structures with spectroscopic redshifts cannot be retrieved (category 2), leading to a recovery rate of almost 93%. 646 additional groups are found in the photometric galaxy dataset (category 3), corresponding to a spurious detection rate of 40%. The number of structures, that are identical in both group catalogs is roughly 16%. Due to the strongly differing input galaxy catalogs, this small percentage does not come as a surprise. Both the category 2 and 3 structures tend to be relatively small. 96.8% of the category 2 and 81.6% of the category 3 groups have less than five members.

Under the assumption that the structures found by the FOF and EXT-FOF algorithms in the spectroscopic galaxy dataset are real, this comparison shows that the EXT-FOF algorithm is capable of finding almost all of the structures contained in this photometric redshift dataset. Roughly 60% of the found structures can be expected to be real. This demonstrates that the EXT-FOF technique is a very conservative cluster finder and can be used for identification of cluster candidates in photometric redshift galaxy datasets.

In principle, an optimization of the recovery rate and the rate of spurious detections is always a compromise. Very high recovery rates usually result in a large number of spurious detections. On the other hand, small spurious detection rates imply fewer recovered clusters. The EXT-FOF technique was created to maximize the recovery rate, meaning every single real structure in the dataset should be retrieved. Relatively high spurious detection rates have to be accepted. This results in what is here called a conservative catalog. Such an approach to structure finding is feasible, if further tests are undertaken in order to separate presumably real clusters from spurious ones. A number of those authenticity tests is shown in Chapter 5.

4.3.4 Analysis of the discrepancies

Some of the discrepancies in the two group catalogs are due to the slightly different composition of the culled input galaxy datasets. After all, the composition of those datasets differs by roughly 10%. However, since this is only a border effect, no further attention is paid to it here.

The really interesting discrepancies can be ascribed to either one, or both, of the following two reasons:

- One is the scattering of the newly created photometric redshifts against the real, spectroscopic ones.
- The other is the fact that the redshift errors are finally taken into account when testing the velocity linking criteria, using relations (3.10) and (3.11).

Due to the Gaussian nature of the simulated photometric redshift distribution, almost 32% of the galaxies have new redshifts, that deviate by more than 1σ from their original spectroscopic ones. However, eqs. (3.10) and (3.11) only allow for a deviation from z_{ini} of roughly 1σ ($V_L/2$ is relatively small compared to $c\delta z_i = c\delta z_j = c\sigma$). So there is a non-negligible amount of galaxies in the photometric redshift dataset, that are found as linked in the spectroscopic redshift cluster catalog, yet cannot be linked together anymore, because of their large scatter in photometric redshift. On the other hand, this scatter can also link galaxies together, that are defined as separated in the spectroscopic redshift dataset. So the scattering is responsible for both, finding and losing links between galaxies, and can result in any type of deviation ranging from category 2 to 8.

Taking the redshift errors into account leads to much larger velocity linking criteria $V_{L,i}$ and $V_{L,j}$, than in the spectroscopic redshift case where $V_L/2$ is used as linking criterion. Thus, the new velocity linking criteria are by far less strict, and the probability of finding two objects as linked is increased. The new velocity linking criteria are responsible for finding additional links and result in deviations of category 3, 5, and 7. In combination with the above mentioned scattering of photometric redshifts they can also result in category 8 discrepancies. The consideration of the photometric redshift errors is the main reason for the obviously larger and more numerous structures found in the photometric redshift structure catalog.

Chapter 5

Authenticity of Structures

Testing the authenticity of the clusters in a structure catalog is crucial for the optimization of the linking parameters. Furthermore, it is very useful for the classification of individual clusters as presumably real or spurious. Since the EXT-FOF algorithm is likely to yield a non-negligible fraction of spurious clusters, as was discussed in Sect. 4.3, having a set of tools for authenticity testing is very convenient.

Four techniques that can be used to assess the quality of the resulting structure catalog and also of the individual structures are explained in this chapter. In Sect. 5.1 attaining spectroscopic redshifts for some of the galaxies and using this information to determine cluster authenticity is discussed. Sect. 5.2 deals with the creation of Voronoi tessellations. This method yields a scale-free measure for local galaxy densities and thus shows clustering tendencies in the galaxy dataset. In Sect. 5.3 a likelihood approach is introduced, that determines probable cluster positions in the galaxy distribution. Finally, in Sect. 5.4, color-magnitude diagrams of the cluster members are explained. These diagrams highlight the elliptical or lenticular population, that make up about (50 – 80) % of the galaxies in clusters (see Sect. 1.1).

5.1 Spectroscopic redshifts

The most reliable method to confirm the cluster or group membership of a set of distant galaxies is to determine their spectroscopic redshifts. Unfortunately, this is also the most expensive technique, since it requires a lot of telescope time.

As was explained in Sect. 3.3.1, a spectroscopic redshift is the most precise measurement of the distance to a galaxy that can be gained in an observed non-local¹

¹There are a number of rather precise distance measurements for local galaxies, involving for example Cepheid light curves, the brightest globular cluster in a galaxy, or supernovae of type Ia (SNIa), to name just a few (Binney & Merrifield, 1998). However, they can either not be used for distant galaxies, or in the case of SNIa, not every galaxy in a given catalog has SNIa observations.

galaxy dataset. Sect. 3.3.1 also showed, that the deviation of the spectroscopic redshifts, or more precisely of the corresponding velocities, of two galaxies that belong to the same cluster are comparable in size to the velocity dispersion of that cluster.

Thus, in order to verify an EXT-FOF cluster candidate, the following approach is used in this work: A cluster candidate is defined as confirmed, if it contains at least N_v galaxies having spectroscopic redshifts $z_{spec,i}$ such that

$$\left| z_{spec,i} - \bar{z}_{cluster} \right| \leq a \cdot \sigma, \quad (5.1)$$

where $\bar{z}_{cluster}$ is the spectroscopic mean redshift of the cluster, a is an arbitrary factor, and σ is a typical structure velocity dispersion.

Since the photometric redshifts are too uncertain to be used for the determination of the real mean cluster redshift, eq. (5.1) is tested for a discrete set of possible $\bar{z}_{cluster}$. The exact choice of the variables N_v , a , and σ has to be optimized with respect to the given cluster dataset. However, a rough estimate of the acceptable range of values for these three parameters, or their dependency on certain cluster properties can be given independently of the cluster dataset:

- Typical velocity dispersions range from $\sigma \approx (100 - 1400) \text{ km s}^{-1}$ from groups to rich clusters, with median values of $\sigma \sim 250 \text{ km s}^{-1}$ and $\sigma \sim 750 \text{ km s}^{-1}$, respectively (Struble & Rood 1991, Bahcall in Dekel & Ostriker 1999, and Sect. 1.1).
- Under the assumption that the redshifts of cluster members can be approximated by a Gaussian distribution, which is at least true for the regular clusters, that make up about 30% of all structures, a total of $\sim 68\%$, $\sim 95\%$, and more than 99% of the cluster galaxies are contained within a $\pm 1\sigma$, $\pm 2\sigma$, and $\pm 3\sigma$ interval around the mean cluster redshift. Consequently, the factor a should be chosen such that $a \in [2, 4]$ so not to exclude too many possible cluster members. For example, Mellier et al. (1988) analyzed the velocity dispersion of the Abell cluster A370 and Quintana & de Souza (1993) examined the dispersion of A3571. The respective velocity dispersions are 1350 km s^{-1} and 1022 km s^{-1} , and their samples of cluster members with spectroscopic redshifts contain objects that have more than 2σ deviations from the mean cluster redshifts. The cluster CL 1358+6245 is discussed in Fisher et al. (1998). It has a velocity dispersion of 1027 km s^{-1} , and the spectroscopic sample of cluster members contains galaxies with deviations of roughly 4σ .
- The number N_v of galaxies that have to fulfill eq. (5.1) should depend on the expected richness of the structure. Since a group should contain less galaxies than a cluster, it is obvious that there will also be less spectra for groups, at least in case that spectra were taken independently of the proposed cluster membership of a galaxy.

In case of the MUNICS survey, the available spectroscopic redshift dataset is a K' band magnitude-limited sample of randomly distributed objects (Feulner et al., 2003), it was not designed to concentrate on cluster candidates. Thus it comprises mainly local to medium-distant galaxies with redshifts between 0.1 and 0.5, as is illustrated in Fig. 1.2, while the range for finding clusters in MUNICS extends to a redshift of unity. Moreover, Table 1.5 shows that only about 11% of the MUNICS objects possess spectroscopic redshifts. It is obvious that this dataset is not perfectly suited for the task of spectroscopic confirmation of the MUNICS clusters, yet it will be used anyway.

The parameters N_v , a , and σ , utilized for the verification of the MUNICS clusters, are chosen as follows. If only few galaxies can be found with similar spectroscopic redshifts, the examined structure might be a group, so a more stringent constraint should be chosen for the velocity dispersion. If however more galaxies are found with similar spectroscopic redshifts, the given structure might be a cluster and a larger velocity dispersion can be used. Hence, two sets of parameters are used for the verification of a structure. In principle, a structure at a higher redshift will inevitably have less visible – and thus less detectable – structure members. Thus a structure having less EXT-FOF members is not necessarily only a group. The above described use of the two confirmation criteria is actually implemented to be on the safe side, i.e. not to be too generous in the confirmation of a structure.

In the following, an EXT-FOF MUNICS structure is said to be confirmed, if

- there are $N_v = 2$ galaxies, that deviate from their spectroscopic mean cluster redshift by less than $a \cdot \sigma = 3 \cdot 250 \text{ km s}^{-1} = 750 \text{ km s}^{-1}$, or
- there exist at least $N_v = 3$ galaxies, with deviations from their spectroscopic mean cluster redshift of less than $a \cdot \sigma = 3 \cdot 750 \text{ km s}^{-1} = 2250 \text{ km s}^{-1}$.

5.2 Voronoi tessellation

The use of Voronoi tessellations to solve astrophysical problems has become very popular over the past two decades. The astrophysical application of these tessellations started with the pioneering work of Kiang (1966), who determined the size-distribution of random Voronoi polygons and polyhedra in two and three dimensions, respectively. Subsequently, the tessellations were used for example, by Icke & van de Weygaert (1987), van de Weygaert & Icke (1989), and van de Weygaert (1994) in their “Fragmenting the Universe” series of papers, dealing with the analysis of superclustering in the universe. Furthermore, a slightly different approach to implement Voronoi regions was made by Zaninetti (1989, 1990, 1991, 1993, 1995), who utilized dynamical Voronoi tessellation to analyze supernova explosions and their eventual interaction in the galactic plane, as well as the distribution of galaxies and asteroids. Several other applications of the Voronoi technique to astronomical issues have been published, like Matsuda & Shima (1984), Ling (1987), Yoshioka & Ikeuchi (1989), Coles (1991), Ikeuchi & Turner (1991), Ebeling & Wiedenmann (1993), Goldwirth et al. (1995),

Ryden (1995), El-Ad et al. (1996), El-Ad & Piran (1997), Doroshkevich et al. (1997), and Meurs & Wilkinson (1999), among others. Voronoi tessellation has also been applied to finding groups and clusters of galaxies in both, two-dimensional (Ramella et al., 2001; Kim et al., 2002), as well as three-dimensional surveys (Marinoni et al., 2002), and proved to be an efficient and reliable tool.

5.2.1 The definition of Voronoi tessellation

The first mention of what is now called Voronoi regions was made by Dirichlet (1850), and Voronoi (1908). An elaborate overview of the basics of Voronoi tessellation is given by van de Weygaert (1994).

In principle, a Voronoi tessellation is defined as follows (after van de Weygaert 1994): Let Θ be a countable set of nuclei $\{\underline{x}_i\}$ in the n -dimensional space \mathbb{R}^n , and $\underline{x}_1, \underline{x}_2, \underline{x}_3, \dots$ be their coordinates. For all nuclei i the Voronoi cell Π_i is defined by the following set of points \underline{x} of the space:

$$\Pi_i := \{\underline{x} \mid d(\underline{x}, \underline{x}_i) < d(\underline{x}, \underline{x}_j) \quad \forall j \neq i\}, \quad (5.2)$$

with $d(\underline{x}, \underline{y})$ being the Euclidean distance between the points \underline{x} and \underline{y} . The complete set of $\{\Pi_i\}$ is a tessellation of \mathbb{R}^n , the Voronoi tessellation $\mathcal{V}(\Theta)$ relative to the set Θ of nuclei.

In other words, the Voronoi tessellation of a two- (or three-) dimensional set of nuclei is a plane- (or space-) partition into convex cells, i.e. polygons (or polyhedra). Each of those cells contains exactly one nucleus and the set of points that is closer to the given nucleus than to any other. The convexity stems from the intersection of the open half-spaces defined by the perpendicular bisectors of the segments connecting the given nucleus with all the other nuclei. This partition is uniquely defined by the set of nuclei.

Two nuclei i and j in a two- (or three-) dimensional set Θ of nuclei, whose Voronoi regions Π_i and Π_j have one line (or face) in common are called contiguous. A network can be obtained by joining all the contiguous pairs of nuclei. A set of three (or four) nuclei, that are contiguous to each other, form a triangle (or tetrahedron). The complete set of all those triangles (or tetrahedra) define a new tessellation, the so-called Delaunay tessellation $\mathcal{D}(\Theta)$ (Delaunay, 1934). This Delaunay tessellation is the dual of the Voronoi tessellation. Like the Voronoi tessellation, the Delaunay partition is in principle defined in the n -dimensional space \mathbb{R}^n .

In three-dimensional space, Voronoi tessellations are made up of three topologically different elements (Martínez & Saar, 2002): Walls, that are the bisecting planes between two nuclei; lines, that are the intersection of three walls; and vertices, where four lines intersect. In two dimensions, the tessellations contain only lines, the perpendicular bisectors between two nuclei, and vertices, at the intersections of three lines.

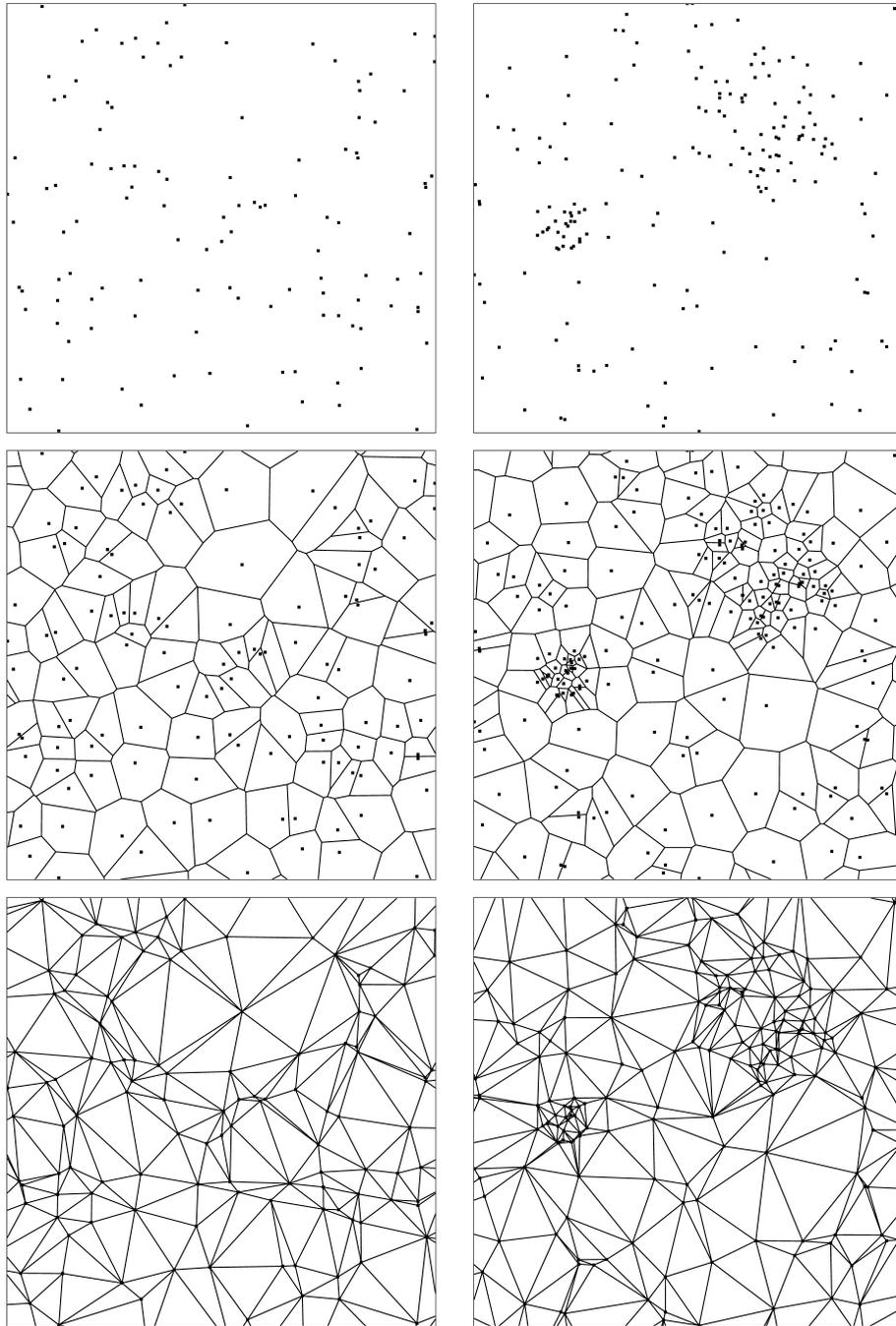


Figure 5.1: Point patterns, Voronoi tessellations and Delaunay meshes for two different two-dimensional point distributions. The **left row** depicts a random distribution of points, while the **right** shows two concentrations of points on a random background. The **upper panel** illustrates the positions of the nuclei, the **middle panel** shows the resulting Voronoi tessellation, and the **lower panel** depicts the corresponding Delaunay mesh.

Fig. 5.1 illustrates the properties of Voronoi and Delaunay tessellations. It depicts the nuclei distributions, Voronoi tessellations and Delaunay meshes for two different types of point distributions. The left row shows a random distribution of nuclei, while on the right side two concentrations of nuclei are plotted on a random background. For simplification this exercise is done in two dimensions.

It becomes obvious that strongly clustered nuclei in a distribution have very small Voronoi cells. Thus, the area A_i of the Voronoi cell around nucleus i is a measure for the local density n_i of the nucleus distribution, with

$$n_i := 1/A_i. \quad (5.3)$$

This measure is local, since the boundaries of the cell are solely defined by the neighboring nuclei and the nucleus i itself. A density contrast δ_i relative to the mean can be defined by

$$\delta_i := n_i / \langle n \rangle, \quad (5.4)$$

where

$$\langle n \rangle := \frac{1}{N} \sum_{i=1}^N n_i \quad (5.5)$$

is the mean density and N is the total number of nuclei.

5.2.2 Application to structure finding

In order to apply the technique of Voronoi tessellation to structure finding, or as is the case here to the authentication of groups and clusters, the galaxies of the survey are identified with the nuclei of the tessellation. Thus the Voronoi area belonging to each nucleus yields a scale-free measure of the local galaxy density.

So far, three structure finding algorithms, that are based on the Voronoi method have been created:

- [Ramella et al. \(2001\)](#) devised the Voronoi galaxy cluster finder (hereafter VGCF) to look for structures in two-dimensional imaging surveys, like the Palomar Distant Cluster Survey (PDCS; [Postman et al. 1996](#)). Their technique uses the projected galaxy positions and the apparent magnitudes. By binning the dataset according to the magnitudes, the authors make a rough deprojection of the galaxy catalog. The Voronoi tessellation is then run on the binned dataset and adjacent Voronoi cells with a density higher than a selected threshold are singled out as cluster candidate centers. The shape of the cluster candidate is then regularized by fitting a circle to each candidate and expanding the radius until the mean

galaxy density inside the circle is lower than the density in the original central concentration. Finally, these candidate catalogs per magnitude bin are reduced into a single cluster catalog by accepting only those candidates that are detected in a minimum number of magnitude bins.

- [Kim et al. \(2002\)](#) created the Voronoi tessellation technique (hereafter VTT), a method designed to detect structures in the Sloan Digital Sky Survey (SDSS; [York et al. 2000](#)). Since the SDSS is a multi-band photometric survey, a binning in color and magnitude is used to deproject the dataset. Except for the pre-selection, the recipe for the VTT is rather similar to that of the VGCF and the details will not be discussed here.
- [Marinoni et al. \(2002\)](#) developed the Voronoi-Delaunay method (hereafter VDM) designed for structure finding in spectroscopic surveys, like the DEEP2 survey ([Davis et al., 2001](#)). This technique combines the three-dimensional density information of the Voronoi diagram with the spatial proximity information of the Delaunay mesh. In the first step of the algorithm the overdense regions of the Voronoi tessellation are marked as candidate locations for clusters. The second step uses the vectors of the Delaunay mesh to identify the central cluster members and to estimate physical properties, like the core density. In the final step these estimates are used to determine the redshift-space window that defines the group members.

In the following, an overview is given of the most important advantages and disadvantages of Voronoi-based cluster finding.

The probably most important advantage of the Voronoi approach is the absolutely nonparametric way of determining local densities. Usually densities are determined by isotropically smoothing the dataset with a scale length given by model predictions or empirical data. As a result, a structure finding method based on these galaxy densities may be selective in favor of spherical clusters with the given scale length. However, a large fraction of clusters is known to have a distinct ellipticity ([Struble & Ftaclas, 1994](#); [Plionis et al., 1991](#); [Wang & Ulmer, 1997](#); [Basilakos et al., 2000](#)). Thus, a cluster finder based on Voronoi tessellations to derive local galaxy densities will be less biased against asymmetric structures with dimensions that deviate from model expectations. Therefore, the Voronoi approach is very close to the dictum “let the data speak for themselves” ([Openshaw, 1984](#)). Moreover, Voronoi-based cluster finders need only one run of the more time-consuming part of their algorithm, which is the tessellation itself. The key parameters, like in the case of the VGCF the density threshold and the number of bins where a cluster candidate needs to be detected, can be chosen afterwards. Consequently the optimization of the key parameters and thus the entire procedure of cluster finding can be done at great speed. Another advantage of this type of structure finding, especially when compared to probabilistic approaches like the matched filter technique, is the fact that this approach yields a list of cluster members that can be used immediately for spectroscopic verification.

Naturally the Voronoi technique also exhibits some disadvantages. Like most cluster finding methods, the Voronoi approach suffers from edge effects. The galaxies that lie next to the geometrical borders of the survey cannot be assigned a closed Voronoi cell. As a result, no area can be determined for those objects and they will have to be removed from the list of possible cluster members. Obviously it is preferable to apply Voronoi based cluster finders to continuous and very dense datasets. Unfortunately, MUNICS is not one of those. Besides, it should be noted that the fringe galaxies of a cluster have larger area cells, since at least one of their contiguous galaxies belongs to the field. Thus the final cluster member identification has to include galaxies with lesser densities, as well.

The above mentioned Voronoi cluster finding algorithms of [Ramella et al. \(2001\)](#) and [Kim et al. \(2002\)](#) are not directly used for the authenticity testing, although a modification of the codes to suit a photometric redshift dataset would be very easy. Both structure finding algorithms contain arbitrary cut-off criteria, like in the case of the VGCF the density threshold or the number of bins in which a cluster candidate has to be detected, and the resulting structure catalogs are influenced by that choice. Consequently, a simple visual inspection of the Voronoi density distribution and of the arrangement of the EXT-FOF cluster members is much closer to the physical reality and comprehends fewer assumptions, than a comparison of the two structure catalogs.

The Voronoi-based authenticity testing that is done for MUNICS follows roughly the idea of [Ramella et al. \(2001\)](#) and [Kim et al. \(2002\)](#). An overlapping photometric redshift binning is used to create two-dimensional subsets of the MUNICS data. Voronoi tessellation is applied to these subsets and the mean density of a random distribution with the same number of galaxies is determined by Monte Carlo calculations. Finally, Voronoi diagrams are created and overdense cells are highlighted. This projected density distribution is then compared to the configuration of EXT-FOF cluster members in the given redshift bin. The galaxies belonging to EXT-FOF structures are expected to occupy preferably the overdense cells. If this is not the case, and a significant fraction of the EXT-FOF cluster members can be found in the low density regions, the linking parameters of the EXT-FOF algorithm are too liberal and have to be reduced. Thus, the Voronoi diagrams help set a rough upper limit to the range of linking parameters and are also capable of accentuating unlikely structures.

To shed some light on the actual implementation of the Voronoi-based authenticity testing for MUNICS, a more detailed description of the individual steps is given as follows:

The dataset used for the creation of the Voronoi tessellation consists of all objects, that are defined as galaxies by their extended profile, as well as their best-fitting SED template. No culling with respect to the size of photometric redshift errors is made.

The central redshifts $z_{central}$ of the bins range from 0 to 1.5 with a spacing of 0.02. The nominal half-width Δz_{bin} of the bins is chosen of the order of the photometric redshift errors $\Delta z_{bin} = 0.055(1 + z_{central})$. To furthermore take into account objects with somewhat deviant redshifts, that also have larger redshift errors, and thus could still be compatible with the given bin, every galaxy that is compatible with the nominal

bin boundaries within its 1σ error is included in the binned dataset. Obviously, the bins overlap well even at low redshifts.

In the next step, the Voronoi tessellation is run on this two-dimensional subset, using the “triangle” C-code of [Shewchuk \(1996\)](#), following [Ramella et al. \(2001\)](#). The galaxies that are lying next to the geometric borders of the survey and therefore have no closed polygons are removed, and the areas and densities are determined for the remaining objects.

Monte Carlo simulations are made to determine the mean density. Thereto 50 random distributions Θ_k with $k \in [1, 50]$ of the galaxies in the given bin are created. The Voronoi tessellations $\mathcal{V}(\Theta_k)$ are determined and the edge nuclei are culled as described above. The mean densities $\langle n \rangle_k$ are determined according to eq. (5.5), with N being the number of nuclei of Θ_k with closed polygons. The final mean density is then given by

$$\langle n \rangle := \frac{1}{50} \sum_{k=1}^{50} \langle n \rangle_k. \quad (5.6)$$

It should be noted, that the so generated mean density is not identical to the mean density of a sample of field galaxies, since the input dataset can include field, as well as cluster galaxies. Thus, the number of galaxies that enter into the Monte Carlo run can be too high and the corresponding mean density can also be increased with respect to a normal underground field density. However, any accumulation of cells that are overdense with respect to this mean density, i.e. $\delta_i > 1$ are also overdense with respect to a typical underground field density. Furthermore, at least the central regions of rich structures will be overdense with respect to this mean density and the EXT-FOF algorithm is anyway more likely to find the central populations of structures than all of their outliers. Thus, galaxies that belong to authentic EXT-FOF clusters are generally expected to lie within the overdense cells. Consequently, an accumulation of galaxies having $\delta_i > 1$ could be an indicator for a cluster center.

To visually inspect the consistency of EXT-FOF structure members with overdense Voronoi polygons, plots are made depicting the Voronoi tessellation for the given redshift bin. Nuclei in overdense cells with $\delta_i > 1$ are marked, as well as the members of EXT-FOF structures. Examples of these plots are shown in Sect. 6.3.2.

5.3 Likelihood approach

The likelihood approach is used to compare observational data with theoretical or empirical models. In principle, this method yields a measure for the plausibility that a dataset can be described by a given model and consequently the best-fitting model parameters can be extracted. The method is based on Bayesian probability theory ([Bayes, 1763](#)), an alternative view of statistics.

Obviously, this likelihood estimation can be put to many uses in the physical sciences. In extragalactic astronomy, for example, likelihoods have been used to determine photometric redshifts, as was already mentioned in Sect. 1.2. Furthermore, the likelihood method has proven to be very useful for the detection of galaxy clusters. This is especially true in the case of structure finding in the two-dimensional imaging surveys, although likelihood based techniques can be applied to every level of precision in the distance information. The first likelihood based cluster finder, the so-called matched filter algorithm, was created by Postman et al. (1996) and was used by many authors, like Olsen et al. (1999) and Donahue et al. (2002). Since then, it has been refined (Schuecker & Boehringer, 1998; Kawasaki et al., 1998; Postman et al., 2002) and modified (Kepner et al., 1999; Lobo et al., 2000; Kim et al., 2002), and has become one of the most popular structure finding techniques.

5.3.1 Bayesian statistics and the likelihood

A thorough overview of Bayesian probability theory and some of its astrophysical implementations is given by Loredó (1995). The following description of the Bayesian approach is based on Lupton (1993) and Bevington & Robinson (2003).

Let D and M be two events. The probability $P(D \cap M)$ of both events happening is

$$P(D \cap M) = P(D|M)P(M) = P(M|D)P(D). \quad (5.7)$$

$P(D)$ or $P(M)$ are the probabilities for the occurrence of the events D or M , respectively. $P(D|M)$ is the probability of D under the condition that event M has occurred.

Eq. (5.7) can be transformed into Bayes' theorem

$$P(M|D) = P(D|M) \frac{P(M)}{P(D)}, \quad (5.8)$$

where $P(M)$ is called a prior probability, $P(M|D)$ is a posterior probability and $P(D|M)$ is referred to as the likelihood.

Bayes' postulate states that in the absence of all other knowledge, the prior probabilities can be assumed to be equal, i.e.

$$P(M) = \text{const.} \quad (5.9)$$

Let for example D represent observed data and M be a set of model parameters used to describe the observations. Utilizing Bayes' theorem and postulate, it becomes possible to estimate the best-fitting model parameters. In order to maximize the probability $P(M|D)$ that the model suits a given dataset, only the total likelihood $P(D|M)$ of the model has to be maximized, since $P(D)$ is a constant, as well.

5.3.2 Application to structure finding

As of today at least three subtypes of the matched filter technique have been developed. The following list contains their basic working principles and major differences:

- [Postman et al. \(1996\)](#) devised the matched filter algorithm (hereafter MF) to search for clusters in the PDCS (see Sect. 5.2.2), an imaging survey in two broadband filters. The algorithm makes use of the projected positions of the galaxies and their magnitudes. A grid is created in the projected coordinates x and y for a set of z -slices. The likelihood is then determined for every gridpoint (x, y, z) to be the center of a galaxy cluster, by comparing the observed data with model predictions. This model contains a term for the background distribution of galaxies, as well as an expression for the distribution of cluster members, known as the filter. The MF likelihood function comprises two filters: A luminosity and a radial distribution, that are typical for clusters. The luminosity distribution is given by a Schechter function, while the radial filter is defined by a King-like profile. A Gaussian approximation is assumed for the likelihood. Clusters are identified by locating peaks in the resulting likelihood map. This is done by using a given set of likelihood thresholds, as well as a minimal number of redshift-slices, in which a cluster candidate has to be detected. In the end, the MF algorithm yields a catalog of clusters with projected coordinates, estimated redshifts and richnesses.
- [Schuecker & Boehringer \(1998\)](#) designed a generalized likelihood filter. This is an improved version of the MF algorithm, that was created to detect the optical counterparts of the ROSAT-ESO-Flux-Limited X-Ray (REFLEX; [Böhringer et al. 2001](#)) clusters. It uses the position and magnitude informations of the galaxies, and adopts the same radial and luminosity profiles, as the MF. In contrast to the MF, the generalized likelihood filter utilizes a mathematically exact expression for the likelihood function, a Poisson distribution. It was shown by [Schuecker & Boehringer \(1998\)](#) that the formulae of the original MF approach can be regained if the background galaxies dominate the total number count. In that case the Poisson distribution can be approximated by a Gaussian, as claimed by the central limit theorem.
- The adaptive matched filter (hereafter AMF) was created by [Kepner et al. \(1999\)](#) to locate structures in the SDSS (see Sect. 5.2.2). Like the other two techniques, it uses the positions and magnitudes of the input objects, yet it can also include redshift informations with arbitrary precision, if available. Consequently, the AMF requires two to three filters. Once more, a King-like radial and a Schechter luminosity profile are chosen. The third filter is a simple redshift cut, which removes all objects that are not compatible within their errors with the given redshift slice. The AMF uses the locations of the input galaxies to create an adaptive grid that provides sufficient spatial resolution even at high redshifts. The actual creation of the structure catalog is done in two steps. First, a coarse likelihood filter is used to locate the clusters. This is basically an application of

the original MF algorithm, that is linear and very quick to compute. In the second step, the cluster redshift and richness are estimated, using a fine likelihood filter. This fine filter assumes a Poisson distribution in the likelihood, as in the case of the generalized likelihood filter.

The matched filter approach excels in finding weak cluster signals in a noise dominated background distribution. In principle, it can be adapted to include any type of model assumption. One of the algorithm's biggest advantages is the fact that it yields an estimate of the cluster redshift, even if the input dataset contains no redshift information. Furthermore, a measure for the richness of the structures is given.

Of course this approach exhibits drawbacks, as well. Like most structure finding methods, the matched filter technique experiences non-negligible border effects. The biggest problem of this approach results from the comparison of real data to a set of artificial model assumptions. By defining a fixed form for the filters, the algorithm is biased against any atypical structures. Since clusters can be found in a variety of shapes, with varying central concentrations, galactic content, etc., as was described in Sect. 1.1, this bias can pose a problem. Especially clusters with large ellipticities, or the presence of pronounced substructures can turn out to be a problem for the radial filter. In case of isolated spherical systems the algorithm can be expected to be very reliable, though. Another disadvantage of this approach is the fact that it does not yield a catalog of presumed cluster members that could be used for spectroscopic verification.

The matched filter incarnations described above are not directly used for the authenticity testing of the MUNICS clusters. Like in the case of the Voronoi-based techniques, these structure finding algorithms contain sets of cut-off criteria, that strongly influence the resulting cluster catalogs. Once more, a visual inspection of the likelihood distribution and of the arrangement of the EXT-FOF cluster members is carried out. This is expected to be closer to the physical reality, than a comparison of the resulting structure catalogs.

The likelihood based authenticity testing created for MUNICS rests on a combination of the original MF algorithm and the AMF. A Gaussian distribution is assumed for the likelihood. Although this is only a rough approximation and not true in case of low background galaxy distributions, it is still sufficiently accurate to locate the peaks in the likelihood. After [Kepner et al. \(1999\)](#), a set of three filters is used: A radial profile, a luminosity function and a redshift cut-off.

In the following, the recipe for the creation of the MUNICS likelihood maps is explained in detail:

Like in the Voronoi approach, only those MUNICS objects are used as part of the dataset that are defined as galaxies by both their morphology and their SED.

A three-dimensional grid is created. The redshift slices of the grid range from 0 to 1.5 with a stepsize of 0.02. Because of the singularity of the angular distance at a redshift of 0.00, the first redshift grid is shifted to 0.01. Finer steps along the redshift axis are useless, since the stepsize is chosen identical to the resolution of the

photometric redshift determination code. In every redshift slice, a grid of $101 \cdot 101$ points is made in the projected directions x and y , covering the observed expansion of the MUNICS fields. Tests have shown that this grid is sufficiently smooth at all given redshifts.

In the next step the likelihood is determined for every gridpoint to be the center of a galaxy cluster (after [Postman et al. 1996](#)):

The model $D_m(x, y, m)$ for the distribution of the galaxies is given by

$$D_m(x, y, m) := b(m) + \Lambda_i s(x, y, x_i, y_i) \phi(m, m^*(z_i)), \quad (5.10)$$

where $b(m)$ is the background distribution of galaxies, Λ_i is a measure for the richness of the structure i , and (x_i, y_i, z_i) are its central coordinates in redshift space. The term $s(x, y, x_i, y_i)$ describes the projected density of galaxies belonging to the structure at the position (x, y) , and $\phi(m, m^*(z_i))$ is the differential cluster luminosity function, with the apparent characteristic K' band magnitude $m^*(z_i)$ at the redshift z_i .

The observed galaxy distribution $D_d(x, y, m)$ can be described by a set of δ -functions

$$D_d(x, y, m) := \sum_{j=1}^N \delta(x_j, y_j, m_j), \quad (5.11)$$

where N is the total number of galaxies, and x_j, y_j , and m_j are the observed projected position and K' band magnitude of galaxy j .

Assuming a Gaussian distribution for the conditional probability that the observed dataset can be described by the given model, yields

$$P(x, y, m) = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[-\frac{[D_m(x, y, m) - D_d(x, y, m)]^2}{2\sigma^2} \right], \quad (5.12)$$

with

$$\sigma^2 = b(m). \quad (5.13)$$

Here $b(m)$ is approximated by the distribution of the data.

The likelihood L is defined by

$$L := \prod_{x, y, m} P(x, y, m). \quad (5.14)$$

Since the likelihood is a product of normalized probabilities, the value of L can become very small and a computer might not be able to deal with the results anymore. Consequently, it is common practice to maximize $\ln L$ instead of the likelihood itself.

After a series of transformations and the maximization with respect to Λ_j , the logarithmic likelihood can be approximated by

$$\ln L \propto \sum_j \frac{1}{b(m_j)} s(x_j, y_j, x_i, y_i) \phi(m_j, m^*(z_i)). \quad (5.15)$$

In other words, this expression describes the maximized log-likelihood for every grid point (x_i, y_i, z_i) to be the center of a galaxy cluster. For a detailed derivation of the maximized logarithmic likelihood see Appendix A.

A King-like profile is chosen as the radial filter, following Postman et al. (1996), Schuecker & Boehringer (1998), and Kepner et al. (1999):

$$s(x_j, y_j, x_i, y_i) := \begin{cases} \frac{1}{v} \left[\left[1 + \left(\frac{r_{ij}}{r_c} \right)^2 \right]^{-\frac{1}{2}} - \left[1 + \left(\frac{r_{co}}{r_c} \right)^2 \right]^{-\frac{1}{2}} \right] & \text{if } r_{ij} < r_{co} \\ 0 & \text{if } r_{ij} \geq r_{co} \end{cases} \quad (5.16)$$

The projected distance r_{ij} between gridpoint i and galaxy j is given by

$$r_{ij} := 2 \sin \frac{\theta_{ij}}{2} d_A(H_0, \Omega_M, \Omega_\Lambda, z_i), \quad (5.17)$$

where θ_{ij} is their angular separation and $d_A(H_0, \Omega_M, \Omega_\Lambda, z_i)$ is the angular distance to the redshift z_i of the grid slice. The core radius is set to $r_c = 0.125 h^{-1} \text{Mpc}$ and the cutoff radius is chosen as $r_{co} = 10 r_c$. This corresponds roughly to the values used by Postman et al. (1996), Schuecker & Boehringer (1998), and Kepner et al. (1999). v is the normalization of the profile

$$v := \int_0^{r_{co}} \left[\left[1 + \left(\frac{r}{r_c} \right)^2 \right]^{-\frac{1}{2}} - \left[1 + \left(\frac{r_{co}}{r_c} \right)^2 \right]^{-\frac{1}{2}} \right] 2\pi r dr, \quad (5.18)$$

so that

$$\iint_{-\infty}^{\infty} s(x, y, x_i, y_i) dx dy = 1. \quad (5.19)$$

A Schechter luminosity function $\phi(m, m^*)$ is chosen. Following Postman et al. (1996), a modified luminosity function $\Phi(m, m^*)$ is created, to place most of the weight of the filter on the bins around the characteristic magnitude m^* , the region of highest contrast between field and cluster galaxies

$$\Phi(m_j, m^*(z_i)) := \frac{1}{\mu} \phi(m_j, m^*(z_i)) 10^{-0.4(m_j - m^*(z_i))}. \quad (5.20)$$

μ is the normalization

$$\mu := \int_0^{m_{lim}} \phi(m, m^*(z_i)) 10^{-0.4(m-m^*(z_i))} dm, \quad (5.21)$$

so that

$$\int_0^{m_{lim}} \Phi(m, m^*(z_i)) dm = 1, \quad (5.22)$$

with m_{lim} as the limiting magnitude of the given MUNICS field.

The parameters of the Schechter function are based on [de Propris et al. \(1999\)](#), who determined K band luminosity functions for galaxies in 38 clusters of varying richness with redshifts in the range of $0.1 < z < 1$. The discrepancies between K band observations and the K' band observations of MUNICS are minor and can be easily neglected, as far as the localization of the likelihood peaks is concerned. The [de Propris et al. \(1999\)](#) infrared-selected galaxy samples reach generally 2 mag fainter than m^* , thus a fixed slope for the faint end of the luminosity function is assumed. The slope is set to $\alpha = -0.9$, that is the value measured in the infrared luminosity function of the Coma cluster for $m < m^* + 3$ mag ([de Propris et al., 1998](#)). [Fig. 5.2](#) illustrates the [de Propris et al. \(1999\)](#) measurements of m^* in ten observed redshift bins and the best-fitting logarithmic expression

$$m^*(z) = 14.01 + 6.45 \ln(1 + z). \quad (5.23)$$

This term is used for the approximation of m^* in the individual z_i -slices. The standard deviation of this fit is 0.789.

Following the idea of [Kepner et al. \(1999\)](#), a redshift filter is included in the determination of the logarithmic likelihood. In principle, this redshift filter removes all galaxies j that are not compatible within their redshift error $w \cdot \delta z_j$ with the given redshift slice z_i . In other words, only those galaxies are taken into account that fulfill

$$\left| z_j - z_i \right| \leq w \cdot \delta z_j. \quad (5.24)$$

The factor w is a weighting factor, that has to be selected considering the size of a typical δz_j and the depth of the survey. If δz_j is relatively small, w should be set to three. In case δz_j constitutes a considerable fraction of the survey depth, as is the case in MUNICS, it has to be set to one.

Including the modified luminosity function and the redshift filter, eq. (5.15) turns into

$$\ln L \propto \sum_{j \mid |z_j - z_i| \leq \delta z_j} \frac{1}{b(m_j)} s(x_j, y_j, x_i, y_i) \Phi(m_j, m^*(z_i)). \quad (5.25)$$

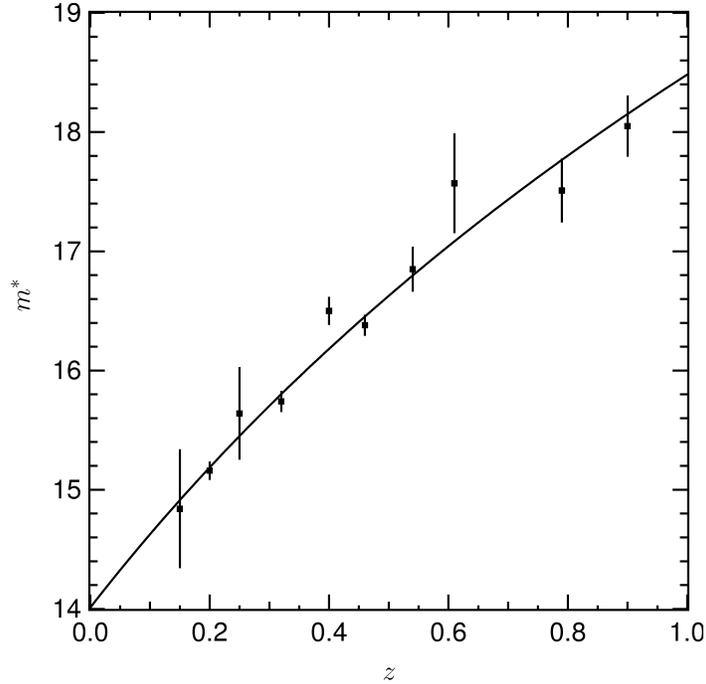


Figure 5.2: The variation of the characteristic apparent magnitude m^* of cluster K band luminosity functions with redshift. The squares represent the [de Propris et al. \(1999\)](#) measurement for m^* in ten redshift bins. Their respective magnitude-errors are given by the vertical lines. The solid curve shows the best-fitting logarithmic approximation to the data (see text).

For the visual inspection of the consistency of EXT-FOF structure members with the likelihood distribution, $\ln L$ maps are created for all redshift slices of the three-dimensional grid. All members of EXT-FOF structures, that roughly correspond to the examined redshift are marked for easy comparison. Authentic EXT-FOF structures can be expected to coincide with local maxima of the likelihood distribution. Examples of these plots are shown in Sect. [6.3.2](#).

5.4 Color-magnitude diagrams

Color-magnitude diagrams have been put to many uses in various fields of astronomy. For stellar astronomy, for example, they can provide valuable information about the evolutionary state of individual stars, and sometimes of their masses. They also yield estimates of the age and distance of (globular) clusters and thus can even influence the cosmological paradigm by setting a limit on the age of the universe. Furthermore, color-magnitude diagrams are very useful for making a rough deprojection of two-dimensional galaxy datasets. This feature was utilized by a number of authors

(Gladders & Yee, 2000; Yee & Gladders, 2002; Kim et al., 2002; Goto et al., 2002) to improve the quality of structure finding algorithms applied to multi-band imaging surveys that do not have photometric redshift estimates.

5.4.1 The cluster red sequence

The majority of early type galaxies, i.e. ellipticals and lenticulars, in clusters lie along a linear relation in a color-magnitude diagram (hereafter CMD). This relation is often referred to as the red sequence. The red sequence is a narrow and almost horizontal line, i.e. it consists of ellipticals with varying magnitudes, yet nearly constant colors (Visvanathan & Sandage, 1977; Annis et al., 1999). It shows a slight negative slope, which implies that fainter early type galaxies possess somewhat bluer colors than their more luminous counterparts (Baum, 1959; Visvanathan & Sandage, 1977; Lugger, 1984; Metcalfe et al., 1994). It was also shown that this slope evolves with redshift (Gladders et al., 1998), and furthermore varies with the richness of the clusters (Stoughton et al., 1998). The overall color of the red sequence is redder for more distant structures, and consequently can be used for an approximate redshift estimation of clusters (Gladders & Yee, 2000).

According to Aragon-Salamanca et al. (1993), Rakos & Schombert (1995), van Dokkum & Franx (1996), Kelson et al. (1997), Ellis et al. (1997), Bender et al. (1998), van Dokkum et al. (1998), Stanford et al. (1998), Kodama et al. (1998), Barrientos (1999), and Jørgensen et al. (1999) among others, the majority of the stellar population of the cluster ellipticals seems to have formed at high redshifts of $z > 2$, and the galaxies then appear to have followed a passive evolution leading to the stringent color and magnitude correlation, that can be observed for all cluster types and redshifts:

- Bower et al. (1992a,b) examined the CMDs of two local structures, the rich Coma cluster, as well as the comparatively poor Virgo cluster. They were able to show that both structures exhibit distinct red sequences with only small scatters.
- The low redshift cluster survey of Lopez-Cruz & Yee (2000) consists of 45 X-ray selected Abell clusters of various Abell richnesses, Bautz-Morgan and Rood-Sastry classes up to a redshift of $z < 0.2$. The authors demonstrated that every cluster has a red sequence and that the sequences seem to be very homogeneous from one cluster to the next.
- Smail et al. (1998) used ten optically selected X-ray luminous clusters with redshifts ranging from $0.22 < z < 0.28$, and Barrientos (1999) examined eight optically selected clusters in the redshift range $0.39 < z < 0.48$. In both studies the clusters have red sequences with very homogeneous scatters and colors.
- Oke et al. (1998) examined a sample of nine optically selected clusters located at redshifts $0.6 < z < 0.9$. Two of these structures were spectroscopically verified as clusters and had detailed information. Both objects show conspicuous red sequences (Oke et al., 1998; Stanford et al., 1998; Gladders et al., 1998).

- The existence and behavior of red sequences for clusters with redshifts above unity is not well studied yet, simply because of the low number of corresponding candidates. However red sequences were found for a number of high redshift structures. E.g. da Costa et al. (1999) showed red sequences in two clusters at redshifts of approximately unity. Benítez et al. (1999) proved the existence of a red sequence in a cluster at $z = 1.01$. Further examples of a strong red sequence were found in clusters at redshifts $z = 1.206$ (Dickinson, 1995), and $z = 1.273$ (Stanford et al., 1997).

5.4.2 Application to structure finding

As described in Sect. 1.1, galaxy clusters are dominated by early type galaxies, and their fraction of ellipticals can be up to 35%. Due to the correlation between the overall color of the red sequence and the redshift of the cluster, CMDs can be used to approximately deproject two-dimensional multi-color surveys.

This characteristic has been used for cluster finding by various authors:

- The cluster red sequence method (hereafter CRS; Gladders & Yee 2000, and Yee & Gladders 2002) was designed to find structures in the Red Sequence Cluster Survey (RCS; Gladders 2002), a two-band imaging survey in the filters R and z' . The CRS algorithm makes use of the projected positions, the $R - z'$ color, its error, and the z' apparent magnitude. Clusters are detected as overdensities in the color and magnitude space, as well as in projected positions: A series of overlapping color-magnitude slices removes fore- and background galaxies. Next, the algorithm searches for peaks in the weighted surface density of the individual subsets.
- The Voronoi tessellation technique of Kim et al. (2002) (see Sect. 5.2.2) also uses a binning in color-magnitude space to deproject the dataset. Their method was developed for the multi-band imaging data of the SDSS, with observations in the u^* , g^* , r^* , i^* , and z^* filters. A series of overlapping color and magnitude cuts was used in the $g^* - r^*$ vs. r^* space and Voronoi tessellations were run on the resulting galaxy subsets.
- Another algorithm was designed to detect structures in the multi-band SDSS imaging data, the cut-and-enhance (hereafter CE) method of Goto et al. (2002). Its use of color and magnitude information is more refined than the Kim et al. (2002) approach. In contrast to the latter, the CE uses ten overlapping color-magnitude cuts in three different color spaces: $r^* - i^*$ vs. r^* , $i^* - z^*$ vs. r^* , and $g^* - r^*$ vs. r^* . It furthermore utilizes four overlapping color-color cuts in $r^* - i^*$ vs. $g^* - r^*$, and $i^* - z^*$ vs. $r^* - i^*$. All of those 34 cuts are done independently of each other. A density enhancement algorithm, that amplifies galaxy pairs, that are close in color, as well as angular separation, is then applied to the individual subsets. In the last step, the 34 cluster candidate lists are merged into a final cluster catalog.

Using red sequences in color-magnitude diagrams as tracers of galaxy clusters has a number of advantages. As was shown in Sect. 5.4.1, red sequences seem to exist independently of cluster richness, shape, radial profile, or luminosity function. Furthermore, the assumptions on colors of cluster members appear to be very robust and relatively independent of cluster type. Consequently, the use of red sequence sensitive methods should yield structure catalogs, that are rather unbiased with respect to any cluster model assumptions. The connection between the red sequence overall color and redshift, on the other hand, turns the red sequence into an interesting tool for the deprojection of elliptical galaxy populations.

However, in case of very sparsely sampled structures, i.e. structures with only few members brighter than the limiting magnitude, the red sequence can become very difficult to detect. This is simply due to the fact, that the sequence usually stands out as an overdensity in color above the fore- and background distribution of other galaxies, while in the case of few visible members, the red sequence signal is dwarfed by the fore- and background noise.

In order to test the authenticity of the MUNICS EXT-FOF clusters, $m_V - m_I$ vs. m_I CMDs are made for every structure candidate.

First, the colors and magnitudes of the EXT-FOF members of a given cluster, that are neither V , nor I band drop-outs are plotted and highlighted. Then, a subset of objects is selected from the MUNICS galaxy catalog. This subset consists of all galaxies i within a projected aperture of $5r_c$ around the center of the EXT-FOF cluster candidate, independent of their own photometric redshift

$$2 \sin \frac{\theta_{ic}}{2} d_A(H_0, \Omega_M, \Omega_\Lambda, \bar{z}) \leq 5r_c, \quad (5.26)$$

where θ_{ic} is the angle between galaxy i and the cluster center, and \bar{z} is the mean redshift of the EXT-FOF cluster. The core radius is again set to $r_c = 0.125 h^{-1} \text{Mpc}$. All of these galaxies, that are neither V , nor I band drop-outs, are then added to the plot, to test whether the EXT-FOF algorithm ignores a significant number of possible cluster candidates.

Finally, a model red sequence for the structure is plotted. The model color-magnitude tracks are taken from Gladders & Yee (2000), and are based on publications by Coleman et al. (1980), and Kodama (1997). Their models are given in AB magnitudes for ten redshifts ranging from 0.1 to 1.0, and for an open cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.2$, and $\Omega_\Lambda = 0$. These models are transformed into the Vega magnitude system used for MUNICS, and are also corrected for a flat cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$ (see Sect. 6.2.1). Fig. 5.3 illustrates the resulting expected variation of the cluster red sequences with redshift in the $m_V - m_I$ vs. m_I color-magnitude diagrams. In order to compare the distribution of EXT-FOF structure members with the red sequence models, the tracks are interpolated and plotted for the structure mean redshift \bar{z} .

Examples of these plots are shown in Sect. 6.3.2.

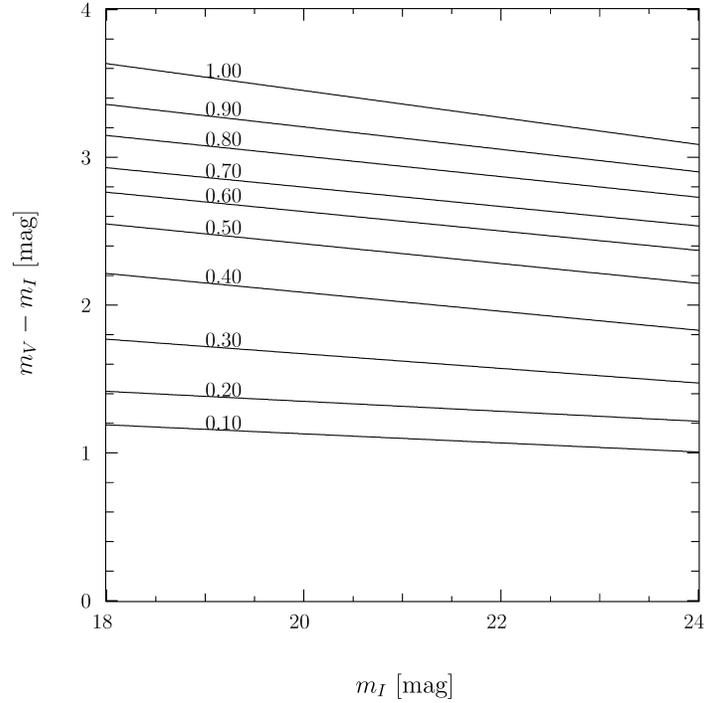


Figure 5.3: $m_V - m_I$ vs. m_I model red sequences. The numbers above the tracks show the redshift.

In case of realistic structures, the majority of the EXT-FOF members are expected to lie on an almost horizontal line. The model track at \bar{z} is not necessarily the best-fitting sequence, simply because of the large uncertainties in the photometric redshifts and the corresponding uncertainty in the cluster mean redshift. If the number of EXT-FOF members is relatively small, it cannot be expected that the red sequence stands out above the fore- and background distribution of galaxies. If the EXT-FOF algorithm omits parts of an obvious red sequence in the CMDs, the chosen linking parameters are too small and have to be increased.

Chapter 6

Structures in the MUNICS Survey

Having explained all the necessary ingredients for structure finding, this chapter shows the application of the EXT-FOF algorithm to the MUNICS survey. Following the recipe described here, the algorithm can be run on any other photometric redshift galaxy catalog, as well.

In Sect. 6.1 an overview of the input galaxy catalog is given. Special emphasis is put on the definition of the geometric borders of the individual MUNICS fields, as well as the identification of detected objects as galaxies. Sect. 6.2 describes the application of the EXT-FOF algorithm to structure finding in the MUNICS survey. The utilized equations, cosmology and other parameters are listed, and the optimization of the linking parameters is presented. In Sect. 6.3 an overview is given of the entire structure catalog. The field S6F5 is used as an example to describe the quality assessment of the individual clusters and a detailed look is put on a few structures of S6F5, showing their Voronoi tessellation and likelihood maps, as well as their CMDs. Sect. 6.4 finally deals with the spectroscopic verification of the EXT-FOF clusters.

6.1 The galaxy catalog

MUNICS is a K' -selected photometric redshift galaxy survey, that has a 50% completeness limit for point sources down to an average of $m_{K'} \approx 19.1$ mag. The survey consists of observations in four optical passbands, B , V , R , I , and two near-infrared bands, J , and K' . The area covered by all six filters in the eight fields with the best photometric homogeneity and seeing is approximately 0.27 deg^2 . The survey's most important properties are already mentioned in Sect. 1.3. A detailed description of the survey motivation and layout, the photometric redshift determination, the completenesses of the individual fields, and the spectroscopic subset is given in [Drory et al. \(2001b, 2003\)](#), [Snigula et al. 2002](#), and [Feulner et al. \(2003\)](#). The dataset used in this thesis was created in August 2003.

6.1.1 The field geometry

The geometrical borders of the MUNICS fields are of complicated shape. They are the result of the superposition of the images taken in the six passbands. The four optical images were taken with the Calar Alto Faint Object Spectrograph (CAFOS) focal reducer in direct imaging mode at the Calar Alto 2.2m telescope. The near-infrared observations were done with the Omega Prime camera at the prime focus of the Calar Alto 3.5m telescope. While the optical images are of circular shape with a diameter of approximately $16'$, the near-infrared images are squares with a 6.75×6.75 arcmin² field of view (Drory et al., 2001b). To use the CAFOS field of view most efficiently, overlapping 2×2 mosaics were created of the near-infrared images in J and K' . As a result, each MUNICS field consists of the overlap area of the four circular CAFOS apertures and the two 2×2 squares of the Omega Prime mosaics, creating a somewhat frayed final image.

Table 1.4 lists the central right ascension and declination of the MUNICS fields, as well as the size of their areas. The geometric layout of the final overlap region of each field can be seen in Fig. 6.3, as well as in the Figures B.1, B.3, B.5, B.7, B.9, B.11, and B.13 in Appendix B.

6.1.2 The spectroscopic sample

The search for structures made up of galaxies naturally requires an input object catalog consisting of galaxies with as little stellar contamination as possible. However, separating the stars from the galaxies is a non-trivial business in a multi-color survey, like MUNICS. Two informations can be used for this separation. One is the morphological classification describing the projected light distribution of the examined object, which is either extended and thus most likely a galaxy, or point-like, i.e. most likely stellar. This information is given by the YODA detection algorithm (Drory, 2003) for every object in each passband. In the MUNICS survey, an object is said to be extended, if it has an extended light distribution in at least one of the three passbands with the best signal to noise. The other available information is the best-fitting spectral energy distribution (SED). The photometric redshift determination code (Bender et al., 2001) used for MUNICS provides the deviations χ_{star}^2 and χ_{gal}^2 of an object's measured magnitudes from the best-fitting stellar and galaxy SEDs. A simple comparison of those deviations establishes the most likely object type.

In order to calibrate the morphology- and SED-based classification of detected objects as stars or galaxies, the spectroscopic sample described in Feulner et al. (2003) is used. The advantage of said dataset is the level of certainty with which objects can be identified as stars or galaxies based on their measured spectra.

The spectroscopic dataset was chosen to follow two basic approaches (Feulner et al., 2003). On the one hand, a K' band magnitude limited sample focused on extended objects was created. On the other hand, a small sample of purely magnitude

Field	Spectra	Galaxies	Stars
S2F1	183	162	21
S2F5	83	71	12
S3F5	10	9	1
S5F1	104	102	2
S5F5	15	14	1
S6F1	9	7	2
S6F5	145	108	37
S7F5	82	57	25
Total	631	530	101

Table 6.1: Spectroscopic sample of MUNICS objects. The second column shows the total number of spectra in the given field. The third and fourth columns list the number of spectra identified as galaxies and stars respectively.

Field	Galaxy & Extended	Galaxy & Point-like	Star & Extended	Star & Point-like
S2F1	154	8	8	13
S2F5	70	1	6	6
S3F5	8	1	0	1
S5F1	100	2	2	0
S5F5	14	0	0	1
S6F1	6	1	1	1
S6F5	103	5	7	30
S7F5	55	2	0	25
Total	510 (96.2 %)	20 (3.8 %)	24 (23.8 %)	77 (76.2 %)

Table 6.2: Image-based classification of the spectroscopic sample. The second and third column list the number of spectroscopically verified galaxies having an extended or point-like morphology, respectively. The fourth and fifth column show the number of verified stars with extended and point-like profiles. The percentages in brackets refer to the total number of galaxy (second and third column) and star (fourth and fifth column) spectra.

selected objects was chosen with the intention of testing the image-based object classification. The resulting spectroscopic subset is representative of the photometric sample. This can also be seen by a comparison of the upper and lower panel of Fig. 6.1. The graphic illustrates the deviations χ_{star}^2 and χ_{gal}^2 of all objects in the spectroscopic (upper panel) and photometric (lower panel) datasets. Both distributions exhibit the same features.

The spectroscopic sample consists of a total of 631 objects, of which 530 have galaxy spectra, and the remaining 101 are stars. Table 6.1 shows the breakdown of the spectroscopic sample into the individual fields and object types.

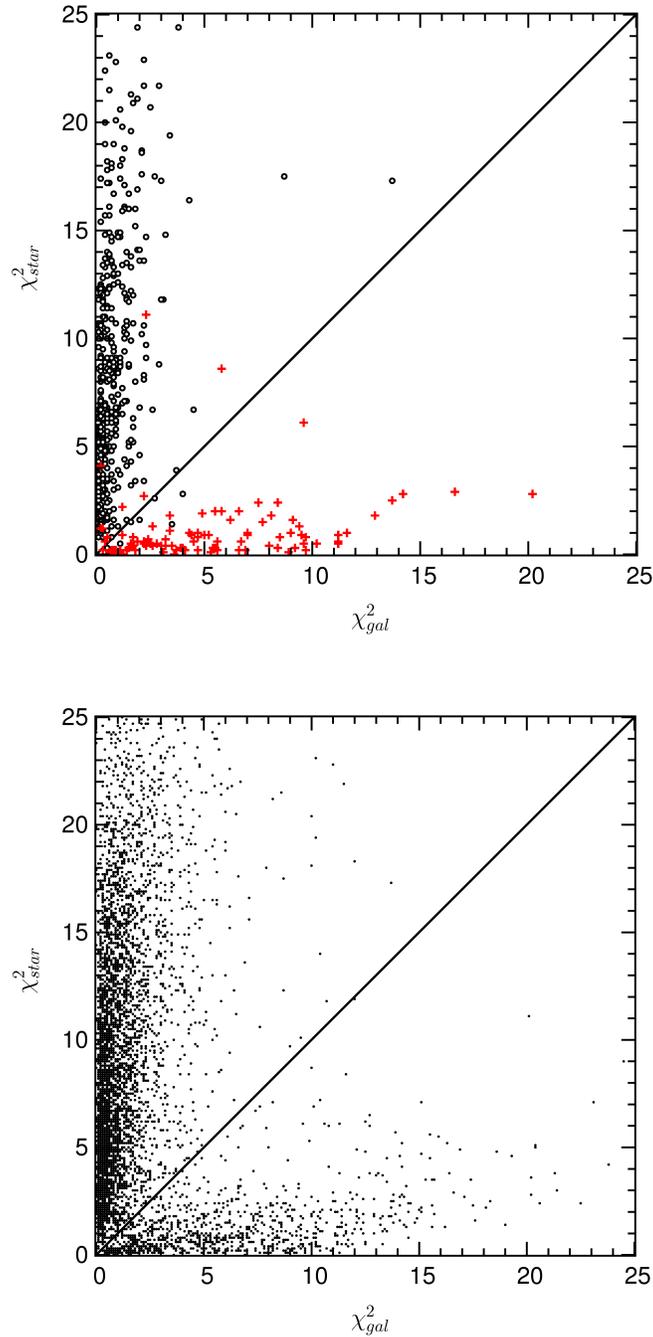


Figure 6.1: SED-based classification of objects in the spectroscopic and photometric datasets. **Upper panel:** This graphic illustrates the deviations χ_{star}^2 and χ_{gal}^2 of the stars (red crosses) and galaxies (black circles) in the spectroscopic sample. **Lower panel:** The graphic shows the deviations from the best-fitting stellar and galaxy SEDs for all objects in the photometric MUNICS survey. The solid line in both graphics shows the separating criterion for the SED-based identification used here.

Field	Galaxy & $\chi_{gal}^2 < \chi_{star}^2$	Galaxy & $\chi_{gal}^2 \geq \chi_{star}^2$	Star & $\chi_{gal}^2 < \chi_{star}^2$	Star & $\chi_{gal}^2 \geq \chi_{star}^2$
S2F1	159	3	2	19
S2F5	70	1	2	10
S3F5	9	0	0	1
S5F1	102	0	1	1
S5F5	13	1	0	1
S6F1	7	0	0	2
S6F5	106	2	2	35
S7F5	57	0	2	23
Total	523 (98.7 %)	7 (1.3 %)	9 (8.9 %)	92 (91.1 %)

Table 6.3: SED-based classification of the spectroscopic sample. The second and third column list the number of spectroscopically verified galaxies with best-fitting galactic and stellar SEDs, respectively. The fourth and fifth column show the number of verified stars with best-fitting galactic and stellar SEDs. The percentages in brackets refer to the total number of galaxy (second and third column) and star (fourth and fifth column) spectra.

Table 6.2 shows a comparison of spectroscopically confirmed object type and light distribution. As can be seen in the table the agreement between a purely image-based classification and the actual object type is very good in case of the galaxies. About 96% of the galaxies in the spectroscopic sample have an extended light distribution. However only 76% of the stars exhibit a point-like profile. The rather high fraction of stars with an extended profile comes mainly from faint stars. The discrimination between a faint star’s signal and the background noise is difficult and the object’s light distribution tends to be mistaken for extended. Yet this effect does not cause any problems in the successful identification of galaxies, as long as the SED-based classification is also taken into account (see Table 6.4).

The upper panel of Fig. 6.1 illustrates the deviations from the best-fitting stellar and galaxy SEDs for the entire spectroscopic dataset. Objects that exhibit galaxy spectra are plotted as black circles, stars are denoted by red crosses. The solid line represents the part of the parameter space, where stellar and galaxy SEDs are equally capable of describing an object’s measured magnitude distribution. This graphic shows clearly that stars and galaxies are well-separated in the examined parameter space. There is absolutely no indication, that the identity relation $\chi_{star}^2 = \chi_{gal}^2$ should not be a suitable separator for the SED-based classification. Table 6.3 lists the number of galaxies and stars with best-fitting galaxy and stellar SEDs. Almost 99% of the galaxies have deviations $\chi_{gal}^2 < \chi_{star}^2$ and 91% of the stars have $\chi_{gal}^2 \geq \chi_{star}^2$.

The combination of the image- and SED-based criterion can be seen in Table 6.4. A total of 509 objects exhibit typical galaxy properties, i.e. an extended light distribution and a best-fitting galaxy SED, while 74 objects have characteristic stellar features, i.e. point-like distributions and best-fitting stellar SEDs. Only 48 ($\sim 7.6\%$) objects

Field	Galaxy & Extended & $\chi_{gal}^2 < \chi_{star}^2$	Star & Point-like & $\chi_{gal}^2 \geq \chi_{star}^2$	Galaxy & Point-like & $\chi_{gal}^2 \geq \chi_{star}^2$	Star & Extended & $\chi_{gal}^2 < \chi_{star}^2$
S2F1	152	11	1	0
S2F5	69	5	0	1
S3F5	8	1	0	0
S5F1	100	0	0	1
S5F5	13	1	0	0
S6F1	6	1	0	0
S6F5	102	30	1	2
S7F5	55	23	0	0
Total	505 (95.3 %)	72 (71.3 %)	2 (2.7 %)	4 (0.8 %)

Table 6.4: Combination of image- and SED-based classification of the spectroscopic sample. The second and third column list the number of spectroscopic galaxies and stars having their object type’s typical morphologies and best-fitting SEDs. Whereas the fourth and fifth column show the number of spectroscopic galaxies and stars with atypical morphological and SED properties. The percentages in brackets refer to the total number of galaxy (second column) and star (third column) spectra, as well as to the amount of objects with typical stellar (fourth column) and galaxy (fifth column) features.

have mixed characteristics. The table shows that about 95% of the galaxies have an extended light distribution and a best-fitting galaxy SED. At the same time, less than 1% of the objects with typical galaxy features are actually stars. 71% of the stars have point-like distributions and best-fitting stellar SEDs, and only about 3% of the objects with characteristic stellar properties are in fact galaxies. Thus, using this classification scheme for the entire photometric dataset, it can be expected that roughly 95% of the real galaxies will be included in the so-created galaxy catalog and that about 1% of the catalog members are actually stars.

6.1.3 Classification of all galaxies

The classification scheme described above is now applied to all objects within the geometric borders, i.e. overlap area of all six passbands, of the MUNICS fields.

The lower panel of Fig. 6.1 shows the deviations from the best-fitting stellar and galaxy SEDs for all objects in the photometric survey. Like in the spectroscopic subset (see upper panel of the same figure), there is a clear tendency for clustering of objects along the axes of the parameter space and a zone of avoidance around the equality line. Consequently, the identity relation $\chi_{star}^2 = \chi_{gal}^2$ once more seems to be a suitable separator for the SED-based classification.

Snigula et al. 2002 demonstrated that a biasing exists against high redshift early type galaxies. Their light distribution follows a de Vaucouleurs profile, which is

Field	Objects	Extended & $\chi_{gal}^2 < \chi_{star}^2$		Extended & $\chi_{gal}^2 \geq \chi_{star}^2$		Point-like & $\chi_{gal}^2 < \chi_{star}^2$		Point-like & $\chi_{gal}^2 \geq \chi_{star}^2$	
		#	%	#	%	#	%	#	%
S2F1	665	479	72.0	41	6.2	47	7.1	98	14.7
S2F5	864	683	79.1	56	6.5	35	4.1	90	10.4
S3F5	789	603	76.4	38	4.8	41	5.2	107	13.6
S5F1	644	516	80.1	32	5.0	34	5.3	62	9.6
S5F5	485	360	74.2	43	8.9	24	4.9	58	12.0
S6F1	552	435	78.8	63	11.4	11	2.0	43	7.8
S6F5	867	677	78.1	48	5.5	45	5.2	97	11.2
S7F5	788	609	77.3	24	3.0	56	7.1	99	12.6
Total	5654	4362	77.1	345	6.1	293	5.2	654	11.6

Table 6.5: Image- and SED-based classification of the photometric dataset. The second column shows the total number of objects in the given field. Columns three and six list the number and percentage of objects having typical galaxy and stellar morphology and SED. While the fourth and fifth column show the objects with mixed morphological and SED features.

steeper in its central region than the exponential profile typical for a spiral galaxy. Consequently, the number of consecutive pixels with a signal to noise above the detection threshold is likely to fall below the accepted minimum and said elliptical cannot be detected. It should be kept in mind, that the resulting galaxy catalog will be most likely depleted of early types for redshifts above unity, leading to a reduced probability of finding structures in that redshift range, since structures contain large fractions of early type galaxies (see Table 1.3).

Table 6.5 shows the number of objects having typical galaxy (column three) and stellar (column six) features, as well as the number of objects with mixed characteristics (columns four and five). Roughly 77% of the entire object catalog are thus classified as galaxies, while almost 12% are identified as stellar. About 11% of the objects exhibit mixed classifications. In the following, only those 4362 objects are used for cluster finding, that have both an extended light distribution and deviations of $\chi_{gal}^2 < \chi_{star}^2$.

Structure finding test runs have been made with input galaxy catalogs consisting of the 5000 objects having extended light distributions or deviations of $\chi_{gal}^2 < \chi_{star}^2$. It proved to be impossible to decide, which of the two galaxy definitions yielded the more credible cluster catalog.

6.2 Finding the structures

At last, this section sees the application of the EXT-FOF technique, that has been described and tested in the Chapters 3 and 4, to the MUNICS galaxy catalog. The

linking parameters, that represent the key to any structure finding based on friends-of-friends algorithms, are adjusted to the MUNICS survey with the goal of creating a conservative catalog. The methods that are used for the linking parameter optimization are described in Chapter 5.

6.2.1 The algorithm

The general make-up of the EXT-FOF algorithm was explained in detail in Sect. 3.2. In the following, the equations representing the linking criteria and other important variables used for structure finding in MUNICS are listed.

Given the depth of the survey, the cosmology has to be included in the determination of the projected distance D_{ij} . As a result, eqs. (3.5) and (3.7) are used, and the projected linking criterion is given by

$$D_{ij} := 2 \sin \frac{\theta_{ij}}{2} d_A(H_0, \Omega_M, \Omega_\Lambda, z_{ini}) \leq D_L. \quad (6.1)$$

The cosmological parameters are set to $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$.

The velocity linking criteria are defined as described in the eqs. (3.8), (3.9), (3.10), and (3.11), and turn into

$$V_i := c |z_i - z_{ini}| \leq \left[(V_L/2)^2 + (c\delta z_i)^2 \right]^{1/2} =: V_{L,i}, \quad (6.2)$$

and

$$V_j := c |z_j - z_{ini}| \leq \left[(V_L/2)^2 + (c\delta z_j)^2 \right]^{1/2} =: V_{L,j}. \quad (6.3)$$

The minimum number of galaxies per group-candidate is again set to $N_{min} = 3$, and the stepsize of the redshift slicing loop is fixed to $\Delta z_{ini} = 10^{-3}$. The used photometric redshift determination code has a resolution of 0.02. Consequently the resolution of the EXT-FOF algorithm has to be at least of the same order of magnitude for z_{ini} to be able to approximate every possible \bar{V} of the input galaxy dataset (see motivation for the introduction of z_{ini} in Sect. 3.2). Numerical tests with stepsizes of 10^{-1} , 10^{-2} , 10^{-3} , 10^{-4} , and 10^{-5} have proven that there is no difference between structure catalogs having $\Delta z_{ini} \leq 10^{-2}$. The slightly refined stepsize of 10^{-3} is chosen to guarantee that no other numerical effects could play a role.

6.2.2 Optimizing the linking parameters

The set of linking parameters has to be customized to suit the input galaxy dataset, as well as to yield the targeted type of structure catalog. For example, in the case of structure finding in a spectroscopic dataset, the size of the linking distance and velocity can be fine-tuned to either detect all structures, or only the rich clusters. In case of photometric redshift catalogs this level of fine-tuning turns out to be impossible. This is simply due to the uncertainty in the distance determination in photometric redshift datasets. Consequently, the approach followed here is to find any type of structure, independent of richness.

Varying the original linking parameters D_L and V_L with redshift on the basis of a K' band luminosity function, analogous to the CFA1 and LCRS FOF or EXT-FOF runs, as described in Sects. 4.1, 4.2, or 4.3, is also impossible in surveys with depths comparable to MUNICS. The resulting linking distances and velocities at the high redshift end of the distribution would simply exceed cluster dimensions, or be far too small to be realistic in case of the low redshift end. Thus, the linking parameters D_L and V_L are not varied with redshift, leading to a structure catalog uncorrected for redshift dependent biases. However, since no value is set on cluster richnesses anyways, and the main goal of this structure finding effort is the creation of a conservative catalog, this biasing poses no problem. As was already explained in Sect. 4.3.3, the authenticity tests of Chapter 5 are used in the end to separate the plausible from the spurious detections.

In Sect. 3.3.3 it was argued that the photometric redshift errors δz will dominate the right sides of eqs. (6.2) and (6.3), making the choice of V_L de facto irrelevant in most cases. Yet it was also mentioned that an effective upper limit can be set on the EXT-FOF velocity linking parameters $V_{L,i}$ and $V_{L,j}$ by removing objects with very large redshift errors from the input galaxy dataset. Thus, a pre-selection step is integrated into the MUNICS EXT-FOF technique, only including galaxies with

$$\delta z \leq \delta z_L, \quad (6.4)$$

where δz_L defines in the end the upper limit for $V_{L,i}$ and $V_{L,j}$.

Consequently, the choice of the linking parameters presents an optimization problem in a three-parameter space, consisting of D_L , V_L , and δz_L . Following the reasoning of Sect. 3.3, D_L and V_L are chosen of the order of typical cluster properties, while δz_L is set to the order of the standard deviation $\sigma(\delta z)$ of the MUNICS photometric redshifts, in order not to remove too many galaxies from the input catalog. Since $\sigma(\delta z)$ is increasing with redshift (see Sect. 1.3), δz_L will also increase with redshift.

The following value sets are used for the linking parameters:

- $D_L \in \{0.5, 1.0, 1.5, 2.0, 2.5, 5.0, 10.0\} \cdot r_c$,
with a core radius of $r_c := 0.125 h^{-1}$ Mpc (see Table 1.2),

Field	$\delta z_L = \sigma(\delta z) \cdot$			
	1.2	1.5	1.7	2.0
S2F1				++
S2F5	++	+		
S3F5		++		
S5F1		+	+	++
S5F5			+	++
S6F1		+	++	
S6F5	+	+	++	
S7F5	+	++	+	

Table 6.6: Optimized redshift error pre-selection parameters for all fields. A “+” denotes an acceptable value for δz_L , while “++” stands for the chosen best-fitting parameter. A blank signifies unacceptable values.

- $V_L \in \{0.5, 1.0, 1.5, 2.0\} \cdot \sigma_c$,
with a typical cluster velocity dispersion $\sigma_c := 1000 \text{ km s}^{-1}$ (see also Table 1.2),
and
- $\delta z_L \in \{0.5, 0.7, 1.0, 1.2, 1.5, 1.7, 2.0, 2.2, 3.0\} \cdot \sigma(\delta z)$,
with $\sigma(\delta z) := 0.055 \cdot (1+z)$, as given in Drory et al. (2003).

The EXT-FOF algorithm is run on all eight fields with the 252 combinations of the three linking parameters. For each of those runs Voronoi and likelihood maps, as well as the CMDs are created, following the procedures described in Sects. 5.2, 5.3, and 5.4. The optimized set of linking parameters is the one, that yields EXT-FOF structures with the highest level of agreement with overdense Voronoi regions, peaks in the likelihood maps and red sequences in the CMDs. This optimization has to be done independently for all fields, due to small inhomogeneities among the fields, e.g. the completeness limits (Table 1.4).

The resulting EXT-FOF structure catalogs show that the used range of linking parameters is absolutely sufficient: An increase of the two value sets for D_L and δz_L in either direction is not necessary, since on both ends of the intervals the structure catalogs are becoming highly unreasonable. As expected, a variation of V_L within physically reasonable intervals does not change the resulting cluster catalog at all. A finer grid in D_L , and δz_L is also not necessary. At the given resolution in linking parameter space it is already very difficult to decide, which of the reasonable parameter combinations fits best.

The optimized parameter set turns out to be:

- $D_L = 1.0 \cdot r_c = 0.125 h^{-1} \text{ Mpc}$ for all eight fields,
- $V_L = 1.0 \cdot \sigma_c = 1000 \text{ km s}^{-1}$ is selected for all fields, while

Field	Selected galaxies	
	#	%
S2F1	358	74.7
S2F5	336	49.2
S3F5	355	58.9
S5F1	346	67.1
S5F5	286	79.4
S6F1	319	73.3
S6F5	457	67.5
S7F5	332	54.5
Total	2789	63.9

Table 6.7: Redshift error pre-selected galaxies. The number of galaxies fulfilling the δz_L pre-selection listed in Table 6.6 is shown, as well as the corresponding percentage of galaxies.

- δz_L varies sensibly from field to field. The best-fitting values for δz_L for each field are given in Table 6.6. A “+” sign denotes an acceptable value, “++” signifies the chosen best-fitting value for δz_L .

The differences between the EXT-FOF catalogs resulting from δz_L parameters marked with a “+” and those finally chosen for the MUNICS cluster catalogs are only minor.

Table 6.7 lists the final number of galaxies included in the EXT-FOF structure finding process. A total of 2789 galaxies, or roughly 64% of all galaxies fulfill the δz_L pre-selection criterion.

6.3 The structure catalog

The resulting structure catalog contains a total of 162 groups and clusters in all eight fields (see column two in Table 6.8). 890 galaxies are found to be members of these structures, corresponding to roughly 20% of the MUNICS galaxies, or 32% of the subset of redshift error pre-selected galaxies (see columns five to seven in Table 6.8). 122 ($\approx 75\%$) of the structures have a mean redshift of $0.3 \leq \bar{z} \leq 0.8$. 621 of the structure galaxies are lying in a redshift bin with $0.3 \leq z_i \leq 0.8$, that corresponds approximately to 27% of the galaxies, or 38% of the pre-selected galaxies in this redshift bin.

Fig. 6.2 illustrates the statistics of the entire structure catalog. The upper panel shows the distribution of galaxies and structures with redshift. The dashed and solid black lines depict the distribution of all galaxies and of the selected ones with their photometric redshifts z_{phot} . The dotted line shows the distribution of the spectroscopic subsample according to their redshifts z_{spec} . The structure distribution is plotted in red

Field	Structs.	$n_s [h^3\text{Mpc}^{-3}]$	$r_s [h^{-1}\text{Mpc}]$	Galaxies in structures		
				#	% Gal.	% Sel. gal.
S2F1	14	$5.3 \cdot 10^{-4}$	12.3	121	25.3	33.8
S2F5	25	$8.7 \cdot 10^{-4}$	10.5	114	16.7	33.9
S3F5	18	$5.9 \cdot 10^{-4}$	11.9	91	15.1	25.6
S5F1	21	$5.9 \cdot 10^{-4}$	11.9	132	25.6	38.2
S5F5	16	$6.0 \cdot 10^{-4}$	11.8	100	27.8	35.0
S6F1	20	$7.5 \cdot 10^{-4}$	11.0	94	21.6	29.5
S6F5	25	$4.2 \cdot 10^{-4}$	13.4	130	19.2	28.4
S7F5	23	$7.8 \cdot 10^{-4}$	10.9	108	17.7	32.5
Mean	–	$6.5 \cdot 10^{-4}$	11.6	–	–	–
Total	162	–	–	890	20.4	31.9

Table 6.8: Overview of the MUNICS structure catalog. The number of structures detected in each field is listed in column two. The structure number densities n_s (column three) and correlation scales r_s (column four) are determined within the redshift bin $0.3 \leq \bar{z} \leq 0.8$ for a cosmology with $H_0 = 70 \text{ km s}^{-1}\text{Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. The number of galaxies found to be structure members is listed in column five, the corresponding fractions of all and pre-selected galaxies are given in columns six and seven.

as a function of the δz -weighted mean redshift \bar{z} . The determination of photometric redshifts of nearby objects is aggravated. If the galaxy is close-by, the Balmer break is not shifted out of the B band filter, and the resulting redshift estimation has an increased uncertainty. With the given MUNICS filter set, galaxies having redshifts larger than $z_{phot} = 0.3$ (blue line) are not influenced by this problem and have good quality photometric redshifts.

The graphic shows that the distribution of all galaxies and of the selected ones are very similar beyond this redshift of 0.3. At a redshift around 0.8, the number of galaxies has dropped to about 50% of its peak frequency, thus the most sensitive range of the MUNICS galaxy survey is roughly covered by $z \in [0.3, 0.8]$. This also corresponds to the range of highest sensitivity for the MUNICS EXT-FOF algorithm, and clusters are found in this domain in all eight fields. Unfortunately, the distribution of the spectroscopic subsample is mainly concentrated at low redshifts, thus a spectroscopic verification of the high redshift MUNICS structures is very unlikely with the given spectroscopic redshift sample.

Structure number densities n_s and structure correlation scales $r_s = n_s^{-1/3}$ are determined for the individual fields over the redshift range $0.3 \leq \bar{z} \leq 0.8$ for the cosmology described in Sect. 6.2.1, using the overlap area of each field as given in Table 1.4. The results are listed in columns three and four of Table 6.8. The mean structure density is $6.5 \cdot 10^{-4} h^3\text{Mpc}^{-3}$. Under the assumption that only about 60% of the detected structures are physically real, as was argued in Sect. 4.3.3, a more probable estimate for n_s is $3.9 \cdot 10^{-4} h^3\text{Mpc}^{-3}$. This value agrees well with the typical group densities

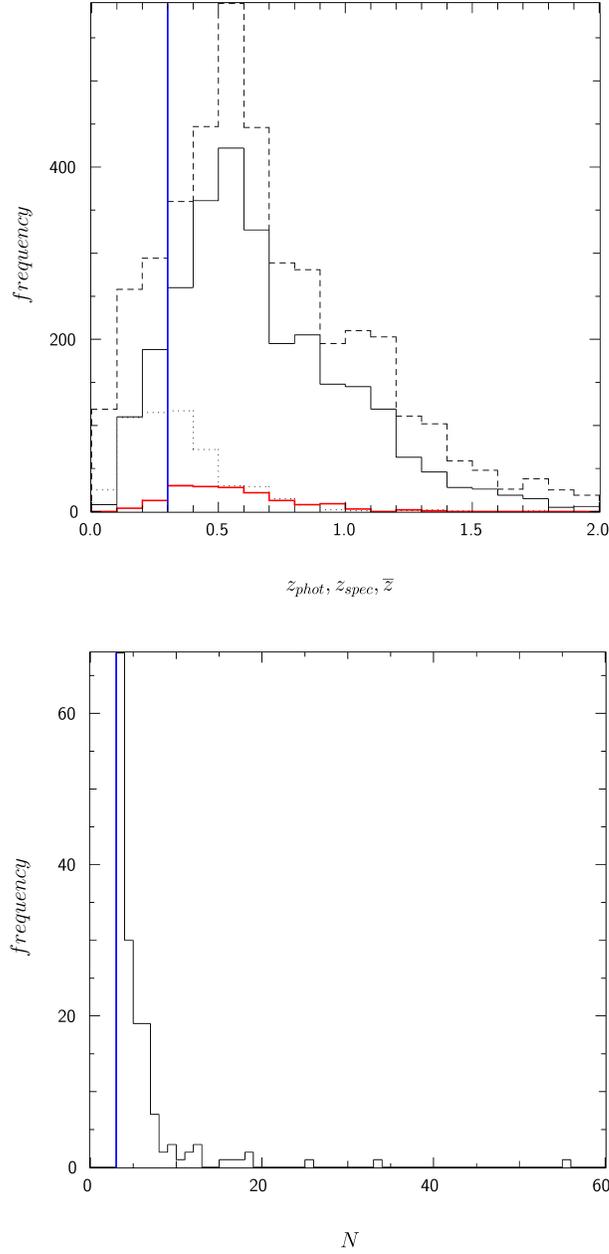


Figure 6.2: Statistics of the entire MUNICS structure catalog. **Upper panel:** The distribution of galaxies and structures with redshift. The dashed (solid) black line shows the distribution of all (redshift error pre-selected) galaxies with their photometric redshifts z_{phot} . The dotted line depicts the distribution of the spectroscopic subsample according to their redshifts z_{spec} . The structure distribution is plotted in red as a function of the mean redshift \bar{z} . The blue line at $z_{phot} = 0.3$ separates the galaxies suffering from the Balmer break problem from the others. **Lower panel:** The distribution of structure members N . The blue line resembles the limiting number of structure candidates $N_{min} = 3$.

of $(10^{-5} - 10^{-3})h^3\text{Mpc}^{-3}$, and is higher than the characteristic cluster densities of $(10^{-6} - 10^{-5})h^3\text{Mpc}^{-3}$ (see Table 1.2). That does not come as a surprise, since being a conservative structure finder, the MUNICS EXT-FOF was tuned to find all types of structures, both groups and clusters. The expected evolution of the group or cluster number density in the given low matter content universe with $\Omega_M = 0.3$ is relatively small within this redshift range (e.g. Bahcall et al. 1997; Bahcall & Fan 1998; Botzler 1999). Especially the less massive structures experience only little evolution. The mean correlation scale is $11.6h^{-1}\text{Mpc}$. Following the above mentioned reasoning, a more probable estimate for r_s is given by $13.7h^{-1}\text{Mpc}$, which is in good agreement with typical group correlation scales of $(13 \pm 2)h^{-1}\text{Mpc}$ (see again Table 1.2). For comparison, typical cluster scales are $(22 \pm 4)h^{-1}\text{Mpc}$ (Table 1.2).

The lower panel of Fig. 6.2 shows the distribution of structure members N . The blue line resembles the limiting number of structure candidates $N_{min} = 3$. Obviously many structures consist of only three galaxies. The possibility of these being the result of chance alignments of physically uncorrelated objects is relatively high, while in general structures that have more members are more trustworthy. On average, each structure exhibits 5.5 galaxies. The structure with by far the most members lies in the field S2F1. It contains 55 galaxies and is located at a low redshift of $\bar{z} = 0.207$. This cluster is real, as will be demonstrated in Sect. 6.4. It consists of a number of spectroscopically verified structures, that can not be separated by the EXT-FOF algorithm due to the low resolution of the photometric redshifts. One of those spectroscopically verified components is a known Abell cluster.

A quality assessment is made for each structure in the catalog, following the recipes described in Chapter 5. This facilitates the identification of high quality candidates and possible spurious detections. In the following, the field S6F5 is used as an example to describe – amongst others – the results of this quality assessment. The properties of the structures in the seven remaining fields are listed in Appendix B.

6.3.1 The field S6F5

The reasons for choosing S6F5 as example field are given as follows: Because of its depth (see Table 1.4), S6F5 has the structure catalog that reaches the highest mean redshift of $\bar{z} \sim 1.3$. Furthermore, as can be seen in Table 6.1, it has a relatively large number of spectroscopic redshifts, which makes the results of the verification attempts very interesting.

Table 6.9 gives an overview of the structures in the field S6F5. The first column lists the name of the structure, the second contains the amount N of members, while the third and fourth show the mean right ascension α and declination δ of the cluster center. The mean structure redshift \bar{z} is given in column five, and is weighted using the redshift errors of the individual structure members. The next three columns deal with the authenticity assessment:

- The analysis of consecutive Voronoi tessellation maps (for details see Sect. 5.2.2) showed that the maps are rather stable against the variation of the central

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	12	11:57:25.35	65:33:16	0.156	3 (2)	5	3	6	+1
2	3	11:58:19.34	65:34:16	0.183	3	5	5	2	0
3	15	11:57:55.01	65:30:15	0.261	1 (4)	2	3	6	0
4	5	11:58:19.14	65:38:29	0.356	1	1	4	4	+1
5	3	11:58:25.42	65:37:32	0.387	3 (1)	1	3	3	-1
6	5	11:58:05.09	65:39:06	0.444	2	2	1	2	0
7	6	11:57:33.83	65:36:04	0.513	1	4	1	2	0
8	4	11:57:06.24	65:36:51	0.534	1	5	3	1	0
9	8	11:57:57.30	65:35:55	0.585	1	2	2	0	0
10	16	11:58:26.74	65:35:04	0.589	1	1	3	6	+1
11	4	11:57:50.71	65:40:47	0.680	1	1	3	0	0
12	5	11:58:25.43	65:31:29	0.689	2	5	2	0	0
13	3	11:58:21.26	65:32:50	0.706	1	5	1	0	0
14	3	11:57:01.54	65:35:27	0.811	1 (1)	4	4	0	0
15	4	11:57:14.92	65:36:48	0.824	4	3	1	0	0
16	5	11:58:14.47	65:32:37	0.845	1	1	2	0	0
17	3	11:57:39.39	65:40:26	0.868	2	3	1	0	0
18	3	11:58:13.24	65:33:48	0.931	1	2	1	0	0
19	3	11:57:13.65	65:33:40	0.939	1	5	5	0	0
20	5	11:58:30.43	65:36:32	0.953	1 (2)	1	3	0	0
21	3	11:58:34.79	65:30:32	0.973	1 (1)	5	4	0	0
22	3	11:57:43.78	65:33:04	0.985	2	5	4	0	0
23	3	11:57:44.76	65:36:40	1.030	1	2	0	0	0
24	3	11:58:26.43	65:35:05	1.254	1	2	0	0	0
25	3	11:57:35.74	65:37:01	1.338	1	5	0	0	0

Table 6.9: Structures in S6F5. Column one lists the name of the EXT-FOF structure. The number N of structure members is given in the second column, the central right ascension α and declination δ are given in columns three and four, and the error weighted mean redshift \bar{z} is shown in column five. Columns six, seven and eight contain the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). See the text for an explanation of the classification. Generally, class 1 (5) implies good (bad) agreement with expected cluster properties. The numbers in brackets in column six list the structure’s edge objects in the Voronoi map (set the case there are any). The amount of available spectroscopic redshifts is listed in column nine. Column ten exhibits information about the spectroscopic verification of the structure: “+1” denotes a confirmed structure, “0” implies that the available spectra support no conclusions yet, and “-1” flags a spurious structure (see also Sect. 6.4 for more details on the verification).

redshift $z_{central}$ from one bin to the next. This proved that the chosen spacing, as well as the width of the redshift bins is sufficient. For each structure’s authenticity testing, the Voronoi map with $z_{central}$ closest to \bar{z} is used. The above mentioned stability against small variations of $z_{central}$ guarantees that the examined Voronoi map is a good representation of the actual local densities, even if \bar{z} can only be determined with limited precision. The numbers N_o and N_u of clus-

ter members lying in overdense and underdense cells are determined, as well as the amount N_e of edge objects, i.e. cluster galaxies that do not have closed polygons. A system of six classes is introduced to quantitatively estimate the level of agreement between the examined structure and the expected array of overdense cells, typical for a real cluster. Class 0 is given to structures, that contain only edge objects ($N = N_e$). Obviously, the Voronoi approach does not yield any information about these objects' authenticity. Class 1 is given to structures, with more than 80% overdense cells, i.e. $N_o/(N_o + N_u) > 0.8$. Class 2 structures fulfill $0.8 \geq N_o/(N_o + N_u) > 0.6$, class 3 have $0.6 \geq N_o/(N_o + N_u) > 0.4$, class 4 is defined by $0.4 \geq N_o/(N_o + N_u) > 0.2$, and class 5 objects fulfill $N_o/(N_o + N_u) \leq 0.2$. Thus, class 1 structures are generally in better agreement with expected cluster properties than class 5 objects. However, this classification scheme is completely independent of the total number of cluster members. Consequently, less weight should be put on this classification for structures with only few members. The Voronoi tessellation classes are listed in column six. The number of edge galaxies N_e is listed in brackets.

- The results of the likelihood estimation are shown in column seven. As described by [Postman et al. \(1996\)](#) the matched filter likelihood of a cluster does not necessarily peak at the exact cluster redshift. Consequently, a series of seven consecutive likelihood maps (for details see Sect. 5.3.2) is used here to check the consistency between EXT-FOF structures and likelihood enhancements. Once more, the map i with z_i closest to \bar{z} is used, as well as the maps with $z_i \pm 0.02$, $z_i \pm 0.04$, and $z_i \pm 0.06$. The classification consists of five categories. If the center (mean α and δ) of the structure lies in an area having a likelihood $\ln L \geq 1$ in at least four of the seven likelihood maps, the structure is given class 1. Class 2 is given to structures that fulfill this criterion in at least one, yet less than four of the maps. To take into account possible small offsets in the projected position of the structure center, two further classes are introduced. It is tested whether the cluster center is lying less than $1r_c = 0.125 h^{-1} \text{Mpc}$ from an enhanced likelihood area with $\ln L \geq 1$. If this criterion is fulfilled in at least four of the maps, the structure is given class 3. If it is only true in at least one, yet less than four of the maps, the structure is given class 4. Class 5 is given to structures that are not compatible with $\ln L \geq 1$ in any of the seven maps. Obviously, class 1 candidates agree better with the model expectations for real clusters than class 5 objects. On the other hand, it must not be forgotten that the utilized models favor clusters and are biased against poorer groups, or atypical structures.
- The CMD classification is listed in column eight. The deviation of the structure members' colors $m_V - m_I$ and magnitudes m_I from the model red sequences at the structure's mean redshift \bar{z} is used as a quantitative measure for the level of agreement between the structure's CMD and typical cluster red sequences (for details see Sect. 5.4.2). Taking into account the individual color errors of the EXT-FOF members, χ_v^2 , the reduced χ^2 deviation between the data and the model track is calculated. Structure members that are V or I band drop-outs are not considered. A system of six classes is used. If a structure has a redshift

of $\bar{z} < 0.1$ or $\bar{z} > 1.0$, no model track is available and consequently no χ_v^2 is determined. In this case no quantitative information can be extracted from the CMD, and the structure is given a CMD class 0. If the deviation has $\sqrt{\chi_v^2} < 5$ the structure is given class 1. Class 2 is given to structures with $5 \leq \sqrt{\chi_v^2} < 10$, class 3 for $10 \leq \sqrt{\chi_v^2} < 15$, class 4 is given to candidates with $15 \leq \sqrt{\chi_v^2} < 20$, and class 5 to structures with $\sqrt{\chi_v^2} \geq 20$. It should be noted though, that a class 5 structure does not necessarily imply a bad agreement of the data with any type of red sequence, it only states that the model track for \bar{z} does not fit the color distribution of the given structure.

In column nine of Table 6.9, the number of structure members having spectroscopic redshifts is listed. Column ten contains information about the spectroscopic confirmation of a structure. For the verification, the criteria described in Sect. 5.1 are used. If a candidate is verified, it is flagged with “+1”. If not all structure members have spectroscopic information yet, and the available spectra do not fulfill the verification criteria, the structure is flagged with “0”. “-1” is given to spurious structures, i.e. if all members have spectroscopic redshifts and they are not compatible with the verification criteria.

Examples of Voronoi and likelihood maps are shown in the Figs. 6.5, 6.7, and 6.9. For examples of CMDs see Figs. 6.6, 6.8, and 6.10.

An overview of the geometric layout of the field S6F5 is illustrated in the upper panel of Fig. 6.3. The overlap area of the images taken in the six passbands is plotted in white. The grey-shaded area was not observed in all filters and thus does not belong to the proper MUNICS survey used here. The image is centered on the central part of the K' band mosaic. The red circles mark the culled galaxies, that are not taken into account for structure finding due to their large errors. The black squares denote the δz_L pre-selected galaxies. Stars are not plotted.

The lower panel of Fig. 6.3 shows the distribution of structures. The culled galaxies are plotted as small crosses. The squares mark the pre-selected galaxies. The ones that do not belong to a structure are plotted black, while structure members are plotted colored. The same color is given to all members of one structure. The color also serves as a rough indicator for the cluster’s mean redshift \bar{z} . Pink marks the closest and dark red the farthest structures. The color scale is shown on the right side of the figure. The numbers at the bottom and top of the scale are the minimum and maximum \bar{z} of this field’s structures. The structure names are plotted in black at the mean α and δ . Note that in S6F5 the structure name “24” is plotted over the “10”, making it illegible.

Fig. 6.4 shows the redshift distribution of galaxies and structures in the field S6F5 (for more details see upper panel of Fig. 6.2).

6.3.2 Examples of clusters in S6F5

To illustrate the authenticity testing, a sample of three interesting structures in S6F5 with $\bar{z} > 0.3$ is examined. Their Voronoi and likelihood maps, as well as their CMDs

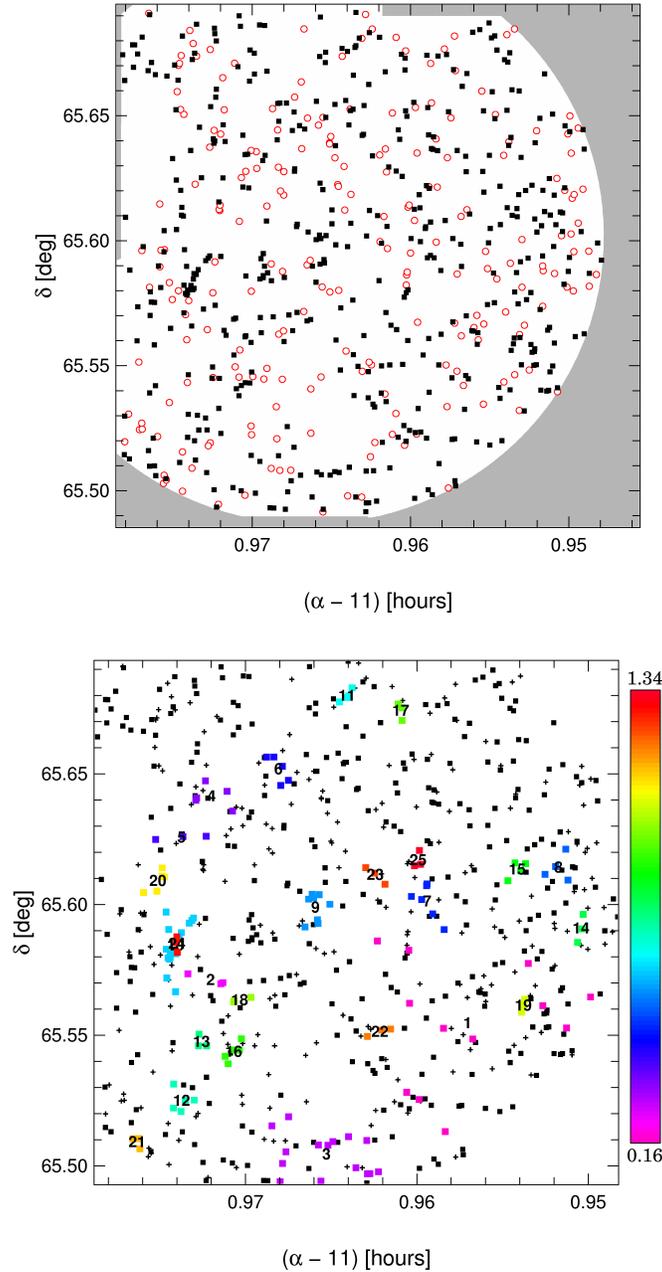


Figure 6.3: Geometrical borders and projected distribution of structures in S6F5. The **upper panel** shows the geometry of the overlap area (white). Red circles mark the culled galaxies, that are not taken into account due to their large errors, and black squares denote the δ_{z_L} pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses signify the culled galaxies, black squares are pre-selected galaxies, that do not belong to any structure. Structure members are plotted colored. All members of a structure are given the same color. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure, as well as the minimum and maximum \bar{z} . Structure names are plotted at the mean α and δ .

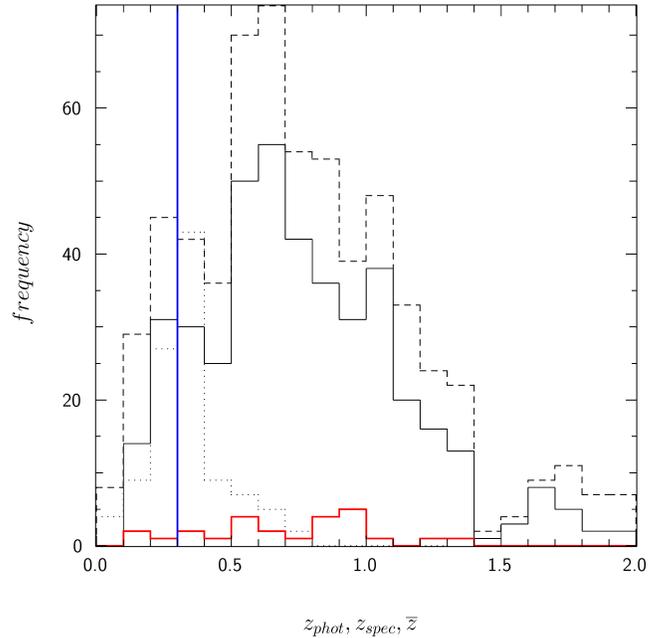


Figure 6.4: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S6F5. See upper panel of Fig. 6.2 for more details.

are shown, and their properties are discussed.

The upper panels of Figs. 6.5, 6.7, and 6.9 depict the Voronoi tessellation maps with $z_{central}$ closest to \bar{z} of the three example structures. All galaxies compatible with the given $z_{central}$ (see Sect. 5.2.2 for details) are used for the Voronoi tessellation process. Underdense cells, i.e. cells with $\delta_i \leq 1$, are marked by small squares, overdense cells are flagged by large triangles. Edge nuclei are also shown as small squares, yet do not have a closed polygon around them. The members of EXT-FOF structures, lying in this map are plotted colored. For easy reference, the same color coding is used as in the lower panel of Fig. 6.3. Field galaxies are black.

The lower panels of Figs. 6.5, 6.7, and 6.9 show the likelihood maps with grid redshift z_i closest to \bar{z} . For details on the determination of the likelihood see Sect. 5.3.2. The likelihood $\ln L$ is plotted in greyscales. White signifies higher likelihoods. Areas having an enhanced likelihood of $\ln L \geq 1$ are marked by black contours (see Sect. 6.3.1). The likelihood scale is shown to the right, as well as the maximum (top) and minimum (bottom) values of $\ln L$. The graphic also depicts all galaxies compatible with the grid redshift, following the recipe used for the creation of the Voronoi maps. These galaxies are plotted as squares. Members of EXT-FOF structures are plotted colored (same color as in Fig. 6.3). The centers of the structures are marked by small triangles and are surrounded by colored rings with an aperture of $1r_c$.

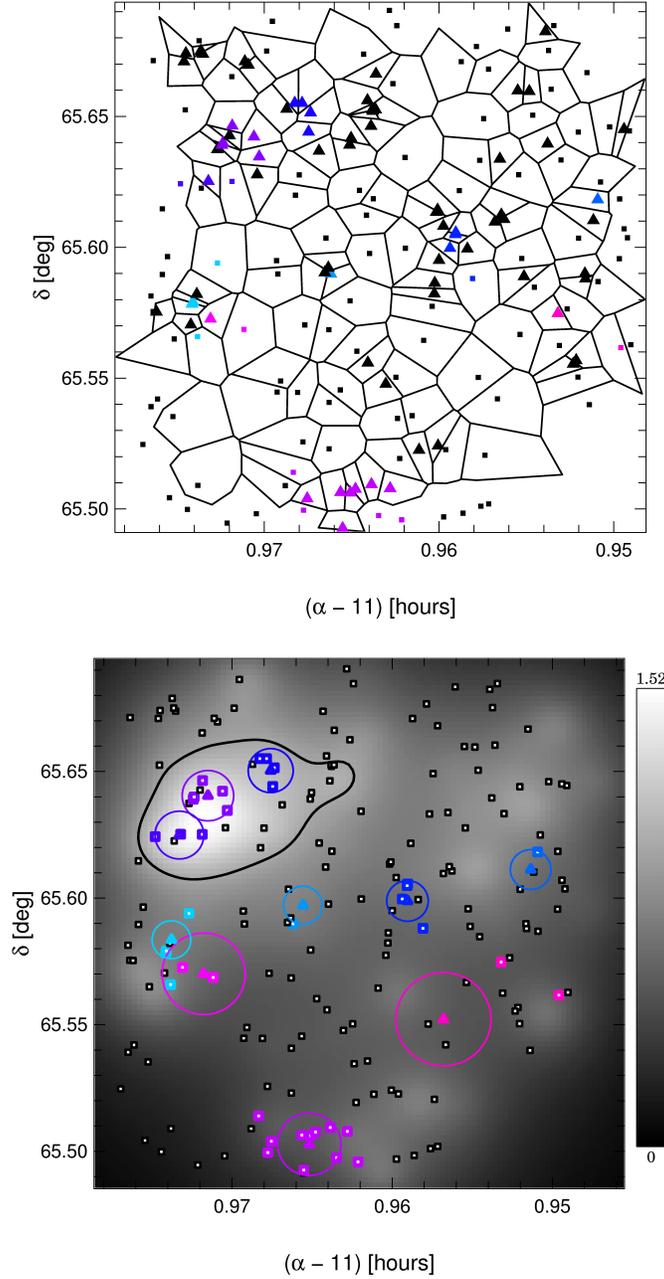


Figure 6.5: Voronoi and likelihood maps for structure 4 in S6F5. **Upper panel:** Voronoi tessellation map for all galaxies compatible with $z_{central} = 0.36$ (see Sect. 5.2.2 for details). Squares mark underdense cells, triangles symbolize overdense ($\delta_i > 1$) cells. All members of EXT-FOF structures contained in this map are plotted colored. **Lower panel:** Likelihood map for $z_i = 0.36$ (see Sect. 5.3.2 for details). The likelihood is given in greyscales. Black contours signify areas with $\ln L \geq 1$ (see Sect. 6.3.1). The likelihood scale is shown to the right. Galaxies compatible with z_i are plotted as squares. All members of EXT-FOF structures are again plotted colored. The structure centers and $1r_c$ apertures are marked by colored triangles and rings.

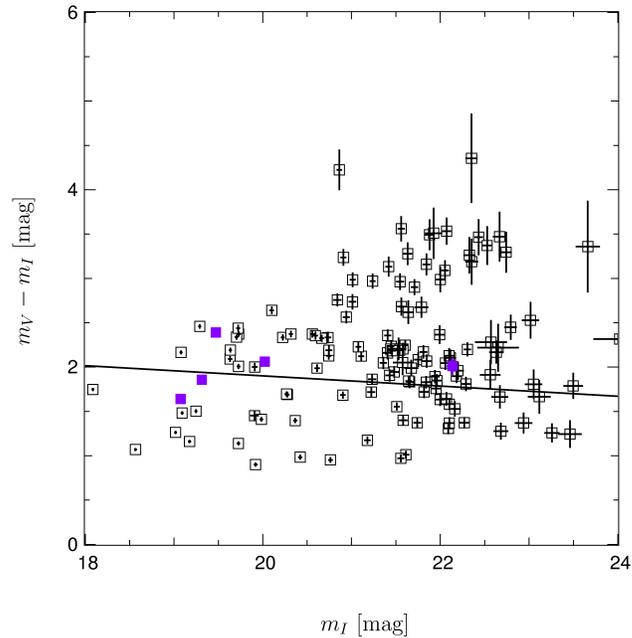


Figure 6.6: CMD of structure 4 in S6F5. The filled colored squares show the colors and magnitudes of the structure members. For easy reference, the same color is used as in the lower panel of Fig. 6.3. Open black squares mark all galaxies lying within a projected aperture of $5r_c$ around the structure center. The crosses show the 1σ error bars. The solid black line is the red sequence model track at \bar{z} (see Sect. 5.4.2 for details).

Figs. 6.6, 6.8, and 6.10 show the color-magnitude diagrams of the clusters. The filled colored squares mark the structure members. Again the same color is used as in the lower panel of Fig. 6.3. The open black squares depict all the galaxies lying within a projected radius of $5r_c$ from the structure center, independent of their redshifts (see Sect. 5.4.2). V or I band drop-outs are not considered. The 1σ error bars are shown as colored or black crosses. The solid black line is the red sequence model track at \bar{z} (Sect. 5.4.2).

As can be seen in Table 6.9, the first of the three examples, structure 4, has five members and a mean redshift of $\bar{z} = 0.356$. It is located in the upper left corner of the field (see lower panel of Fig. 6.3) at a position of $\alpha \approx 11.972^{\text{h}}$ and $\delta \approx 65.64^\circ$, and is marked in violet. Fig. 6.5 shows the Voronoi and likelihood maps for $z_{\text{central}} = z_i = 0.36$ and the CMD is given in Fig. 6.6. The Voronoi tessellation shows that structure 4 agrees well with the expected cluster behavior. All structure members are lying in consecutive overdense cells, and are even partially surrounded by further overdense cells. Some of these belong to the structures 5 and 6. The structure does not have any edge nuclei and is given a Voronoi classification of class 1. The likelihood classification of the structure also yielded class 1. The structure center

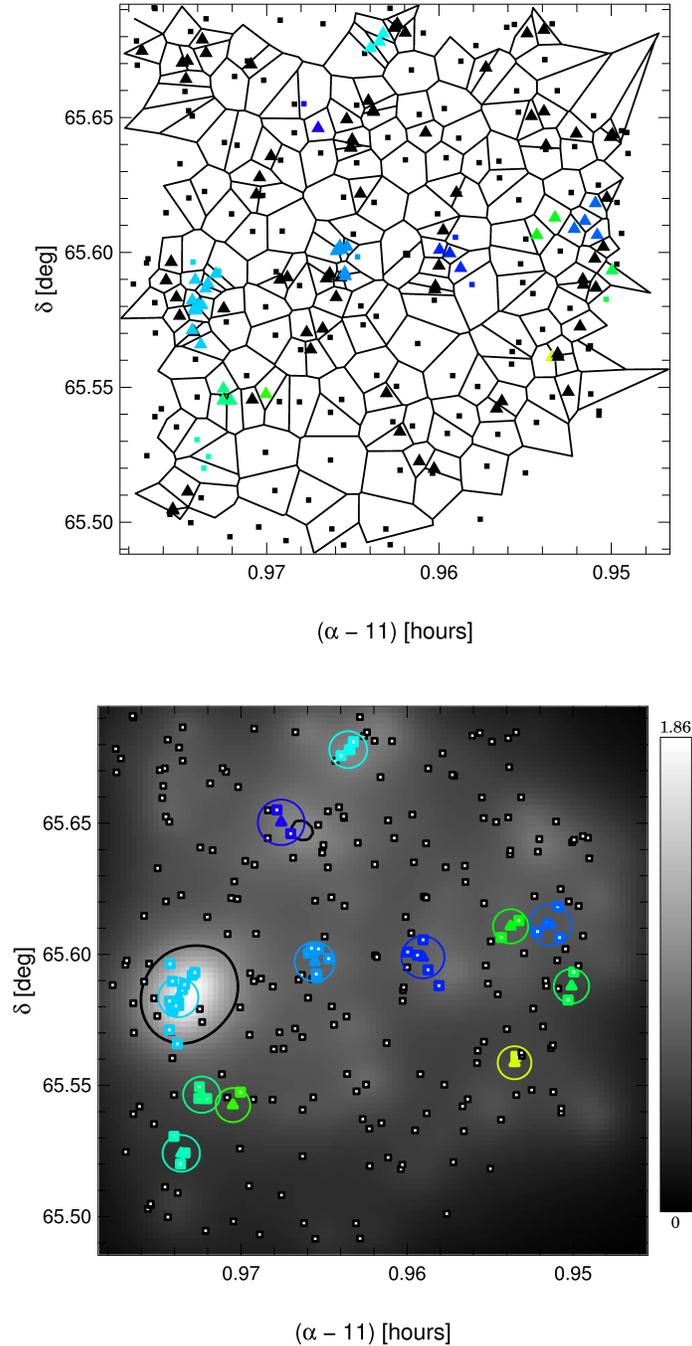


Figure 6.7: Voronoi and likelihood maps for structure 10 in S6F5. **Upper panel:** Voronoi tessellation map for $z_{central} = 0.60$. Underdense (overdense) cells are marked by squares (triangles). Structure members are plotted colored. **Lower panel:** Likelihood map for $z_i = 0.60$. The likelihood is given in greyscales. Black contours signify areas with $\ln L \geq 1$. Galaxies compatible with z_i are plotted as squares. Structure members are plotted colored, as well as structure centers (triangles) and $1r_c$ apertures (rings). See Fig. 6.5 for more details.

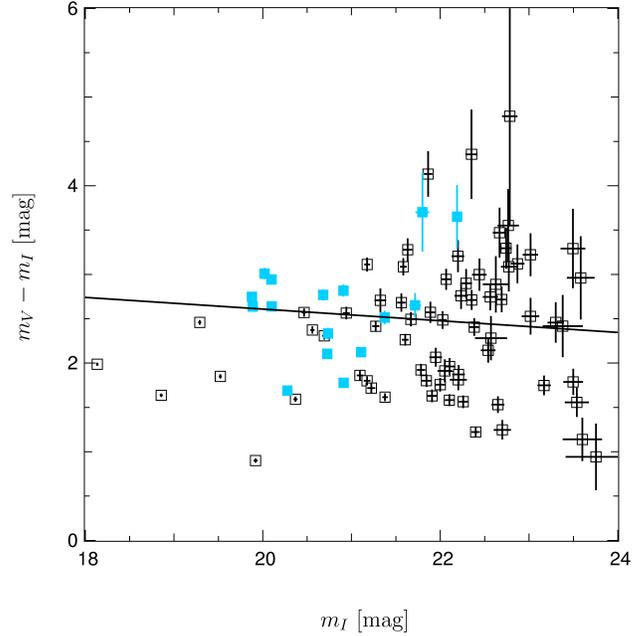


Figure 6.8: CMD of structure 10 in S6F5. Colored (black) squares show the colors and magnitudes of the structure members (all galaxies within $5r_c$). The black line is the red sequence model track at \bar{z} . See Fig. 6.6 for more details.

coincides with the enhanced likelihood areas in six of the seven maps. The maximum likelihood reached by structure 4 in these seven maps is $\ln L = 1.56$ in the bin with $z_i = 0.38$. The peak of this likelihood enhancement is shifted slightly in the direction of structure 5, which is also visible in the lower panel of Fig. 6.5. The spread of the color-magnitude distribution is relatively large, while the color errors of the individual galaxies are small, thus the CMD is given a classification of 4. Four of the structure members also have spectroscopic redshifts. Three of them correspond well with each other, resulting in the spectroscopic verification of structure 4. The true mean spectroscopic redshift of this verified cluster is $\bar{z}_{cluster} = 0.388$ (see Table 6.11) and the deviation from the EXT-FOF mean redshift is $\bar{z}_{cluster} - \bar{z} = 0.032$.

There is one more thing worth mentioning about structure 4: Structure 5 (marked in blue) with three members is sitting to the lower left of structure 4 at almost the same redshift of $\bar{z} = 0.387$. It has an acceptable Voronoi class of 3, a good likelihood class 1, and a CMD class 3, and is the only structure in the entire MUNICS survey that has been proven spurious by spectroscopy. However, one of the members has a spectroscopic redshift of $z_{spec,i} = 0.386$, being in excellent agreement with the real $\bar{z}_{cluster} = 0.388$ of the neighboring structure 4, while having a projected distance of the order of $2r_c$ from the center of structure 4. Thus it is obvious that this galaxy is also part of the cluster at $\bar{z}_{cluster} = 0.388$. This might also be an explanation for the above mentioned slight shift of the likelihood peak towards structure 5.

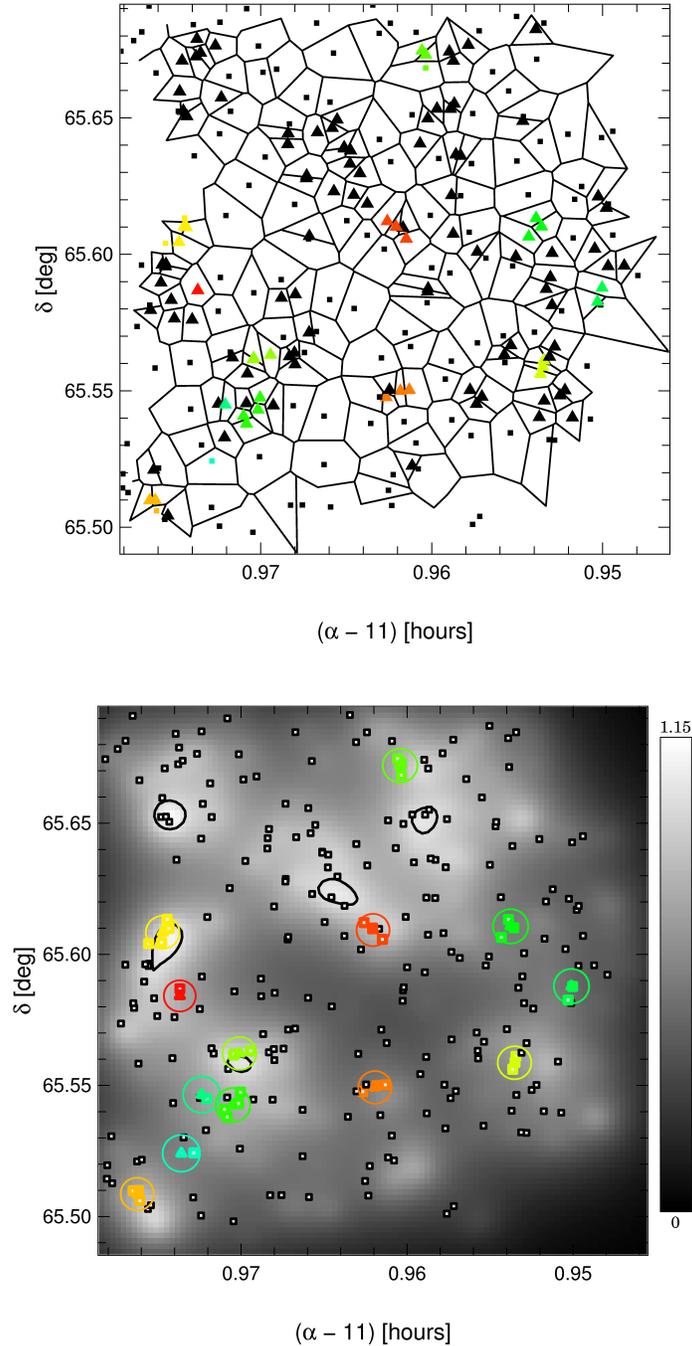


Figure 6.9: Voronoi and likelihood maps for structure 20 in S6F5. **Upper panel:** Voronoi tessellation map for $z_{central} = 0.96$. Underdense (overdense) cells are marked by squares (triangles). Structure members are plotted colored. **Lower panel:** Likelihood map for $z_i = 0.96$. The likelihood is given in greyscales. Black contours signify areas with $\ln L \geq 1$. Galaxies compatible with z_i are plotted as squares. Structure members are plotted colored, as well as structure centers (triangles) and $1r_c$ apertures (rings). See Fig. 6.5 for more details.

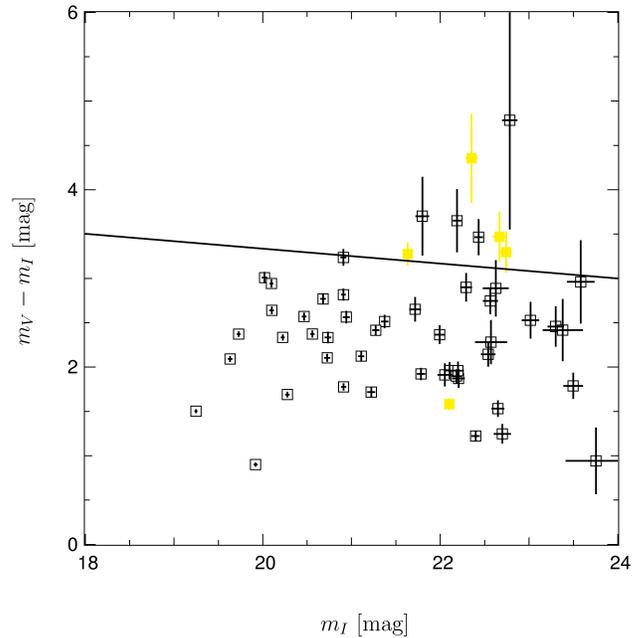


Figure 6.10: CMD of structure 20 in S6F5. Colored (black) squares show the colors and magnitudes of the structure members (all galaxies within $5r_c$). The black line is the red sequence model track at \bar{z} . See Fig. 6.6 for more details.

Structure 10 contains 16 members at a mean redshift of $\bar{z} = 0.589$. It is located to the left of the field at a position of $\alpha \approx 11.974^{\text{h}}$ and $\delta \approx 65.58^\circ$, and is marked in light blue. Fig. 6.7 shows the Voronoi and likelihood maps for $z_{\text{central}} = z_i = 0.60$ and the CMD is given in Fig. 6.8. The Voronoi tessellation shows that structure 10 is in good agreement with typical cluster behavior. 14 of 16 galaxies are lying in overdense cells, leading to a Voronoi class of 1. More overdense cells are lying directly to the left and right. The likelihood classification of the structure also resulted in class 1. The structure center coincides with enhanced likelihood areas in all seven maps. The maximum likelihood reached by the structure in these maps is $\ln L = 2.21$ in the bin with $z_i = 0.64$. The position of the peak of this likelihood enhancement is almost identical to the structure center. Although the CMD of the structure has a relatively large spread, many of the objects are located close to the red sequence or have a larger error, resulting in a CMD classification of 3. Six spectroscopic redshifts are available for this structure. Five of them correspond well with each other, leading to the spectroscopic verification of structure 10. The true mean spectroscopic redshift of this cluster is $\bar{z}_{\text{cluster}} = 0.600$ (see Table 6.11) and the deviation from the EXT-FOF mean redshift is $\bar{z}_{\text{cluster}} - \bar{z} = 0.011$. The lower panel of Fig. 6.3 showed, that another structure occupies this projected position. This structure 24 is located at a mean redshift of $\bar{z} = 1.254$, making it highly unlikely that members of this structure are actually part of cluster 10.

Structure 20 has the highest redshift of the three examples, has five members and is located at a mean redshift of $\bar{z} = 0.953$. It lies in the upper left corner of the field, roughly between the other two example clusters, has a central position of $\alpha \approx 11.975^{\text{h}}$ and $\delta \approx 65.61^{\circ}$, and is marked in yellow. Fig. 6.9 shows the Voronoi and likelihood maps for $z_{\text{central}} = z_i = 0.96$ and the CMD is given in Fig. 6.10. This example demonstrates the problem with using Voronoi tessellations, if the examined structure is located close to the edge. Two of the five galaxies are edge objects, while the remaining three are all lying in consecutive overdense cells, resulting in a class 1 Voronoi classification. Further overdense cells are located to the lower left. The structure also has a likelihood classification of 1. Its center coincides with enhanced likelihood areas in four of the maps. The maximum likelihood reached by structure 20 in these maps is $\ln L = 1.16$ in the bin with $z_i = 0.98$. The position of the likelihood peak is very close to the structure center, yet moved slightly to the lower right. The structure is given a CMD class 3. Most of its members are compatible with the red sequence, except for one rather blue galaxy with only a small color error. No spectra are available for this high redshift candidate and consequently no verification tests can be made.

6.4 Spectroscopic verification

In order to confirm the MUNICS structures, the recipe based on spectroscopic redshifts, that was described in Sect. 5.1 is used.

Of the 162 structures in the catalog a total of 33 (20.4%) possesses at least two galaxies with spectra (see columns two and three in Table 6.10). Only these structures can be used for the confirmation tests, since at least two spectroscopic redshifts are necessary for a verification. 15 of these structures fulfilled the confirmation criteria, corresponding to 45.5% of the structures viable for testing (column five of Table 6.10). Only one (3%) of the EXT-FOF candidates, structure 5 in S6F5, having three members and lying right next to a verified structure, is proven to be spurious (see column six of Table 6.10). All three galaxies have spectroscopic redshifts that are not compatible with each other. For the remaining 17 candidates, i.e. 51.5%, no conclusions about their authenticity can be drawn yet. As described in Sect. 4.3.3, roughly 60% of the EXT-FOF structures are expected to be real, while 40% should be spurious. This analysis shows, that so far no contradictions can be found to the expected rates of real and spurious detections. A detailed field-to-field analysis can be seen in Table 6.10.

Table 6.11 takes a closer look at the subset of confirmed clusters. It lists the field and name of the verified structure (columns one and two), the amount N of members according to the EXT-FOF run (column three), the number of spectra (column four), and the photometric mean redshift \bar{z} (column five). The spectroscopic mean redshifts \bar{z}_{cluster} of the real spectroscopic clusters, and the number N_{spec} of spectroscopic cluster members are given in columns six and seven. The table shows that the 15 verified EXT-FOF structures contain in fact a total of 21 spectroscopically real groups or clusters. Twelve of the EXT-FOF candidates seem to possess only one real cluster each, while there are three EXT-FOF structures, that are composed of a number of real clusters.

Field	Structs.	≥ 2 Spec.	Spec. in structs.	Verified	Spurious
S2F1	14	6	68	4	0
S2F5	25	7	23	3	0
S3F5	18	0	0	0	0
S5F1	21	7	34	3	0
S5F5	16	1	2	0	0
S6F1	20	0	1	0	0
S6F5	25	8	32	3	1
S7F5	23	4	13	2	0
Total	162	33 (20.4 %)	173 (32.6 %)	15 (45.5 %)	1 (3.0 %)

Table 6.10: Overview of the spectroscopic verification statistics. In column two the total number of structures per field is given. Column three shows the number of structures that have at least two galaxies with spectroscopically determined redshifts. The number in brackets shows the corresponding fraction of all structures. The amount of structure members having spectroscopic redshifts is given in column four, the percentage in brackets refers to the total amount of galaxy spectra. Column five contains the number of verified EXT-FOF structures, while the number of spurious detections is listed in column six. The fractions in brackets refer to the total number of structures with at least two spectra.

These are the structures 1 and 7 in S2F1 and structure 2 in S5F1. This effect does not come as a surprise, since it was already predicted in Sect. 4.3.3. There it was demonstrated that some of the photometric redshift structures found by the EXT-FOF will consist of several real structures (i.e. category 7). This phenomenon is an intrinsic property of structure finding in photometric redshift datasets. After all, the photometric redshift determination is less precise by almost two orders of magnitude than the actual spread of a real cluster. If the resulting cluster catalog is to be conservative, this effect has to be accepted.

One of these three structures suffering from the resolution problem, the structure 1 in S2F1, deserves a bit more attention. While the other two are composed of only two real clusters, this structure contains five confirmed clusters. It is by far the most member rich structure in the entire MUNICS catalog and consists of a populous cluster at $\bar{z}_{cluster} = 0.110$ with $N_{spec} = 22$ and four less populous structures reaching out to a redshift of $\bar{z}_{cluster} = 0.369$. A comparison with published cluster catalogs showed that this more populous structure at $\bar{z}_{cluster} = 0.110$ is in fact part of the Abell cluster A0412 (Abell et al., 1989) with a mean redshift of 0.107. The right ascension 03:06:15.5 $\approx 3.104^h$ and declination -00:10:27 $\approx -0.174^\circ$ of the Abell cluster puts the central position slightly left of the field S2F5, that is sitting directly to the south of S2F1. Consequently, another part of the Abell cluster can be identified with the confirmed structure 1 in S2F5, having $N_{spec} = 2$ and $\bar{z}_{cluster} = 0.108$. A0412 is classified as medium compact in the Zwicky system (see Table 1.3) and has an Abell richness class of $R = 0$ (Table 1.1).

Field	Structure	N	Spectra	\bar{z}	$\bar{z}_{cluster}$	N_{spec}
S2F1	1	55	43	0.207	0.110	22
					0.149	4
					0.241	7
					0.283	2
	3	7	6	0.328	0.369	3
					0.272	2
					0.355	2
4	4	3	0.355	0.267	2	
				0.396	4	
S2F5	1	5	3	0.232	0.424	2
					0.108	2
	2	9	3	0.340	0.368	2
					0.458	2
	12	4	2	0.458	0.366	2
S5F1	2	25	8	0.254	0.175	2
					0.281	3
					0.296	2
S6F5	5	12	4	0.296	0.156	2
					0.391	2
	10	4	3	0.387	0.391	2
					0.204	2
S7F5	1	8	2	0.268	0.388	3
					0.589	5
					0.600	5
S7F5	11	5	2	0.519	0.219	2
					0.495	2

Table 6.11: Spectroscopically verified structures. Columns one and two list the field and name of the verified structure. In column three the number N of structure members found by the EXT-FOF is given. Column four shows the number of structure members having spectra. Column five contains the photometric mean redshift \bar{z} . Column six lists the spectroscopic mean redshifts $\bar{z}_{cluster}$ of the real spectroscopic clusters. The number N_{spec} of spectroscopic cluster members is shown in column seven.

Even though this low-redshift regime is not the most sensitive area of the MUNICS EXT-FOF algorithm, the Abell cluster is detected. This example shows one small problem: Since the cluster search is done in each field individually, there is no way to combine the two EXT-FOF components of A0412, namely the structures 1 of the fields S2F1 and S2F5 in the purely photometric MUNICS structure catalog. This problem can be solved by an a posteriori combination of said structures, once they are spectroscopically verified and their $\bar{z}_{cluster}$ can be compared. As already explained in Sect. 6.2.2 using one set of linking parameters for all MUNICS fields, and thus circumventing this problem, is not acceptable. On first sight this effect might seem irritating, yet the probability that this problem does show up in the really interesting redshift range $\bar{z} \geq 0.3$ is not very high, simply due to the corresponding small angular extensions of high redshifted clusters.

Fig. 6.11 illustrates the deviations between the photometric mean structure redshifts

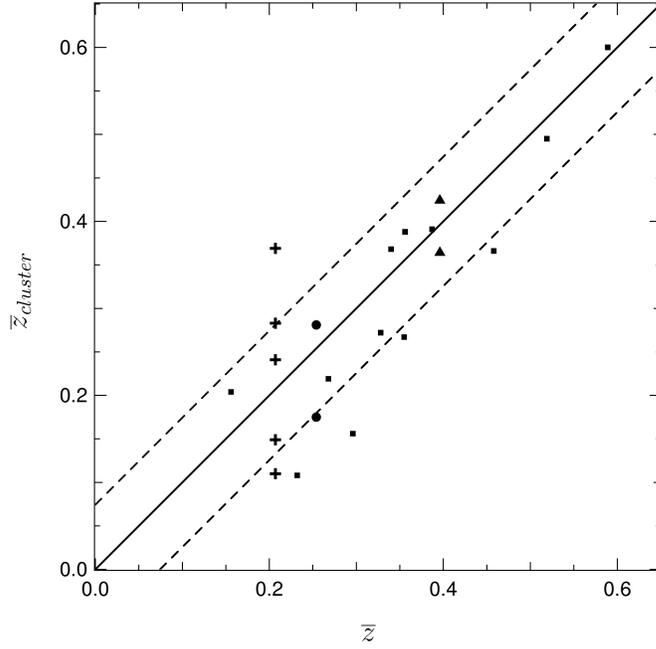


Figure 6.11: Comparison of the photometric mean structure redshift \bar{z} and its spectroscopic counterpart $\bar{z}_{cluster}$ for the verified structures. Squares mark the twelve EXT-FOF structures that contain only one real spectroscopic cluster. The crosses show the splitting of structure 1 in S2F1 into five spectroscopically real components. Triangles mark the two spectroscopic clusters in structure 7 of S2F1, while the circles represent the two real clusters in structure 2 in S5F1. The solid black line is given by $\bar{z}_{cluster} = \bar{z}$, while the dashed lines show the deviation $\sigma_z = 0.074$.

\bar{z} and their spectroscopic counterparts $\bar{z}_{cluster}$ for the confirmed clusters. The three EXT-FOF structures split into multiple real clusters are marked by crosses (structure 1 in S2F1), triangles (structure 7 of S2F1), and circles (structure 2 in S5F1). The twelve EXT-FOF structures that contain only one real spectroscopic cluster are marked by squares.

The standard deviation σ_z between the photometric and spectroscopic mean redshifts of all 21 real spectroscopic clusters is $\sigma_z = 0.074$. If only the eight spectroscopic clusters with $N_{spec} \geq 3$ are used, the deviation is only marginally smaller with $\sigma_z = 0.073$. Both deviations are determined within a redshift range of $0.1 \leq \bar{z}_{cluster} \leq 0.6$. Considering the typical photometric redshift errors of the galaxies, the level of agreement between the photometric and spectroscopic mean redshifts is remarkable. For comparison, the results of a spectroscopic analysis by [Holden et al. \(1999\)](#) of a subset of 16 of the PDCS ([Postman et al., 1996](#)) clusters are shown: Note, the technique used for finding the PDCS structures is the matched filter algorithm (see Sect. 5.3.2 for more details). [Holden et al. \(1999\)](#) used a confirmation criterion of having at least two spectroscopic redshifts within 1500 km s^{-1} of the median of a structure's spectro-

scopic distribution. This agrees roughly to a combination of the two criteria introduced in Sect. 5.1. The standard deviation σ_z of the Holden et al. (1999) clusters with at least two galaxies fulfilling the confirmation criterion is $\sigma_z = 0.20$ within $0.1 \leq \bar{z}_{cluster} \leq 0.5$. If only the structures are used that have at least three galaxies in agreement with the criterion, the deviation reduces to $\sigma_z = 0.06$ within $0.1 \leq \bar{z}_{cluster} \leq 0.35$. Obviously, the redshift deviations in the MUNICS EXT-FOF catalog can compete with the deviations of the matched filter PDCS catalog.

Up to now, mainly low-redshift structures can be tested, which is simply due to the small number of high redshift galaxies contained in the spectroscopic sample. A future observing strategy, to improve the test statistics and to collect more data on scientifically relevant objects, will be to concentrate primarily on plausible high redshift cluster candidates, i.e. structures located at redshifts above 0.5, possessing good Voronoi, likelihood and CMD classifications, and if possible more than three members. Taking spectra of these EXT-FOF structure members will certainly improve the test statistics. Once high-redshift structures are confirmed it will be very important to observe as many of their member galaxies as possible. A good starting point might be to select objects having photometric redshifts compatible with $\bar{z}_{cluster}$, that are also lying within an aperture of 2 – 3 core radii of the cluster center. Thus velocity dispersions, virial masses, cluster luminosities, and mass-to-light ratios, as well as richnesses and cluster morphologies could be determined.

Chapter 7

Summary

Clusters of galaxies, that are the most massive collapsed objects as of today, constitute excellent laboratories for the study of various important topics of extragalactic astronomy. For example, the interaction of galaxies with each other, or with the intergalactic material can be examined. Furthermore, the distribution of the dark matter, that dominates the cluster mass, can be determined. It is also of interest to examine the evolution of galactic properties, like the luminosity function, in both high- and low-density environments in order to understand the role of environmental effects. These studies yield valuable information on the formation and evolution of galaxies and their conglomerates.

Galaxy clusters have also been playing a fundamental role in our understanding of structure formation and cosmology. Since the evolution of the cluster mass-function is highly sensitive to cosmological parameters, the type of dark matter and the biasing of dark against baryonic matter (e.g. [Press & Schechter 1974](#); [Eke et al. 1996](#)), a comparison of the observed mass-function with models yields strong constraints on structure formation scenarios and the cosmological paradigm.

With the development of the photometric redshift determination technique (e.g. [Baum 1962](#); [Benítez 2000](#); [Bender et al. 2001](#)), approximate redshifts became available for all galaxies in a photometric multi-band survey, without having to do time-consuming spectroscopic follow-up observations. The relatively large errors in this type of redshift evaluation pose a problem for a reliable cluster identification, though. While photometric redshift galaxy surveys have increased in popularity over the past years, a paucity of techniques suitable for cluster finding in this type of survey does exist. The new structure finding technique introduced in this work, the extended friends-of-friends (EXT-FOF) algorithm, is designed to deal with photometric redshift surveys.

One of these photometric redshift surveys is MUNICS, the Munich Near-Infrared Cluster Survey ([Drory et al., 2001b](#)), a K' band selected galaxy survey with observations in the optical B , V , R , and I , and the near-infrared J and K' passbands. The part of the survey with the best photometric homogeneity and seeing covers an area of roughly 0.27 deg^2 down to an average depth of $m_{K'} \approx 19.1 \text{ mag}$ (50% completeness

limit for point sources; [Snigula et al. 2002](#)). The prime motivation for constructing this survey was the detection of early type galaxies at redshifts $0 \leq z \leq 1.5$ for evolutionary studies, and the identification of groups or clusters of galaxies up to redshifts around unity. The application of the new EXT-FOF method to the MUNICS survey is one of the main topics of this thesis.

In order to find galaxy groups or clusters in any type of dataset, it is crucial to know their typical observable properties, e.g. the spatial distribution and colors of the structure galaxies, or the extension and luminosity of the X-ray emission of the hot intracluster gas. After all, structure finding is always based on the search for either one or more of those characteristic structure features. Chapter 1 gives a summary of the composition, classification, and the most important characteristics of galaxy groups and clusters, in order to aid the reader in understanding the choice of techniques and criteria used for structure finding, and to ease the interpretation of the resulting cluster catalogs. The chapter also contains a short introduction to the method of photometric redshift determination and to the basics of the MUNICS survey.

In Chapter 2 the basics of structure finding in galaxy datasets are explained. The actual identification of galaxy groups or clusters, either in an observed or simulated catalogue, is a non-trivial enterprise. A reliable detection of large-scale structures in an observed galaxy catalogue is even more complicated if the galaxies have only photometric redshifts or, even worse, none at all. Up to now, many different techniques have been created for the purpose of structure finding. All of them have their advantages and disadvantages and are constructed to fit the specific features of their surveys, like the wavelength of the observed radiation, or whether redshifts were determined for all objects, and if so, what type of redshift determination was used. This chapter illustrates the fundamental problems of constructing structure catalogs, and shows how the features of galaxy surveys influence the choice of the cluster finding technique. It furthermore includes a list of the most common and recent algorithms and their fundamental working principles.

Chapter 3 explains the friends-of-friends (FOF; [Huchra & Geller 1982](#)) technique, one of the most common and reliable cluster finding methods, and introduces the newly designed EXT-FOF algorithm, that is a tailor-made solution for dealing with large photometric redshift errors. The FOF was originally designed to find number overdensities in spectroscopic galaxy surveys and has also been modified to look for structures in simulated galaxy datasets. The algorithm is very straightforward and has no need for complicated cluster models. In principle, it detects overdensities in the galaxy distribution, by looking for galaxy pairs, that are closer to one another than a given cut-off separation. Due to the relatively large errors, that are inherent to the photometric redshift determination, this algorithm cannot be applied to photometric redshift surveys, though. Taking the redshift errors into account would inevitably lead to structures that are unreasonably extended along the redshift axis. Consequently, a modified algorithm, the EXT-FOF, based on the original FOF was created. The basic working principle of the EXT-FOF is the limitation to galaxies for structure finding, that are compatible within their redshift errors with given redshift values. This so-called “redshift-slicing” limits the flexibility of the structure finder and thus heavily

reduces the probability of finding unreasonably elongated structures. The EXT-FOF is the major structure finding tool used in this thesis. This chapter also explains the intricacies of choosing suitable linking parameters, a set of cut-off criteria intrinsic to all types of friends-of-friends algorithms.

The testing of this new algorithm is demonstrated in Chapter 4. This is done in two steps. First, the FOF and EXT-FOF are applied to two spectroscopic redshift datasets, the first Center for Astrophysics Redshift Survey (CFA1; [Huchra et al. 1983](#); [Geller & Huchra 1983](#)) and the Las Campanas Redshift Survey (LCRS; [Shectman et al. 1996](#); [Tucker et al. 2000](#)), and the resulting structure catalogs are compared to one another. It is proven, that in this case both algorithms yield almost identical results. Second, the EXT-FOF is applied to the LCRS with simulated photometric redshifts. The cluster catalog is then compared to the structures resulting from the application of the EXT-FOF to the unmodified spectroscopic dataset. It is shown that in case of photometric surveys, the EXT-FOF yields a rather conservative structure catalog, i.e. the recovery rate for finding clusters is very high (almost 93%), while the rate of spurious detections is still acceptable (about 40%). Furthermore, it is shown that all of the discrepancies in the group compositions are well-understood, and bring about the applicability of the EXT-FOF for the photometric redshift surveys. As expected, a general tendency for finding larger groups was discovered in the photometric redshift case.

Chapter 5 provides a number of tools, that allow the optimization of the above mentioned linking parameters, as well as a rough separation of plausible clusters from any spurious detections. These tools include spectroscopic follow-ups of some of the galaxies, and the usage of this information to determine cluster authenticity. This is by far the best, yet also most expensive tool, as far as telescope time is concerned. Another approach is the creation of Voronoi tessellations. This method yields a scale-free measure for local galaxy densities and thus shows clustering tendencies in the galaxy dataset. The huge advantage of this tool is its independency of cluster models. On the other hand, Voronoi tessellations cannot give any information about edge objects, i.e. galaxies lying too close to the geometric borders of the survey. Thus an application of this technique to dense and contiguous surveys is preferred. The calculation of a probability density, the likelihood, for having a cluster at a given coordinate in the dataset is also described. This method derives from Bayesian statistics, is the basis for the matched filter (MF; [Postman et al. 1996](#)) cluster finding technique, and is very sensitive to even weak cluster signals. Since the determination of this likelihood is based on a comparison of the observed data with cluster models, the results are biased against atypical structures, though. Another method to determine the authenticity of the structure catalog is the creation of color-magnitude diagrams (CMDs). They illustrate the presence of a “red sequence” in clusters, consisting of early type galaxies, that make up about (50 – 80)% of a cluster’s galaxy population. On the one hand, red sequences seem to be very universal, on the other hand, the red sequence is very difficult to detect, if a structure consists only of a handful of members. This collection of methods proves to be a highly valuable tool for authenticity testing, due to the intrinsic strengths and disadvantages of each technique that complement each other.

Chapter 6 deals with the application of the EXT-FOF algorithm to the MUNICS

survey. First, the input galaxy dataset is discussed in detail. Using the subset of MUNICS objects having spectra (about 11% of the objects), the criteria used for classifying objects as galaxies are explained and proven to be valid. After explaining the set of formulae used for the run of the MUNICS EXT-FOF algorithm, the selection process for the optimized linking parameters is demonstrated. Finally, the MUNICS structure catalog is presented, containing a total of 162 structures in an area of roughly 0.27 deg^2 , with mean redshifts $0.13 \leq \bar{z} \leq 1.34$. This corresponds to a mean structure number density of $6.5 \cdot 10^{-4} h^3 \text{ Mpc}^{-3}$ within the most sensitive redshift range for MUNICS structures $0.3 \leq \bar{z} \leq 0.8$, and a flat cosmology with $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. Considering the above mentioned 40% of spurious detections this density reduces to $3.9 \cdot 10^{-4} h^3 \text{ Mpc}^{-3}$. This value agrees well with the typical group densities of $(10^{-5} - 10^{-3}) h^3 \text{ Mpc}^{-3}$, and is higher than the characteristic cluster densities of $(10^{-6} - 10^{-5}) h^3 \text{ Mpc}^{-3}$ (Bahcall in Dekel & Ostriker 1999). That does not come as a surprise, since being a conservative structure finder, the MUNICS EXT-FOF was tuned to find all types of structures, both groups and clusters. One of the eight disconnected fields of the MUNICS survey, the field S6F5, is used as an example to describe the quality assessment of the individual clusters and a detailed look is put on a few structures of S6F5, showing their Voronoi tessellation and likelihood maps, as well as their CMDs. The few available spectroscopic redshifts are used for the verification of the structures, which is aggravated by the predominantly low redshifts of the spectroscopic sample. Of the 33 structures having enough spectra to try a verification, 15 (45.5%) can be confirmed, only one (3.0%) is proven spurious. For the remaining 17 candidates (51.5%), no conclusions about their authenticity can be drawn yet. This analysis shows, that so far no contradictions can be found to the above mentioned expected rates of real (60%) and spurious (40%) detections. It is also demonstrated that some of the photometric redshift structures found by the EXT-FOF consist of several real spectroscopic structures, which agrees well with the results found by the EXT-FOF testing in Chapter 4. A comparison with published cluster catalogs showed that one of the verified structures is in fact part of the Abell cluster A0412 (Abell et al., 1989) with a mean redshift of 0.107, even though this low-redshift regime is not the most sensitive area of the MUNICS EXT-FOF algorithm. For the subset of confirmed clusters, the standard deviation σ_z between the photometric mean structure redshifts and their spectroscopic counterparts $\bar{z}_{cluster}$ is determined to be $\sigma_z = 0.074$ within a redshift range of $0.1 \leq \bar{z}_{cluster} \leq 0.6$. Considering the typical photometric redshift errors of the galaxies of $\sigma(\delta z) = 0.055(1+z)$, the level of agreement between the photometric and spectroscopic mean structure redshifts is remarkable. For comparison, the standard deviation σ_z between the MF algorithm's estimated redshifts and the spectroscopic redshifts measured by Holden et al. (1999) in a subset of the Palomar Distant Cluster Survey (PDCS; Postman et al. 1996) is $\sigma_z = 0.20$ within $0.1 \leq \bar{z}_{cluster} \leq 0.5$. The redshift deviations in the MUNICS EXT-FOF catalog can obviously compete with the deviations of the PDCS matched filter catalog.

In total, it can be said that the newly created EXT-FOF algorithm is a very efficient tool for structure finding, capable of detecting real structures in photometric redshift datasets. The algorithm itself is well understood and rather straightforward.

It yields almost identical results to its older sibling, the FOF, in case of spectroscopic datasets, and gives a conservative catalog for photometric redshift datasets. In contrast to Voronoi based cluster finders, the EXT-FOF can also be applied to sparsely sampled galaxy surveys. Unlike any matched filter or likelihood based techniques, the EXT-FOF does not include cluster models, and thus is relatively unbiased towards atypical structures. All of these properties make the EXT-FOF algorithm a valuable addition to the family of structure finding methods. The resulting MUNICS cluster catalog complies well with typical or expected structure properties. The precision of the structure redshift determination is remarkable, and can compete well with the redshift estimation of the matched filter algorithm.

A future observing strategy to improve the MUNICS cluster catalog and to collect more data on scientifically relevant objects, will be to concentrate primarily on plausible high redshift cluster candidates. Taking spectra of these EXT-FOF structure members will certainly improve the verification statistics. Once high-redshift structures are confirmed it will be very important to observe as many of their member galaxies as possible. A good starting point might be to select objects having photometric redshifts compatible with the spectroscopic cluster redshift, that are also lying within an aperture of some core radii of the cluster center. Thus velocity dispersions, virial masses, cluster luminosities, and mass-to-light ratios, as well as richnesses and cluster morphologies could be determined. Furthermore, the EXT-FOF can be applied to other photometric redshift surveys, like the Fors Deep Field (FDF; [Heidt et al. 2003](#)).

Bibliography

- Abell, G. O. 1958, *ApJS*, 3, 211 [2](#), [14](#)
- Abell, G. O., Corwin, H. G., & Olowin, R. P. 1989, *ApJS*, 70, 1 [15](#), [97](#), [104](#)
- Annis, J., et al. 1999, *AAS*, 31, 1391 [67](#)
- Aragon-Salamanca, A., Ellis, R. S., Couch, W. J., & Carter, D. 1993, *MNRAS*, 262, 764 [67](#)
- Böhringer, H., et al. 2001, *A&A*, 369, 826 [61](#)
- Bahcall, N. A. 1975, *ApJ*, 198, 249 [3](#), [4](#)
- Bahcall, N. A. 1977, *ARA&A*, 15, 505 [3](#), [4](#), [6](#), [7](#), [26](#)
- Bahcall, N. A. 1981, *ApJ*, 247, 787 [4](#)
- Bahcall, N. A., & Cen, R. 1992, *ApJ*, 398, L81 [2](#)
- Bahcall, N. A., & Fan, X. 1998, *ApJ*, 504, 1 [84](#)
- Bahcall, N. A., Fan, X., & Cen, R. 1997, *ApJ*, 485, L53 [2](#), [84](#)
- Bahcall, N. A., Lubin, L. M., & Dorman, V. 1995, *ApJ*, 447, L81 [5](#)
- Barrientos, L. F. 1999, Ph.D. thesis, University of Toronto [67](#)
- Basilakos, S., Plionis, M., & Maddox, S. J. 2000, *MNRAS*, 316, 779 [57](#)
- Baum, W. A. 1959, *PASP*, 71, 106 [67](#)
- Baum, W. A. 1962, in *IAU Symp. 15: Problems of Extra-Galactic Research*, Vol. 15, 390 [8](#), [101](#)
- Bautz, L. P., & Morgan, W. W. 1970, *ApJ*, 162, L149 [7](#)
- Bayes, T. 1763, *Phil. Trans. Roy. Soc. London*, 53, 370 [59](#)
- Bender, R., et al. 2001, in *ESO/ECF/STSci Workshop on Deep Fields*, 327 [8](#), [72](#), [101](#)
- Bender, R., Saglia, R. P., Ziegler, B., Belloni, P., Greggio, L., Hopp, U., & Bruzual, G. 1998, *ApJ*, 493, 529 [67](#)

- Benítez, N. 2000, *ApJ*, 536, 571 8, 101
- Benítez, N., Broadhurst, T., Rosati, P., Courbin, F., Squires, G., Lidman, C., & Magain, P. 1999, *ApJ*, 527, 31 68
- Bevington, P. R., & Robinson, D. K. 2003, *Data Reduction and Error Analysis for the Physical Sciences* (Boston: McGraw-Hill) 60
- Binney, J., & Merrifield, M. 1998, *Galactic Astronomy* (Princeton: Princeton University Press) 51
- Bode, P., Bahcall, N. A., Ford, E. B., & Ostriker, J. P. 2001, *ApJ*, 551, 15 2
- Botzler, C. S. 1999, Diploma thesis, Ludwig-Maximilian University, Munich 2, 84
- Botzler, C. S., et al. 2003a, *Astrophysics and Space Science*, 284, 393
- Botzler, C. S., Snigula, J., Bender, R., & Hopp, U. 2003b, *MNRAS*, accepted
- Botzler, C. S., Snigula, J., Bender, R., Hopp, U., & Goranova, Y. 2004, *MNRAS*, in preparation
- Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992a, *MNRAS*, 254, 601 67
- Bower, R. G., Lucey, J. R., & Ellis, R. S. 1992b, *MNRAS*, 254, 589 67
- Brunner, R. J., Connolly, A. J., & Szalay, A. S. 1999, *ApJ*, 516, 563 8
- Butcher, H., & Oemler, A. 1984, *ApJ*, 285, 426 1
- Carrol, S. M., Press, W. H., & Turner, E. L. 1992, *ARA&A*, 30, 499 22, 42
- Cole, S., & Lacey, C. 1996, *MNRAS*, 281, 716 15, 17
- Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, *ApJS*, 43, 393 69
- Coles, P. 1991, *Nature*, 349, 288 53
- Couch, W. J., Ellis, R. S., MacLaren, I., & Malin, D. F. 1991, *MNRAS*, 249, 606 15
- Croft, R. A. C., & Efstathiou, G. 1994, *MNRAS*, 267, 390 27
- da Costa, L., et al. 1999, *A&A*, 343, L29 68
- Dalton, G. B., Maddox, S. J., Sutherland, W. J., & Efstathiou, G. 1997, *MNRAS*, 289, 263 15
- Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, *ApJ*, 292, 371 17, 27
- Davis, M., Newman, J. A., Faber, S. M., & Phillips, A. C. 2001, in *ESO/ECF/STScI Workshop on Deep Fields*, 241 57
- de Propris, R., Eisenhardt, P. R., Stanford, S. A., & Dickinson, M. 1998, *ApJ*, 503, L45 65

- de Propriis, R., Stanford, S. A., Eisenhardt, P. R., Dickinson, M., & Elston, R. 1999, *AJ*, 118, 719 [1](#), [65](#), [66](#)
- Dekel, A., & Ostriker, J. P., ed. 1999, *Formation of Structure in the Universe* [4](#), [5](#), [6](#), [7](#), [27](#), [52](#), [104](#)
- Delaunay, B. 1934, *Bull. Acad. Sci. USSR (VII)*, 793 [54](#)
- Dickinson, M. 1995, in *ASP Conf. Ser. 86: Fresh Views of Elliptical Galaxies*, 283 [68](#)
- Dirichlet, G. 1850, *J. reine angew. Math.* 40, 216 [54](#)
- Donahue, M., et al. 2002, *ApJ*, 569, 689 [60](#)
- Doroshkevich, A. G., Gottlober, S., & Madsen, S. 1997, *A&AS*, 123, 495 [54](#)
- Dressler, A. 1978, *ApJ*, 226, 55 [3](#), [4](#)
- Dressler, A. 1980, *ApJ*, 236, 351 [1](#)
- Drory, N. 2003, *A&A*, 397, 371 [9](#), [10](#), [72](#)
- Drory, N., Bender, R., Feulner, G., Hopp, U., Maraston, C., Snigula, J., & Hill, G. J. 2003, *ApJ*, 595, 698 [9](#), [10](#), [11](#), [71](#), [80](#)
- Drory, N., Bender, R., Snigula, J., Feulner, G., Hopp, U., Maraston, C., Hill, G. J., & de Oliveira, C. M. 2001a, *ApJ*, 562, L111 [1](#)
- Drory, N., Feulner, G., Bender, R., Botzler, C. S., Hopp, U., Maraston, C., Mendes de Oliveira, C., & Snigula, J. 2001b, *MNRAS*, 325, 550 [8](#), [71](#), [72](#), [101](#)
- Ebeling, H., & Wiedenmann, G. 1993, *Ph. Rev. E*, 47, 704 [53](#)
- Efstathiou, G., Frenk, C. S., White, S. D. M., & Davis, M. 1988, *MNRAS*, 235, 715 [17](#)
- Eisenstein, D. J., & Hut, P. 1998, *ApJ*, 498, 137 [16](#)
- Eke, V. R., Cole, S., & Frenk, C. S. 1996, *MNRAS*, 282, 263 [1](#), [101](#)
- El-Ad, H., & Piran, T. 1997, *ApJ*, 491, 421 [54](#)
- El-Ad, H., Piran, T., & da Costa, L. N. 1996, *ApJ*, 462, L13 [54](#)
- Ellis, R. S., Smail, I., Dressler, A., Couch, W. J., Oemler, A. J., Butcher, H., & Sharples, R. M. 1997, *ApJ*, 483, 582 [67](#)
- Fernández-Soto, A., Lanzetta, K. M., & Yahil, A. 1999, *ApJ*, 513, 34 [8](#)
- Feulner, G., Bender, R., Drory, N., Hopp, U., Snigula, J., & Hill, G. J. 2003, *MNRAS*, 342, 605 [1](#), [9](#), [53](#), [71](#), [72](#)
- Fisher, D., Fabricant, D., Franx, M., & van Dokkum, P. 1998, *ApJ*, 498, 195 [52](#)

- Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. 1988, *ApJ*, 327, 507 [27](#)
- Fried, J. W., et al. 2001, *A&A*, 367, 788 [1](#)
- Geller, M. J. 1990, in *STScI Symposium Series: Clusters of Galaxies*, 25 [3](#)
- Geller, M. J., & Huchra, J. P. 1983, *ApJS*, 52, 61 [17](#), [19](#), [30](#), [31](#), [32](#), [103](#)
- Giuricin, G., Marinoni, C., Ceriani, L., & Pisani, A. 2000, *ApJ*, 543, 178 [17](#)
- Gladders, M. 2002, Ph.D. thesis, University of Toronto [68](#)
- Gladders, M. D., Lopez-Cruz, O., Yee, H. K. C., & Kodama, T. 1998, *ApJ*, 501, 571 [67](#)
- Gladders, M. D., & Yee, H. K. C. 2000, *AJ*, 120, 2148 [16](#), [67](#), [68](#), [69](#)
- Goldwirth, D. S., da Costa, L. N., & van de Weygaert, R. 1995, *MNRAS*, 275, 1185 [53](#)
- Goto, T., et al. 2002, *AJ*, 123, 1807 [16](#), [67](#), [68](#)
- Gourgoulhon, E., Chamaroux, P., & Fouque, P. 1992, *A&A*, 255, 69 [15](#)
- Governato, F., Babul, A., Quinn, T., Tozzi, P., Baugh, C. M., Katz, N., & Lake, G. 1999, *MNRAS*, 307, 949 [16](#)
- Heidt, J., et al. 2003, *A&A*, 398, 49 [8](#), [105](#)
- Henry, P., & Tucker, W. 1979, *ApJ*, 229, 78 [1](#)
- Holden, B. P., Nichol, R. C., Romer, A. K., Metevier, A., Postman, M., Ulmer, M. P., & Lubin, L. M. 1999, *AJ*, 118, 2002 [99](#), [100](#), [104](#)
- Huchra, J., Davis, M., Latham, D., & Tonry, J. 1983, *ApJS*, 52, 89 [19](#), [29](#), [103](#)
- Huchra, J. P., & Geller, M. J. 1982, *ApJ*, 257, 423 [10](#), [17](#), [19](#), [102](#)
- Icke, V., & van de Weygaert, R. 1987, *A&A*, 184, 16 [53](#)
- Ikeuchi, S., & Turner, E. L. 1991, *MNRAS*, 250, 519 [53](#)
- Jing, Y., & Fang, L. 1994, *ApJ*, 432, 438 [2](#)
- Jones, C., & Forman, W. 1992, in *NATO ASIC Proc. 366: Clusters and Superclusters of Galaxies*, 49 [5](#)
- Jørgensen, I., Franx, M., Hjorth, J., & van Dokkum, P. G. 1999, *MNRAS*, 308, 833 [67](#)
- Kawasaki, W., Shimasaku, K., Doi, M., & Okamura, S. 1998, *A&AS*, 130, 567 [60](#)
- Kelson, D. D., van Dokkum, P. G., Franx, M., Illingworth, G. D., & Fabricant, D. 1997, *ApJ*, 478, L13 [67](#)

- Kepner, J., Fan, X., Bahcall, N., Gunn, J., Lupton, R., & Xu, G. 1999, *ApJ*, 517, 78
15, 60, 61, 62, 64, 65
- Kiang, T. 1966, *Zeitschrift für Astrophysik*, 64, 433 53
- Kim, R. S. J., et al. 2002, *AJ*, 123, 20 16, 54, 57, 58, 60, 67, 68
- King, I. R. 1972, *ApJ*, 174, L123 3
- Kodama, T. 1997, Ph.D. thesis, University of Tokyo 69
- Kodama, T., Arimoto, N., Barger, A. J., & Aragon-Salamanca, A. 1998, *A&A*, 334,
99 67
- Koo, D. C. 1985, *AJ*, 90, 418 8
- Lacey, C., & Cole, S. 1994, *MNRAS*, 271, 676 15, 17, 27
- Lidman, C. E., & Peterson, B. A. 1996, *AJ*, 112, 2454 15
- Lin, H., Kirshner, R. P., Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., &
Schechter, P. L. 1996, *ApJ*, 464, 60 43
- Lineweaver, C. H., Tenorio, L., Smoot, G. F., Keegstra, P., Banday, A. J., & Lubin, P.
1996, *ApJ*, 470, 38 41
- Ling, E. N. 1987, Ph.D. thesis, Sussex University, Brighton 53
- Lobo, C., Iovino, A., Lazzati, D., & Chincarini, G. 2000, *A&A*, 360, 896 15, 60
- Longair, M. S., ed. 1998, *Galaxy Formation* 4, 6, 7
- Lopez-Cruz, O., & Yee, H. K. C. 2000, *ApJ*, submitted 67
- Loredo, T. J. 1995, Ph.D. thesis 60
- Lubin, L. M., & Bahcall, N. A. 1993, *ApJ*, 415, L17 1
- Lubin, L. M., Cen, R., Bahcall, N. A., & Ostriker, J. P. 1996, *ApJ*, 460, 10 5
- Lugger, P. M. 1984, *ApJ*, 278, 51 67
- Lupton, R. 1993, *Statistics in Theory and Practice* (Princeton: Princeton University
Press) 60
- Marinoni, C., Davis, M., Newman, J. A., & Coil, A. L. 2002, *ApJ*, 580, 122 16, 54,
57
- Martínez, V. J., & Saar, E. 2002, *Statistics of the Galaxy Distribution* (Chapman &
Hall/CRC) 54
- Marzke, R., et al. 1999, in *ASP Conf. Ser. 191: Photometric Redshifts and the Detec-
tion of High Redshift Galaxies*, 148 8

- Materne, J. 1978, *A&A*, 63, 401 15
- Matsuda, T., & Shima, E. 1984, *Progress of Theoretical Physics*, 71, 855 53
- McCarthy, P., et al. 2001b, in *ESO/ECF/STScI Workshop on Deep Fields*, 247
- McCarthy, P. J., et al. 2001a, *ApJ*, 560, L131
- Mellier, Y., Soucail, G., Fort, B., & Mathez, G. 1988, *A&A*, 199, 13 52
- Merchán, M. E., Maia, M. A. G., & Lambas, D. G. 2000, *ApJ*, 545, 26 17
- Metcalf, N., Godwin, J. G., & Peach, J. V. 1994, *MNRAS*, 267, 431 67
- Meurs, E. J. A., & Wilkinson, M. I. 1999, in *ASP Conf. Ser. 176: Observational Cosmology*, 333 54
- Nilson, P. 1973, *Uppsala General Catalogue of Galaxies (Uppsala: Astronomiska Observatorium)* 30
- Oke, J. B., Postman, M., & Lubin, L. M. 1998, *AJ*, 116, 549 67
- Olsen, L. F., et al. 1999, *A&A*, 345, 681 60
- Openshaw, S. 1984, in *Spatial Analysis and GIS*, ed. S. Fotheringham & P. Rogerson (London: Taylor and Francis), 83 57
- Paredes, S., Jones, B. J. T., & Martinez, V. J. 1995, *MNRAS*, 276, 1116 15
- Peebles, P. J. E. 1993, *Principles of Physical Cosmology (Princeton: Princeton University Press)* 1
- Plionis, M., Barrow, J. D., & Frenk, C. S. 1991, *MNRAS*, 249, 662 15, 57
- Postman, M., Lauer, T. R., Oegerle, W., & Donahue, M. 2002, *ApJ*, 579, 93 60
- Postman, M., Lubin, L. M., Gunn, J. E., Oke, J. B., Hoessel, J. G., Schneider, D. P., & Christensen, J. A. 1996, *AJ*, 111, 615 15, 56, 60, 61, 63, 64, 86, 99, 103, 104, 118
- Postman, M., Lubin, L. M., & Oke, J. B. 2001, *AJ*, 122, 1125 1
- Press, W. H., & Schechter, P. 1974, *ApJ*, 187, 425 1, 2, 101
- Quintana, H., & de Souza, R. 1993, *A&AS*, 101, 475 52
- Rakos, K. D., & Schombert, J. M. 1995, *ApJ*, 439, 47 67
- Ramella, M., Boschin, W., Fadda, D., & Nonino, M. 2001, *A&A*, 368, 776 16, 54, 56, 58, 59
- Ramella, M., Geller, M. J., & Huchra, J. P. 1989, *ApJ*, 344, 57 17
- Ramella, M., Geller, M. J., Pisani, A., & da Costa, L. N. 2002, *AJ*, 123, 2976 17

- Ramella, M., Pisani, A., & Geller, M. J. 1997, *AJ*, 113, 483 17, 30
- Rood, H. J., & Sastry, G. N. 1971, *PASP*, 83, 313 7
- Ryden, B. S. 1995, *ApJ*, 452, 25 54
- Schuecker, P., & Boehringer, H. 1998, *A&A*, 339, 315 15, 60, 61, 64
- Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., Lin, H., Kirshner, R. P., & Schechter, P. L. 1996, *ApJ*, 470, 172 40, 103
- Shewchuk, J. 1996, in *First Workshop on Applied Computational Geometry*, ACM (<http://www.cs.cmu.edu/~quake/triangle.html>) 59
- Slezak, E., Bijaoui, A., & Mars, G. 1990, *A&A*, 227, 301 15
- Smail, I., Edge, A. C., Ellis, R. S., & Blandford, R. D. 1998, *MNRAS*, 293, 124 67
- Snigula, J., Drory, N., Bender, R., Botzler, C. S., Feulner, G., & Hopp, U. 2002, *MNRAS*, 336, 1329 8, 9, 71, 76, 102
- Stanford, S. A., Eisenhardt, P. R., & Dickinson, M. 1998, *ApJ*, 492, 461 67
- Stanford, S. A., Elston, R., Eisenhardt, P. R., Spinrad, H., Stern, D., & Dey, A. 1997, *AJ*, 114, 2232 68
- Stoughton, C., Annis, J., Tucker, D. L., Hashimoto, Y., McKay, T. A., & Smith, J. A. 1998, *AAS*, 30, 888 67
- Struble, M. F., & Ftaclas, C. 1994, *AJ*, 108, 1 1, 57
- Struble, M. F., & Rood, H. J. 1991, *ApJS*, 77, 363 4, 52
- Sunyaev, R. A., & Zeldovich, Y. B. 1972, *Comments on Astrophysics*, 4, 173 5
- Trasarti-Battistoni, R. 1998, *A&AS*, 130, 341 17
- Trentham, N. 1998, *MNRAS*, 294, 193 1
- Tucker, D. L., et al. 1997, *MNRAS*, 285, L5 43
- Tucker, D. L., et al. 2000, *ApJS*, 130, 237 17, 41, 44, 103
- Tully, R. B. 1980, *ApJ*, 237, 390 15
- Tully, R. B. 1987, *ApJ*, 321, 280 15
- Turner, E. L., & Gott, J. R. 1976, *ApJS*, 32, 409 15
- Ueda, H., Itoh, M., & Suto, Y. 1993, *ApJ*, 408, 3 2
- Valageas, P., Lacey, C., & Schaeffer, R. 2000, *MNRAS*, 311, 234 17
- van de Weygaert, R. 1994, *A&A*, 283, 361 53, 54

- van de Weygaert, R., & Icke, V. 1989, *A&A*, 213, 1 53
- van Dokkum, P. G., & Franx, M. 1996, *MNRAS*, 281, 985 67
- van Dokkum, P. G., Franx, M., Kelson, D. D., & Illingworth, G. D. 1998, *ApJ*, 504, L17 67
- Visvanathan, N., & Sandage, A. 1977, *ApJ*, 216, 214 67
- Voronoi, G. 1908, *J. reine angew. Math.* 134, 198 54
- Wang, Q. D., & Ulmer, M. P. 1997, *MNRAS*, 292, 920 57
- White, D. A., & Fabian, A. C. 1995, *MNRAS*, 273, 72 5
- White, S. D. M., Navarro, J. F., Evrard, A. E., & Frenk, C. S. 1993, *Nature*, 366, 429 5
- Williams, R. E., et al. 2000, *AJ*, 120, 2735 8
- Williams, R. E., et al. 1996, *AJ*, 112, 1335 8
- Wolf, C., Dye, S., Kleinheinrich, M., Meisenheimer, K., Rix, H.-W., & Wisotzki, L. 2001b, *A&A*, 377, 442
- Wolf, C., et al. 2001a, *A&A*, 365, 681
- Yee, H. K. C., & Gladders, M. D. 2002, in *ASP Conf. Ser. 257: High-Z Clusters, Missing Baryons, and CMB Polarization*, 109 67, 68
- York, D. G., et al. 2000, *AJ*, 120, 1579 57
- Yoshioka, S., & Ikeuchi, S. 1989, *ApJ*, 341, 16 53
- Zaninetti, L. 1989, *A&A*, 224, 345 53
- Zaninetti, L. 1990, *A&A*, 233, 293 53
- Zaninetti, L. 1991, *A&A*, 246, 291 53
- Zaninetti, L. 1993, *A&A*, 276, 255 53
- Zaninetti, L. 1995, *A&AS*, 109, 71 53
- Zwicky, F. 1957, *Morphological Astronomy* (Berlin: Springer) 3
- Zwicky, F., Herzog, E., & Wild, P. 1961-1968, *Catalogue of Galaxies and of Clusters of Galaxies*, 6 volumes (Pasadena: California Institute of Technology) 6, 15, 29
- Zwicky, F., & Zwicky, M. A. 1971, *Catalogue of Selected Compact Galaxies and of Post-Eruptive Galaxies* (Guemligen) 29

Appendix A

Maximizing the Likelihood

Let $D_m(x, y, m)$ be a model for the distribution of galaxies

$$D_m(x, y, m) := b(m) + \Lambda_i s(x, y, x_i, y_i) \phi(m, m^*(z_i)), \quad (\text{A.1})$$

where $b(m)$ is the background distribution of galaxies, Λ_i is a measure for the richness of the structure i , and (x_i, y_i, z_i) are its central coordinates in redshift space. The term $s(x, y, x_i, y_i)$ describes the projected density of galaxies belonging to the structure at the position (x, y) , and $\phi(m, m^*(z_i))$ is the differential cluster luminosity function.

The observed galaxy distribution $D_d(x, y, m)$ can be described by a set of δ -functions

$$D_d(x, y, m) := \sum_{j=1}^N \delta(x_j, y_j, m_j), \quad (\text{A.2})$$

where N is the total number of galaxies.

Assuming that the dataset contains enough background galaxies, a Gaussian distribution can be used to determine the conditional probability that the observed dataset can be described by the given model. In the case of MUNICS this is only a rough approximation, yet still sufficiently accurate to locate the peaks in the likelihood.

$$P(x, y, m) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{[D_m(x, y, m) - D_d(x, y, m)]^2}{2\sigma^2}\right], \quad (\text{A.3})$$

with

$$\sigma^2 = b(m). \quad (\text{A.4})$$

The likelihood L is given by

$$L := \prod_{x,y,m} P(x,y,m). \quad (\text{A.5})$$

Thus the logarithmic likelihood $\ln L$ can be written as

$$\begin{aligned} -2 \ln L &= -2 \ln \prod_{x,y,m} P(x,y,m) = \\ &= -2 \sum_{x,y,m} \ln P(x,y,m) = \\ &= 2 \sum_{x,y,m} \left[\ln(\sqrt{2\pi}\sigma) + \frac{[D_m(x,y,m) - D_d(x,y,m)]^2}{2\sigma^2} \right] \rightarrow \\ &\rightarrow 2 \iiint \left[\ln(\sqrt{2\pi}\sigma) + \frac{[D_m(x,y,m) - D_d(x,y,m)]^2}{2\sigma^2} \right] dx dy dm = \\ &= \text{const} + \iiint \frac{[D_m(x,y,m) - D_d(x,y,m)]^2}{\sigma^2} dx dy dm, \end{aligned} \quad (\text{A.6})$$

where the integration covers the entire modeled area and magnitude space.

The second expression on the right hand side of eq. (A.6) can be expressed as

$$\begin{aligned} \iiint \frac{[D_m(x,y,m) - D_d(x,y,m)]^2}{\sigma^2} dx dy dm &= \\ &= \iiint \frac{D_m^2(x,y,m)}{\sigma^2} dx dy dm + \iiint \frac{D_d^2(x,y,m)}{\sigma^2} dx dy dm \\ &\quad - 2 \iiint \frac{D_m(x,y,m) D_d(x,y,m)}{\sigma^2} dx dy dm, \end{aligned} \quad (\text{A.7})$$

with

$$\begin{aligned} \iiint \frac{D_m^2(x,y,m)}{\sigma^2} dx dy dm &= \iiint \sigma^2 dx dy dm \\ &\quad + \iiint \frac{\Lambda_i^2 s^2(x,y,x_i,y_i) \phi^2(m,m^*(z_i))}{\sigma^2} dx dy dm \\ &\quad + 2 \iiint \Lambda_i s(x,y,x_i,y_i) \phi(m,m^*(z_i)) dx dy dm, \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned}
\iiint \frac{D_d^2(x, y, m)}{\sigma^2} dx dy dm &= \\
&= \iiint \frac{\sum_j \delta(x_j, y_j, m_j) \sum_k \delta(x_k, y_k, m_k)}{\sigma^2} dx dy dm = \\
&= \iiint \frac{\sum_j \delta(x_j, y_j, m_j)}{\sigma^2} dx dy dm = \\
&= \iiint \frac{D_d(x, y, m)}{\sigma^2} dx dy dm, \tag{A.9}
\end{aligned}$$

and

$$\begin{aligned}
\iiint \frac{D_m(x, y, m) D_d(x, y, m)}{\sigma^2} dx dy dm &= \\
&= \iiint D_d(x, y, m) dx dy dm \\
&\quad + \iiint \frac{\Lambda_i s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m)}{\sigma^2} dx dy dm. \tag{A.10}
\end{aligned}$$

Using eqs. (A.7), (A.8), (A.9), and (A.10) the logarithmic likelihood is

$$\begin{aligned}
\ln L \propto & \iiint \Lambda_i^2 \frac{s^2(x, y, x_i, y_i) \phi^2(m, m^*(z_i))}{2\sigma^2} dx dy dm \\
& - \iiint \Lambda_i \left[\frac{s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m)}{\sigma^2} \right. \\
& \quad \left. - s(x, y, x_i, y_i) \phi(m, m^*(z_i)) \right] dx dy dm + \text{const.} \tag{A.11}
\end{aligned}$$

The expression for Λ_i is determined by setting $\partial \ln L / \partial \Lambda_i = 0$:

$$\begin{aligned}
\frac{\partial \ln L}{\partial \Lambda_i} \propto & \iiint \Lambda_i \frac{s^2(x, y, x_i, y_i) \phi^2(m, m^*(z_i))}{\sigma^2} dx dy dm \\
& - \iiint \left[\frac{s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m)}{\sigma^2} \right. \\
& \quad \left. - s(x, y, x_i, y_i) \phi(m, m^*(z_i)) \right] dx dy dm = \\
& = 0, \tag{A.12}
\end{aligned}$$

thus

$$\begin{aligned} \Lambda_i = & \iiint \left[\frac{s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m)}{\sigma^2} \right. \\ & \left. - s(x, y, x_i, y_i) \phi(m, m^*(z_i)) \right] dx dy dm \\ & \cdot \left[\iiint \frac{s^2(x, y, x_i, y_i) \phi^2(m, m^*(z_i))}{\sigma^2} dx dy dm \right]^{-1}. \end{aligned} \quad (\text{A.13})$$

Inserting eq. (A.13) into (A.11) and neglecting the constant expressions yields

$$\begin{aligned} \ln L \propto & \left[\iiint \left[\frac{s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m)}{\sigma^2} \right. \right. \\ & \left. \left. - s(x, y, x_i, y_i) \phi(m, m^*(z_i)) \right] dx dy dm \right]^2 \\ & \cdot \left[\iiint \frac{s^2(x, y, x_i, y_i) \phi^2(m, m^*(z_i))}{\sigma^2} dx dy dm \right]^{-1}. \end{aligned} \quad (\text{A.14})$$

According to Postman et al. (1996) maximizing eq. (A.14) is roughly equivalent to only maximizing the first term in the numerator. Thus

$$\begin{aligned} \ln L \propto & \iiint \frac{1}{\sigma^2} s(x, y, x_i, y_i) \phi(m, m^*(z_i)) D_d(x, y, m) dx dy dm = \\ = & \iiint \frac{1}{b(m)} s(x, y, x_i, y_i) \phi(m, m^*(z_i)) \sum_j \delta(x_j, y_j, m_j) dx dy dm = \\ = & \sum_j \frac{1}{b(m_j)} s(x_j, y_j, x_i, y_i) \phi(m_j, m^*(z_i)). \end{aligned} \quad (\text{A.15})$$

Appendix B

MUNICS Structures

B.1 The field S2F1

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	55	03:06:39.10	-00:00:51	0.207	2 (11)	1	5	43	+1
2	6	03:06:52.27	00:01:22	0.314	1	1	5	4	0
3	7	03:06:38.55	-00:00:15	0.328	1	1	4	6	+1
4	4	03:06:35.24	00:05:21	0.355	4	2	3	3	+1
5	4	03:06:26.50	-00:00:50	0.385	1	1	3	1	0
6	3	03:06:22.89	-00:00:18	0.390	2	1	1	1	0
7	11	03:06:32.53	-00:01:50	0.396	1	1	2	6	+1
8	5	03:06:55.06	00:04:34	0.547	2	2	1	1	0
9	3	03:06:38.90	-00:03:25	0.620	2	1	2	0	0
10	3	03:06:24.54	00:05:19	0.651	2	5	1	0	0
11	9	03:06:42.94	-00:04:41	0.724	1 (2)	2	2	1	0
12	4	03:06:34.69	-00:00:52	0.735	1	2	2	0	0
13	4	03:06:54.72	00:05:30	0.757	1 (1)	5	1	0	0
14	3	03:06:40.78	00:03:30	0.799	2	5	3	2	0

Table B.1: Structures in S2F1. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

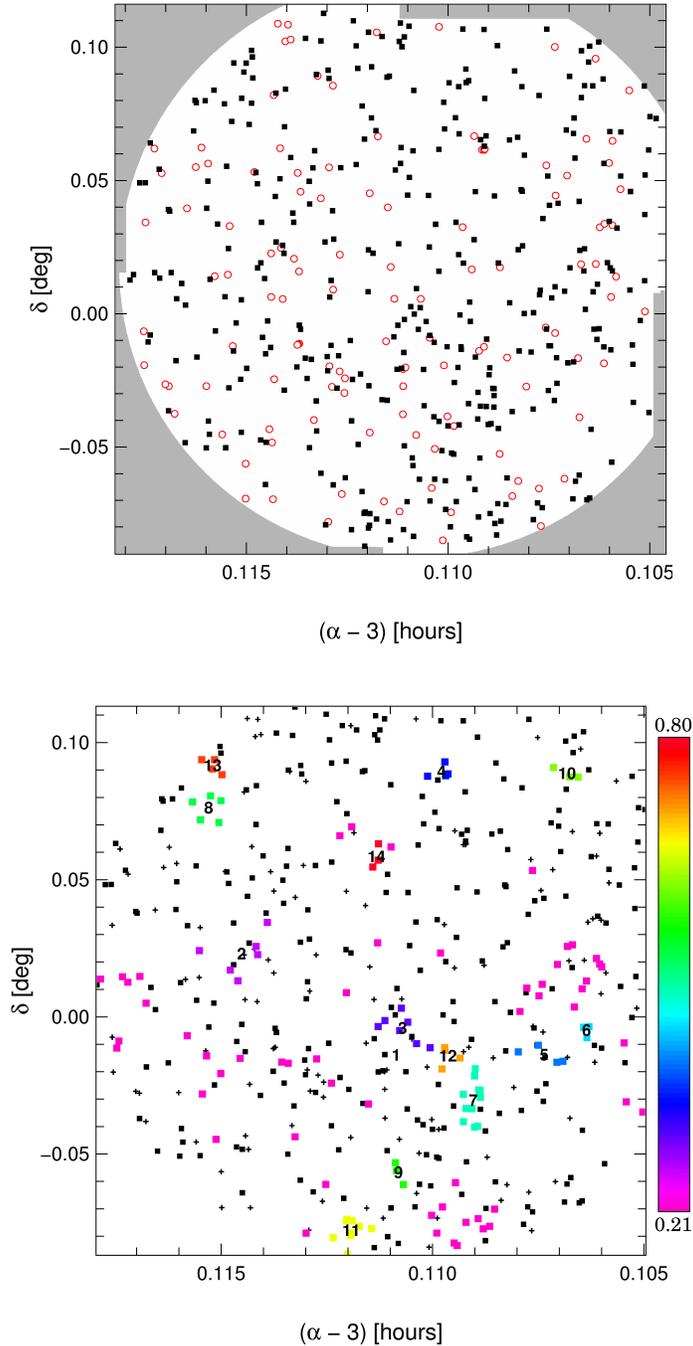


Figure B.1: Geometrical borders and projected distribution of structures in S2F1. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

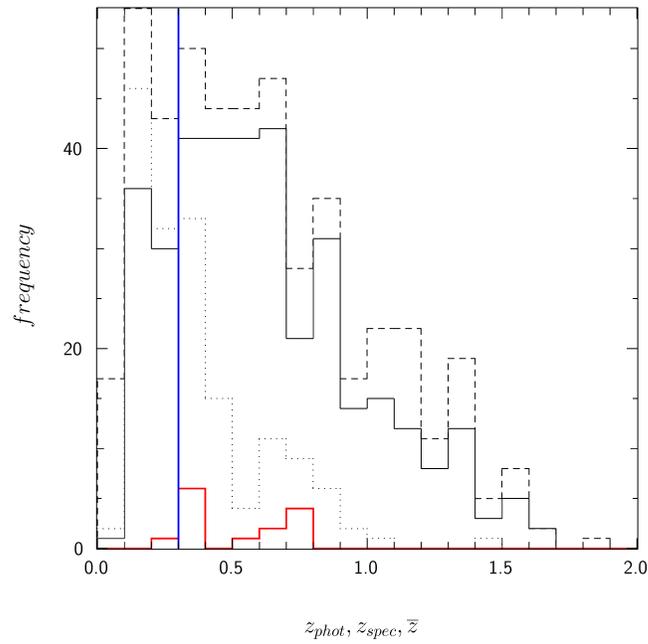


Figure B.2: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S2F1. See upper panel of Fig. 6.2 for more details.

B.2 The field S2F5

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	5	03:06:28.65	-00:08:38	0.232	2 (2)	1	5	3	+1
2	9	03:06:43.58	-00:11:20	0.340	1	1	4	3	+1
3	3	03:06:37.39	-00:09:05	0.348	2	1	2	2	0
4	3	03:06:51.97	-00:16:20	0.349	1	1	3	2	0
5	3	03:06:48.53	-00:12:41	0.361	2	1	5	1	0
6	3	03:06:56.97	-00:10:02	0.395	4	1	2	0	0
7	6	03:06:40.94	-00:16:02	0.407	2	1	4	2	0
8	3	03:06:53.94	-00:15:11	0.413	2	1	5	0	0
9	3	03:06:44.31	-00:14:52	0.421	2	1	3	2	0
10	4	03:06:39.90	-00:17:34	0.445	2	1	5	1	0
11	3	03:06:48.13	-00:17:14	0.452	2	1	1	0	0
12	4	03:06:56.06	-00:11:36	0.458	2	1	3	2	+1
13	4	03:06:49.08	-00:14:26	0.461	3	1	5	1	0
14	18	03:06:35.23	-00:18:17	0.467	1 (3)	1	2	0	0
15	6	03:06:45.45	-00:13:24	0.469	1	1	3	1	0
16	5	03:06:53.88	-00:18:04	0.473	3 (1)	1	2	1	0
17	4	03:06:44.09	-00:16:58	0.516	1	1	2	0	0
18	3	03:06:51.00	-00:13:25	0.557	2	1	2	0	0
19	3	03:06:24.11	-00:15:20	0.609	4	5	3	1	0
20	6	03:06:34.15	-00:14:48	0.669	3	1	1	0	0
21	3	03:06:45.14	-00:14:34	0.679	2	1	1	1	0
22	3	03:06:35.90	-00:13:37	0.703	1	1	2	0	0
23	4	03:06:58.92	-00:10:03	0.793	1	5	2	0	0
24	3	03:06:55.24	-00:09:47	0.940	2	5	1	0	0
25	3	03:06:57.26	-00:13:56	0.959	1	5	2	0	0

Table B.2: Structures in S2F5. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

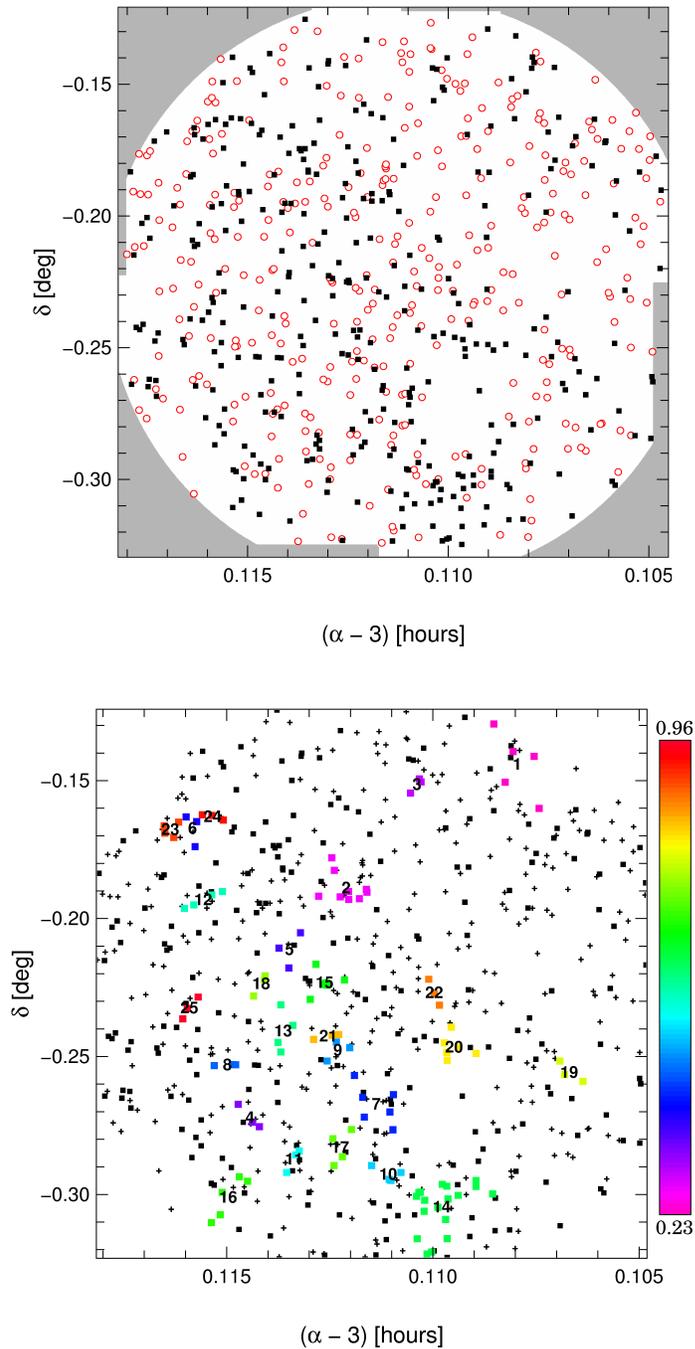


Figure B.3: Geometrical borders and projected distribution of structures in S2F5. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

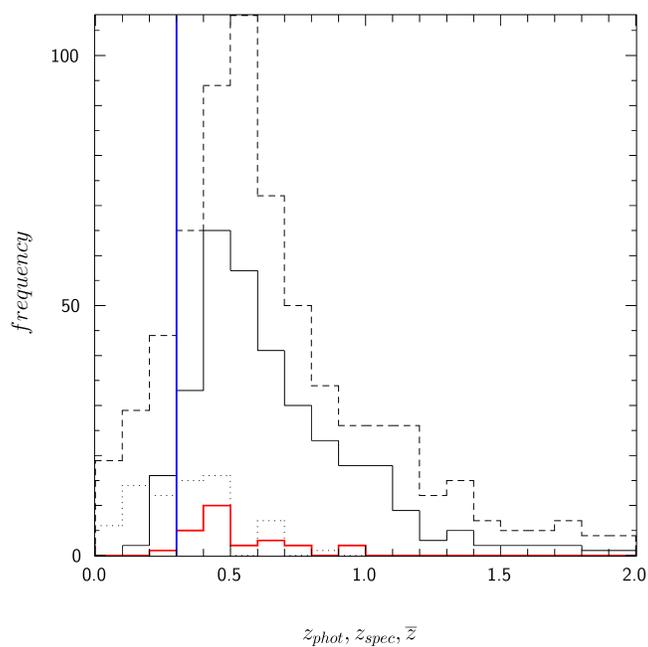


Figure B.4: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S2F5. See upper panel of Fig. 6.2 for more details.

B.3 The field S3F5

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	6	09:03:44.58	30:03:27	0.206	1	1	2	0	0
2	3	09:03:59.93	30:00:21	0.278	1	2	5	0	0
3	5	09:03:54.40	30:07:03	0.419	2 (2)	5	3	0	0
4	3	09:03:33.79	30:03:21	0.433	1	1	5	0	0
5	6	09:04:01.88	30:06:20	0.503	1 (1)	2	4	0	0
6	4	09:03:39.56	30:05:49	0.507	3	1	1	0	0
7	3	09:03:38.52	29:57:30	0.542	5 (1)	4	1	0	0
8	18	09:03:28.38	30:05:34	0.550	1	1	3	0	0
9	6	09:03:48.34	29:59:21	0.553	3	4	3	0	0
10	3	09:03:20.01	30:04:28	0.580	2	2	2	0	0
11	5	09:03:41.15	30:04:41	0.594	2	1	4	0	0
12	4	09:03:36.39	30:03:42	0.609	1	1	4	0	0
13	3	09:03:31.84	29:57:47	0.642	1 (1)	5	1	0	0
14	7	09:03:31.10	30:06:57	0.642	1 (3)	2	1	0	0
15	6	09:03:51.70	30:02:15	0.678	1	1	1	0	0
16	3	09:03:25.23	30:04:11	0.854	1	5	1	0	0
17	3	09:03:19.02	30:03:18	0.900	1 (1)	5	1	0	0
18	3	09:03:33.12	29:57:20	1.285	1 (1)	3	0	0	0

Table B.3: Structures in S3F5. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

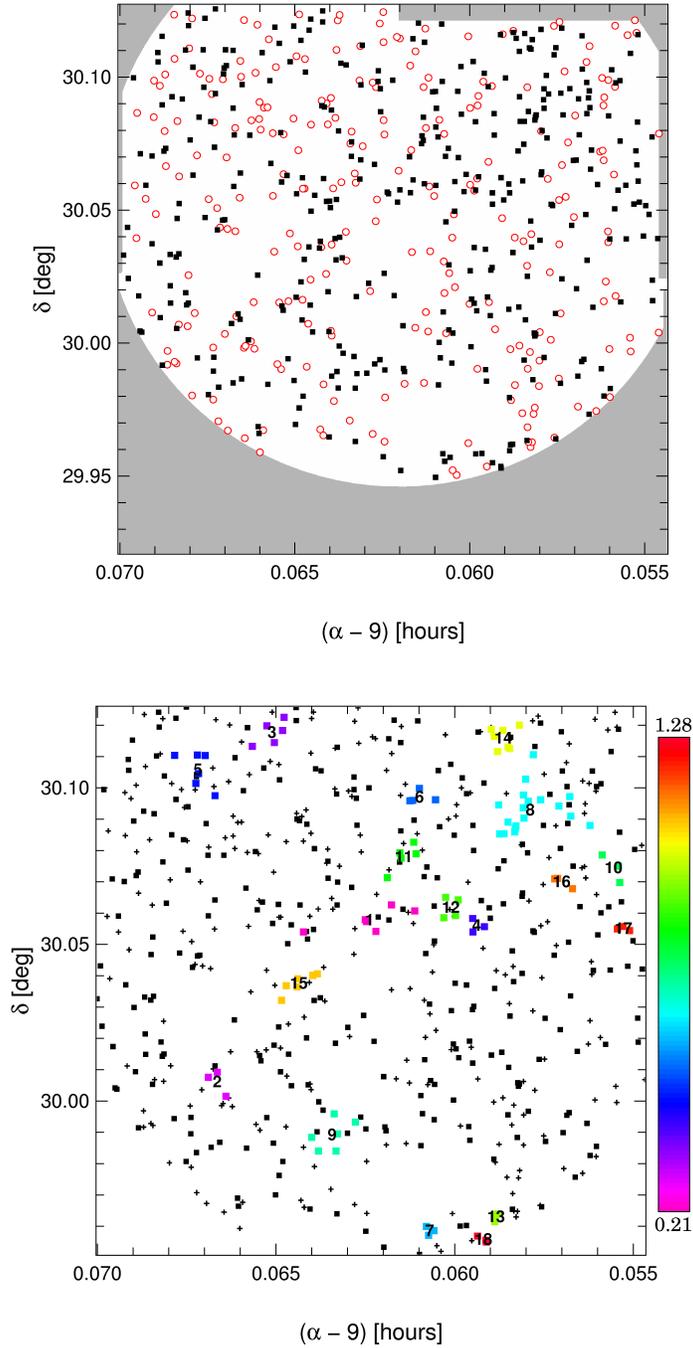


Figure B.5: Geometrical borders and projected distribution of structures in S3F5. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

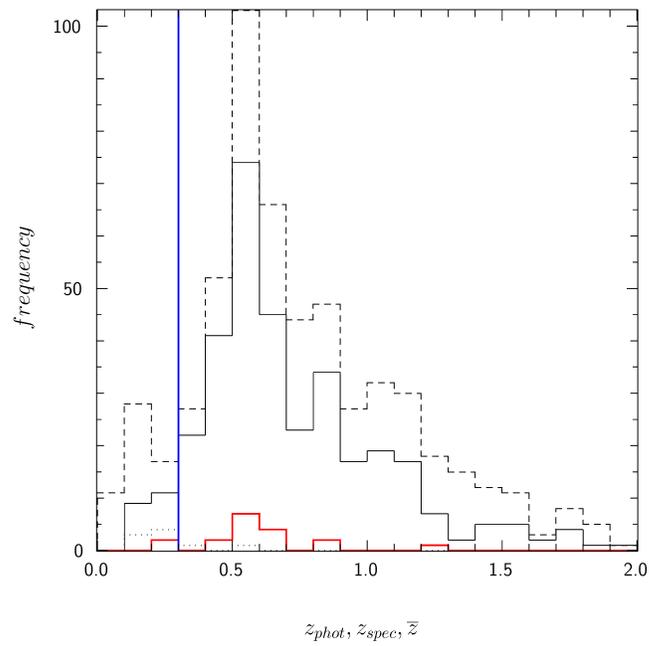


Figure B.6: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S3F5. See upper panel of Fig. 6.2 for more details.

B.4 The field S5F1

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	5	10:23:33.20	39:44:47	0.239	3 (3)	1	5	2	0
2	25	10:24:14.33	39:46:02	0.254	1 (1)	1	5	8	+1
3	3	10:24:03.08	39:48:24	0.265	2	1	4	1	0
4	5	10:23:32.42	39:49:20	0.292	4 (2)	1	5	3	0
5	12	10:23:36.28	39:51:30	0.296	2 (3)	1	4	4	+1
6	12	10:24:22.87	39:45:22	0.334	1	1	5	2	0
7	6	10:23:51.78	39:51:38	0.346	3	1	5	1	0
8	3	10:24:26.53	39:48:03	0.348	3 (1)	1	3	1	0
9	3	10:24:14.86	39:49:47	0.368	5	5	1	0	0
10	4	10:23:48.30	39:43:52	0.387	5 (1)	2	5	3	+1
11	5	10:23:39.64	39:49:01	0.393	2	1	5	0	0
12	4	10:23:48.35	39:48:28	0.449	2	1	3	2	0
13	11	10:24:16.02	39:46:24	0.496	1	2	5	1	0
14	4	10:23:41.06	39:45:05	0.544	2	5	5	1	0
15	4	10:24:06.63	39:49:47	0.559	3	5	5	1	0
16	6	10:23:58.21	39:42:55	0.562	1	5	4	0	0
17	4	10:24:19.32	39:51:13	0.588	2	5	3	1	0
18	3	10:24:05.44	39:51:07	0.664	1	5	3	1	0
19	3	10:23:42.68	39:48:56	0.717	1	5	1	1	0
20	3	10:24:05.04	39:46:55	0.871	1	5	5	0	0
21	7	10:24:07.10	39:52:20	1.005	1	2	0	1	0

Table B.4: Structures in S5F1. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

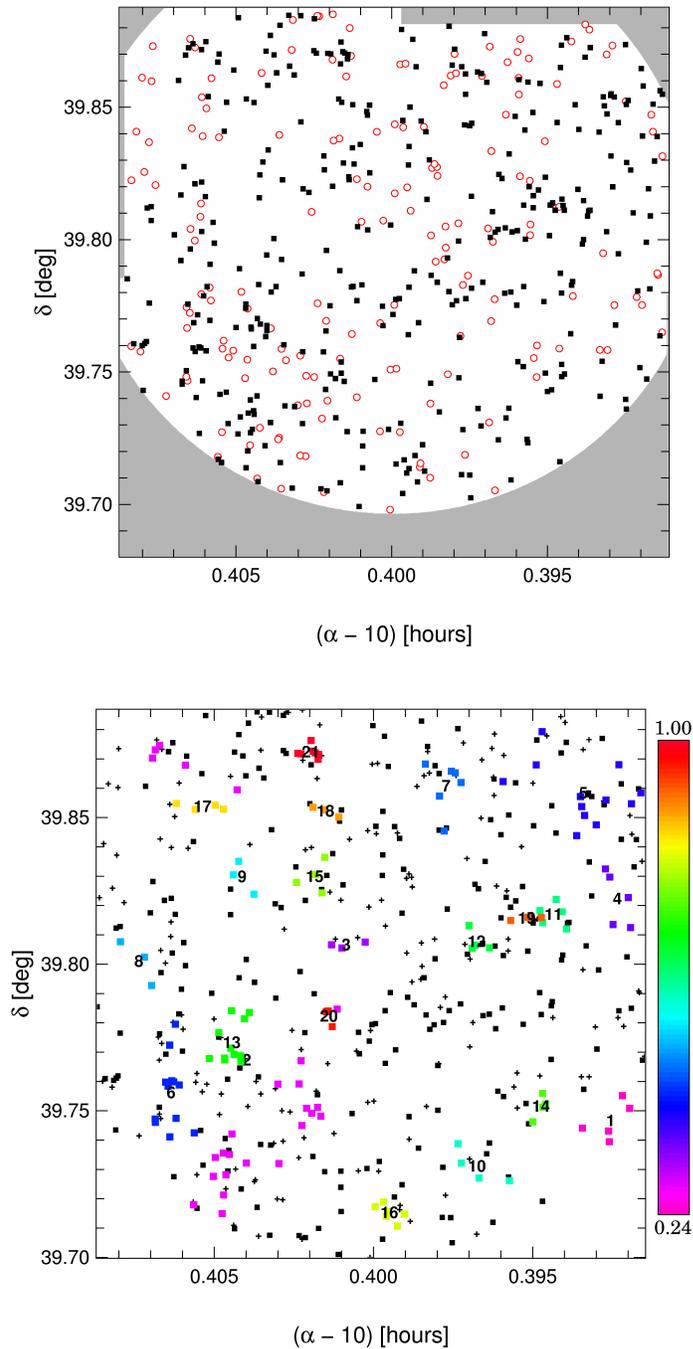


Figure B.7: Geometrical borders and projected distribution of structures in S5F1. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

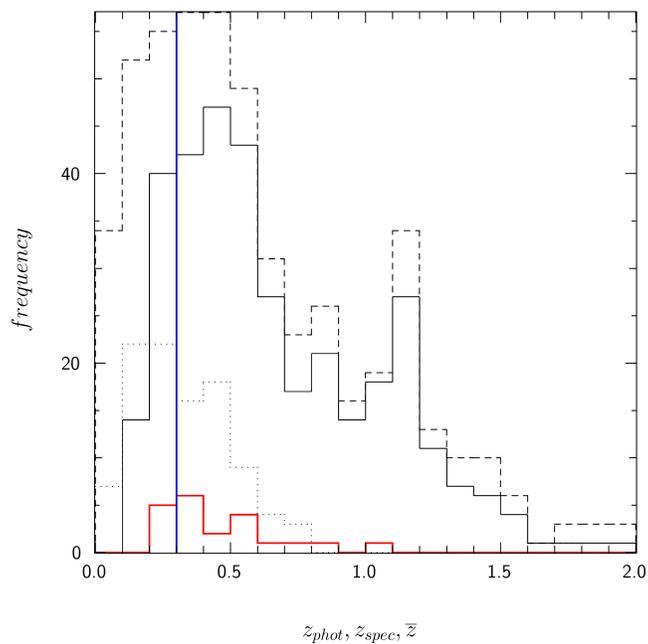


Figure B.8: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S5F1. See upper panel of Fig. 6.2 for more details.

B.5 The field S5F5

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	33	10:25:24.05	39:48:25	0.193	2 (4)	1	5	2	0
2	6	10:25:31.91	39:45:25	0.256	1	1	4	0	0
3	4	10:25:36.51	39:50:46	0.304	4 (1)	2	2	0	0
4	7	10:24:55.27	39:50:51	0.326	2 (2)	1	5	0	0
5	4	10:25:36.59	39:47:30	0.375	1	2	3	0	0
6	3	10:25:07.41	39:51:19	0.395	3 (1)	1	2	0	0
7	5	10:25:13.34	39:44:33	0.403	2	2	2	0	0
8	3	10:25:06.24	39:43:28	0.420	1	4	2	0	0
9	9	10:25:07.76	39:48:42	0.430	1	1	2	0	0
10	7	10:25:16.21	39:46:34	0.495	1	1	2	0	0
11	4	10:25:04.00	39:49:43	0.586	1	2	1	0	0
12	3	10:25:23.64	39:52:33	0.664	1 (2)	5	1	0	0
13	3	10:25:41.65	39:45:48	0.694	1 (2)	5	1	0	0
14	3	10:24:57.46	39:48:55	0.701	1	5	1	0	0
15	3	10:24:56.48	39:43:24	0.716	1 (2)	5	1	0	0
16	3	10:24:53.15	39:44:18	1.063	1	3	0	0	0

Table B.5: Structures in S5F5. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

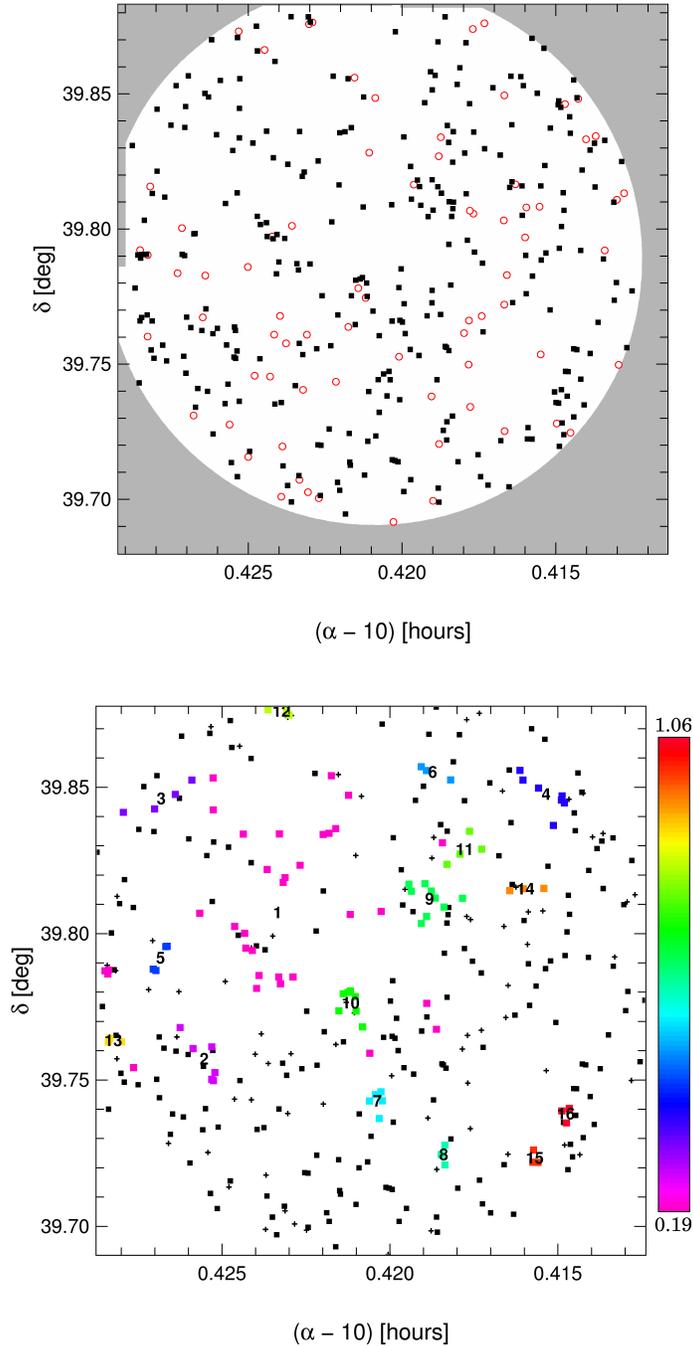


Figure B.9: Geometrical borders and projected distribution of structures in S5F5. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

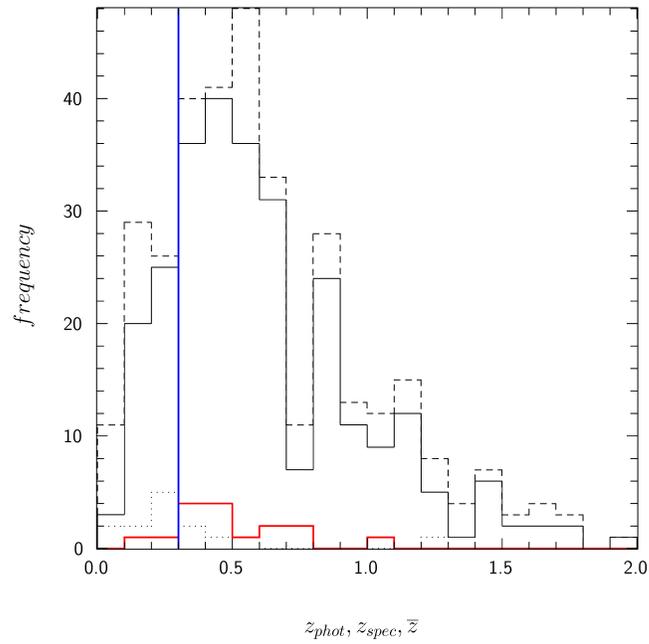


Figure B.10: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S5F5. See upper panel of Fig. 6.2 for more details.

B.6 The field S6F1

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	10	11:55:47.23	65:35:39	0.129	2 (1)	5	5	1	0
2	6	11:56:40.33	65:39:38	0.280	1 (2)	5	5	0	0
3	3	11:55:33.06	65:31:57	0.341	1 (1)	3	4	0	0
4	3	11:56:06.80	65:32:32	0.362	1	1	3	0	0
5	5	11:56:11.39	65:33:22	0.368	3	1	4	0	0
6	6	11:55:01.64	65:34:10	0.377	1 (2)	5	1	0	0
7	4	11:56:03.40	65:29:46	0.386	0 (4)	2	2	0	0
8	4	11:56:16.32	65:31:45	0.389	1	1	2	0	0
9	7	11:56:24.08	65:30:29	0.418	1 (1)	1	4	0	0
10	6	11:56:26.75	65:37:20	0.496	1	2	4	0	0
11	5	11:56:38.88	65:36:08	0.573	5 (4)	5	1	0	0
12	3	11:56:35.24	65:37:11	0.600	1 (1)	2	3	0	0
13	5	11:56:21.06	65:36:15	0.612	3	5	3	0	0
14	5	11:56:24.17	65:30:37	0.627	3 (1)	1	3	0	0
15	3	11:56:35.12	65:32:38	0.629	3 (1)	5	4	0	0
16	4	11:55:18.11	65:40:06	0.643	1 (1)	5	2	0	0
17	6	11:55:09.58	65:39:37	0.652	1 (1)	5	1	0	0
18	3	11:56:01.86	65:33:46	0.721	2	1	1	0	0
19	3	11:56:17.75	65:39:14	0.738	1	5	4	0	0
20	3	11:56:00.94	65:37:26	0.757	4	2	4	0	0

Table B.6: Structures in S6F1. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

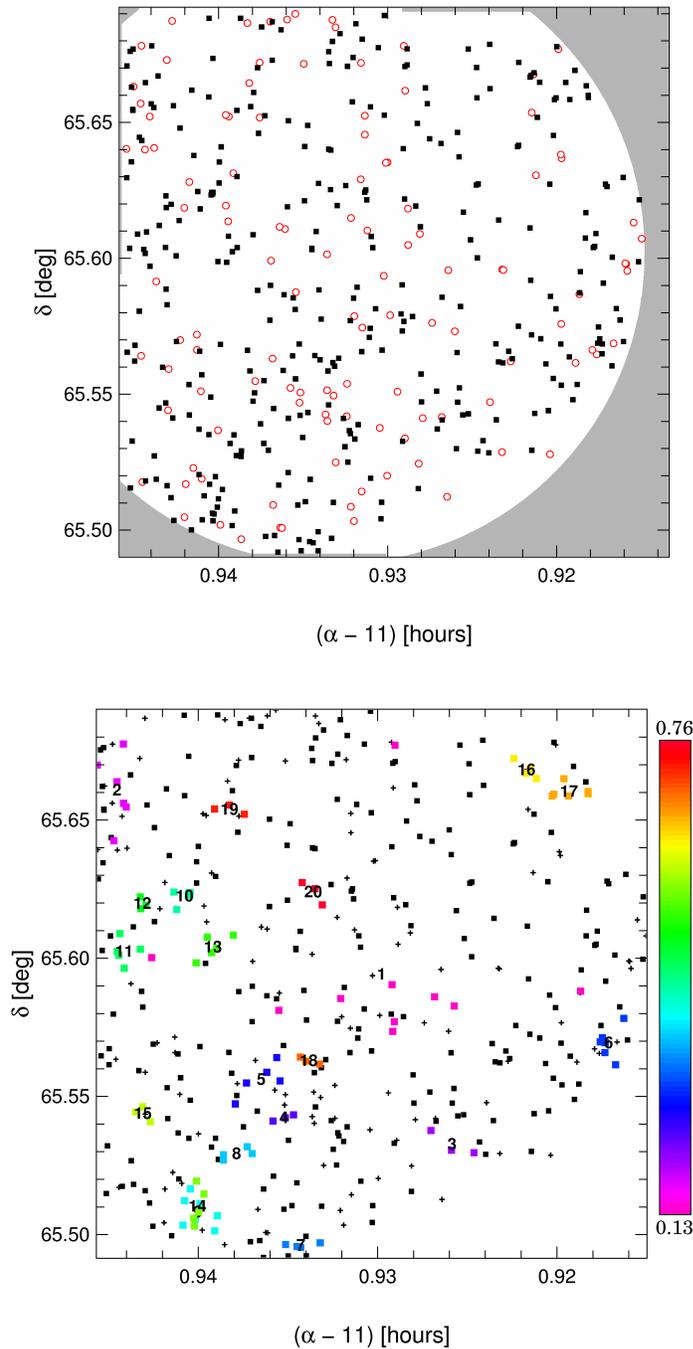


Figure B.11: Geometrical borders and projected distribution of structures in S6F1. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . The structure name “14” is plotted over the “9”. See Fig. 6.3 for more details.

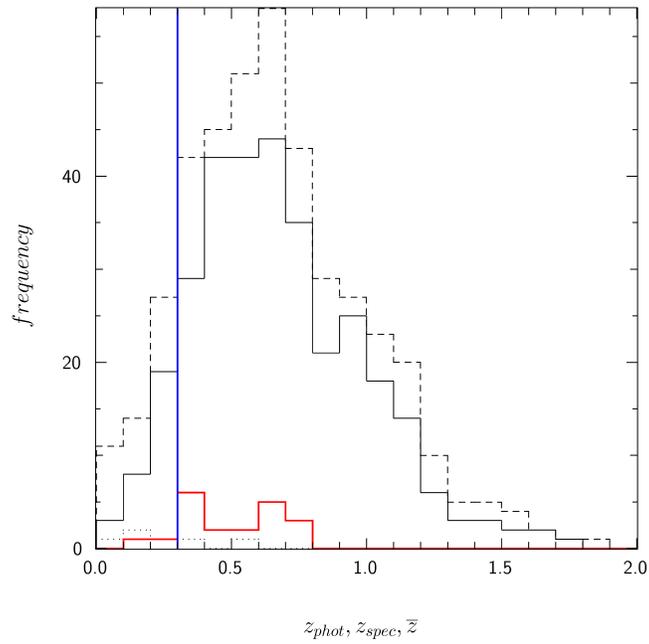


Figure B.12: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S6F1. See upper panel of Fig. 6.2 for more details.

B.7 The field S7F5

Struct.	N	α (J2000)	δ (J2000)	\bar{z}	VT	LA	CMD	Spec.	Ver.
1	8	13:34:33.85	16:56:19	0.268	1 (1)	2	2	2	+1
2	6	13:34:55.43	16:52:59	0.354	1	1	3	2	0
3	7	13:34:51.11	16:45:46	0.409	2 (3)	5	1	1	0
4	4	13:34:26.58	16:47:14	0.436	2 (1)	4	1	1	0
5	3	13:34:43.69	16:52:41	0.438	1	1	2	0	0
6	3	13:34:35.66	16:52:58	0.443	1	1	2	0	0
7	3	13:34:39.10	16:46:40	0.466	2	5	1	0	0
8	4	13:34:23.21	16:53:08	0.467	3	1	2	0	0
9	6	13:34:34.47	16:47:15	0.471	2	4	4	1	0
10	3	13:34:21.70	16:47:40	0.485	1 (2)	2	1	0	0
11	5	13:34:21.99	16:54:03	0.519	1	1	1	2	+1
12	3	13:34:29.85	16:55:27	0.528	1	1	1	0	0
13	3	13:35:00.11	16:55:06	0.528	5	4	2	0	0
14	3	13:34:50.37	16:47:02	0.531	5	5	1	1	0
15	3	13:34:39.78	16:50:36	0.545	5	5	2	0	0
16	4	13:34:33.78	16:56:50	0.562	1 (1)	1	1	0	0
17	3	13:34:48.21	16:53:52	0.576	2	2	3	0	0
18	17	13:34:35.23	16:55:29	0.627	1	1	1	3	0
19	6	13:34:34.02	16:53:41	0.649	1	1	1	0	0
20	4	13:34:30.11	16:52:11	0.653	2	2	1	0	0
21	3	13:34:47.49	16:55:49	0.825	1	5	1	0	0
22	4	13:34:51.82	16:53:01	0.924	1	1	1	0	0
23	3	13:34:27.90	16:51:24	0.934	1	3	2	0	0

Table B.7: Structures in S7F5. The table lists the number of structure members N , the mean right ascension α , declination δ , and redshift \bar{z} . It contains information about the quality assessments of the Voronoi tessellation (VT), the likelihood approach (LA) and the color-magnitude diagram (CMD). Additionally, the number of spectroscopic redshifts per structure is included, as well as the verification status. See Table 6.9 for more details.

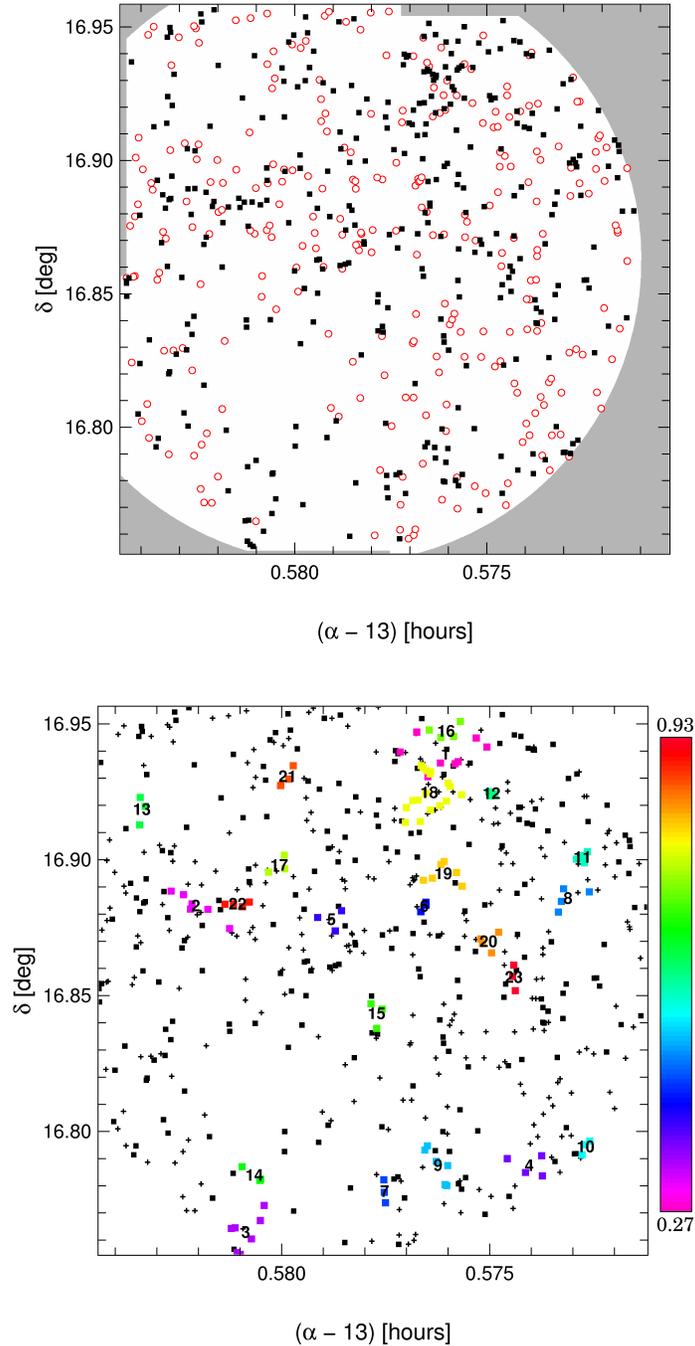


Figure B.13: Geometrical borders and projected distribution of structures in S7F5. The **upper panel** shows the geometry of the overlap area (white), with red circles marking the culled, and black squares denoting the pre-selected galaxies. The **lower panel** illustrates the structure distribution. Crosses (squares) signify the culled (pre-selected) galaxies. Structure members are plotted colored. The color indicates the structure's \bar{z} . The color scale is shown on the right side of the figure. Structure names are plotted at the mean α and δ . See Fig. 6.3 for more details.

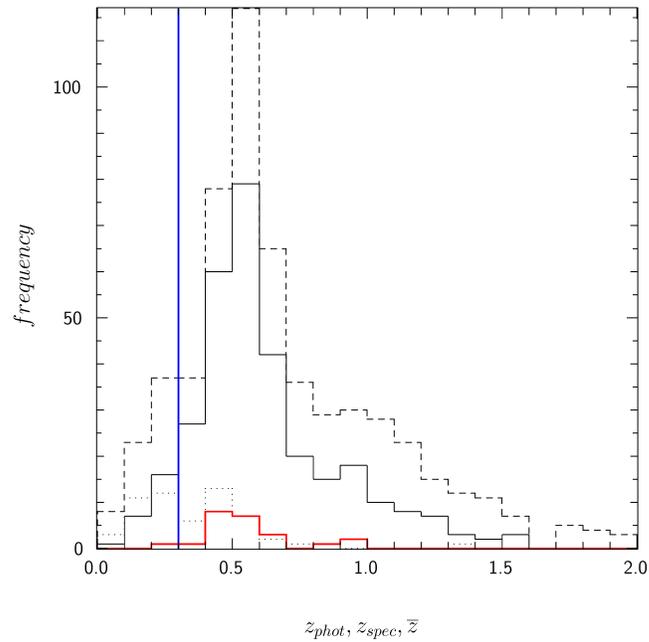


Figure B.14: Redshift distribution of all (dashed black line), pre-selected (solid black line), and spectroscopic (dotted black line) galaxies, as well as structures (red line) in S7F5. See upper panel of Fig. 6.2 for more details.

Acknowledgments

My thesis supervisor Ralf Bender deserves my deeply felt gratitude for his support, especially for convincing the stipend committee to give me a stipend, and offering me the SFB position. Without him I could not even have started my Ph.D. thesis.

I want to thank Ulrich Hopp, my advisor, deeply for supporting my ideas, allowing me the freedom to solve problems my way, for always putting much effort into understanding my – for some too – mathematical and systematic approach to science, for trusting in my formulae, codes, and calculations, for his honesty and good advice and simply for being a good friend.

Special thanks are in order to Jan Snigula, co-author of my paper(s), patient listener to any of my cluster finding ideas ranging from pure brainstorming to bright even at 01:00 in the night, or in the middle of a rapid on some wild New Zealand river, for bearing my enthusiasm as well as my disappointments, for being the excellent administrator of my institute PC and of our beloved Beowulf cluster, for teaching me that a good computer code is more than a couple of commands that “do the job”, and most of all for being one of the best friends I ever had.

The other members of the MUNICS team, my observer colleagues, Niv Drory, Georg Feulner, and Yulia Goranova, whose work on the MUNICS galaxies was the basis for my structure finding efforts, deserve my honestly felt gratitude and appreciation. Quite naturally, I could not have done anything without their results.

I am grateful to my various roommates, Sarah Bühler, Niv Drory, Yulia Goranova, Johannes Koppenhöfer, Claudia Maraston, Daniel Thomas, Jens Thomas, and Martin Tschimmel for the good working atmosphere.

My fellow extragalactic astronomy Ph.D. students at the USM, Georg Feulner, Jürgen Fliri, Armin Gabasch, Claus Gössl, Yulia Goranova, Aleksii Halkola, Arno Riffeser, Jan Snigula, and Jens Thomas, deserve my thanks for many lively discussions, wonderful parties, and simply for the nice time we shared.

I would like to explicitly thank the senior members of the extragalactic group at the USM and MPE, Ralf Bender, Laura Greggio, Ulrich Hopp, Claudia Maraston, Marc

Neeser, Roberto Saglia, Mara Salvato, Stella Seitz, and Daniel Thomas, who have taught me a lot about astronomy. Their comments about my work were highly valued, and I always felt them to be extremely constructive.

Laura Greggio, Harald Lesch, and Regina Schulte-Ladbeck deserve my special, deeply felt gratitude for their interest and believe in my scientific work, for their good advice, and also for their support.

I furthermore would like to thank John Huchra for discussing his CFA1 galaxy catalog with me, as well as Douglas Tucker for his extensive comments on his LCRS cluster catalog. Further thanks are in order to Hans Böhringer and Peter Schuecker for the fascinating discussion about their likelihood filter technique, and their interest in my EXT-FOF algorithm and the MUNICS clusters.

I would also like to explicitly thank the Bavarian State for granting me the Ph.D. stipend, as well as the Sonderforschungsbereich 375 - Astroteilchenphysik of the Deutsche Forschungsgesellschaft for supporting the MUNICS project. Further thanks are due to the Ω Cam project for allowing me to use and test their Beowulf cluster. I could not have completed my work without that incredible machine.

At this point I would like to thank all of my friends, both at the institute and at home, for sharing some years of our lives together, for helping me deal with the loss of my grandmother and also with my parents' diseases, for all the wonderful evenings spent in the movies, for the canoeing trips, martial arts and dinner sessions, for every single minute of their friendship and love.

Last but definitely not least, my deepest thanks go to my parents and grandmothers. Only their love, encouragement, endless support and constant believe in me and my abilities enabled me to do what I always dreamed of.

Pursue that which you treasure. If you must concede, let it be only to a majestic mountain.

– traditional Maori proverb

