A search for distant clusters of galaxies in the COSMOS field

Michael Zatloukal

Max-Planck-Institut für Astronomie

Heidelberg 2008

DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Dipl.-Phys. Michael Zatloukal

born in Stuttgart, Germany

Oral examination: 22 October 2008

A search for distant clusters of galaxies in the COSMOS field

Referees: PD Dr. Hermann-Josef Röser
Prof. Dr. Matthias Bartelmann

To Chia Lin

Zusammenfassung

Die genaue Entwicklung von Galaxienhaufen und ihren Mitgliedsgalaxien im Rotverschiebungsbereich oberhalb von $z = 1$ ist momentan immer noch unklar, da nur wenige entfernte Galaxienhaufen bekannt sind. Um diese wichtige Epoche zu untersuchen und die Entwicklung der Schlüsselcharakteristika, die wir heute in Galaxienhaufen beobachten, zu verfolgen, stellt der HIROCS-Survey eine statistisch signifikante Haufenstichprobe im Rotverschiebungsbereich $0.5 \leq z \leq 1.5$ zusammen. Für diese Dissertation wurde die Haufenselektionsfunktion von HIROCS mit Hilfe eines künstlichen Himmels basierend auf semi-analytischen Galaxienentwicklungssimulationen detailliert untersucht. Öffentliche und proprietäre Daten im COSMOS-Feld wurden kombiniert, und die folgende Haufensuche resultierte in einem Katalog von Haufen im Rotverschiebungsbereich $0.5 \leq z \leq 1.59$. Darin enthalten ist die erste größere, konsistent selektierte Stichprobe von entfernten ($z \geq 1.2$) Galaxienhaufen. Die Farbentwicklung von röntgenhellen und -dunklen Galaxienhaufen wurde verglichen mit dem Ergebnis, dass letztere bei allen Rotverschiebungen weniger passive Mitgliedsgalaxien als ihre röntgenhellen Gegenstücke besitzen. Spektroskopische Nachfolgeuntersuchungen von vier Haufenkandidaten bei $z \approx 0.85$, darunter drei röntgendunklen, wurden durchgeführt, um mögliche daraus resultierende Auswahlssytematiken zu untersuchen. Mit Hilfe der Stichprobe entfernter Haufen wurde gezeigt, dass Galaxienhaufen im Rotverschiebungsbereich $z \geq 1.2$ vornehmlich blaue Galaxien mit aktiver Sternentstehung und einer signifikanten Anzahl wechselwirkender Mitglieder besitzen, dass ihre Frühtyp-Haufengalaxien aber alt sind und den Großteil ihrer Sternpopulationen bei $z_f \geq 3$, möglicherweise $z_f \geq 5$, gebildet haben, in hervorragender Übereinstimmung mit Vorhersagen aus Galaxienentwicklungssimulationen.

Abstract

The evolution of galaxy clusters and their members above $z \geq 1$ is largely unclear at present due to the small number of known distant clusters. To probe this important period and witness the development of key cluster characteristics observed today, the HIROCS survey is establishing a statistically significant $0.5 \leq z \leq 1.5$ cluster sample. For this thesis, the HIROCS cluster selection function was studied in detail using a mock sky based on semi-analytical galaxy evolution simulations. Public and proprietary data were combined for the COSMOS field. A cluster search was performed, yielding a $0.5 \leq z \leq 1.59$ catalog of clusters, including the first larger, uniformly selected sample of distant ($z \geq 1.2$) clusters published. Comparing the color evolution of X-ray bright and -dark clusters, the latter were found to contain fewer passive galaxies than their X-ray counterparts. Spectroscopic follow-up observations of four $z \approx 0.85$ candidates, three of them X-ray dark, were conducted to probe cluster detection biases related to this. Using the distant cluster sample, it was shown that $z \geq 1.2$ clusters predominantly contain blue galaxies with active star formation and a significant fraction of interacting members, but that their early type galaxies are old and formed the bulk of their stars at $z_f \geq 3$, possibly $z_f \geq 5$, in excellent agreement with predictions from simulations.

Contents

1 Introduction

The large-scale structure of the distribution of galaxies in the universe resembles a foam-like web, as several redshift surveys have shown. Galaxies preferably populate the filaments, while there are huge voids with extremely low galaxy density in between them. In the picture of hierarchical structure growth, now accepted as the standard model for the evolution of structure in the universe, this very prominent large-scale structure traces back to quantum fluctuations in the early, hot and dense phase of the universe. Then, $\sim 10^{-43}$ sec after the Big Bang, inflation may have boosted its volume by a factor of $\sim 10^{30}$. This picture is still under debate, but solves the so-called "horizon problem", the fact that the temperature of the cosmic microwave background (CMB) is the same everywhere, even in regions that otherwise could never have interacted, explains how almost instant separation of initial quantum fluctuations can keep them from being canceled out by interaction, thus imprinting the temperature fluctuations that we see today on the CMB, and results in the universe being isotropic and flat, the state that is being observed now. The blown-up quantum fluctuations then served as initial conditions for the subsequent formation of structure, which is dominated by the dark matter. Interacting only through gravity, its behavior can be modeled relatively easily. Indeed, simulations of the dark matter halos have been very successful in reflecting the structure of the universe as seen at all redshifts accessible by observations. Figure 1.1, taken from such a simulation, gives snapshots of the dark matter distribution at a redshift of $z = 1.4$ (top row) and today (bottom row) in slices of thickness $15\,\mathrm{Mpc}/h$. The baryonic mass, and thus also the galaxies, only trace the underlying dark matter halo distribution seen there, without influencing it substantially. As seen by observations, the galaxies live in the filaments of the dark matter, accumulating in their crossing points through gravitational attraction. Smaller galaxy accumulations form galaxy groups, while the galaxy clusters, being the most massive gravitationally bound structures in the universe, reside in the major crossings of the filaments. Dark matter halos and thus galaxies are being accreted by these centers of mass, streaming in along the filaments, leading to structure growth and contrast enhancement of the web structure over time, as can be seen by comparing the top and bottom rows of Figure 1.1.

The size scales of these web structures are on the order of Mpc in comoving coordinates; while the dimension of voids is $\sim 25\,\mathrm{Mpc}$, large galaxy clusters with up to > 1000 member galaxies span $3 - 10\,\mathrm{Mpc}$, and the smaller galaxy groups are $1 - 2\,\mathrm{Mpc}$ across. As first discovered by Zwicky (1933), the dark matter also dominates the mass budget of galaxy clusters, making up $\sim 80\%$ of their total mass and forming a potential well for the other constituents to live in. The galaxies only account for a mass fraction of $\sim 5\%$, while $\sim 15\%$ of the cluster mass are contributed by the hot, $\sim 10^6 - 10^8\,\mathrm{K}$ intra-cluster gas. This gas shines as diffuse emission in the X-rays through thermal bremsstrahlung (free-free) emission of X-ray photons. Even though galaxy clusters

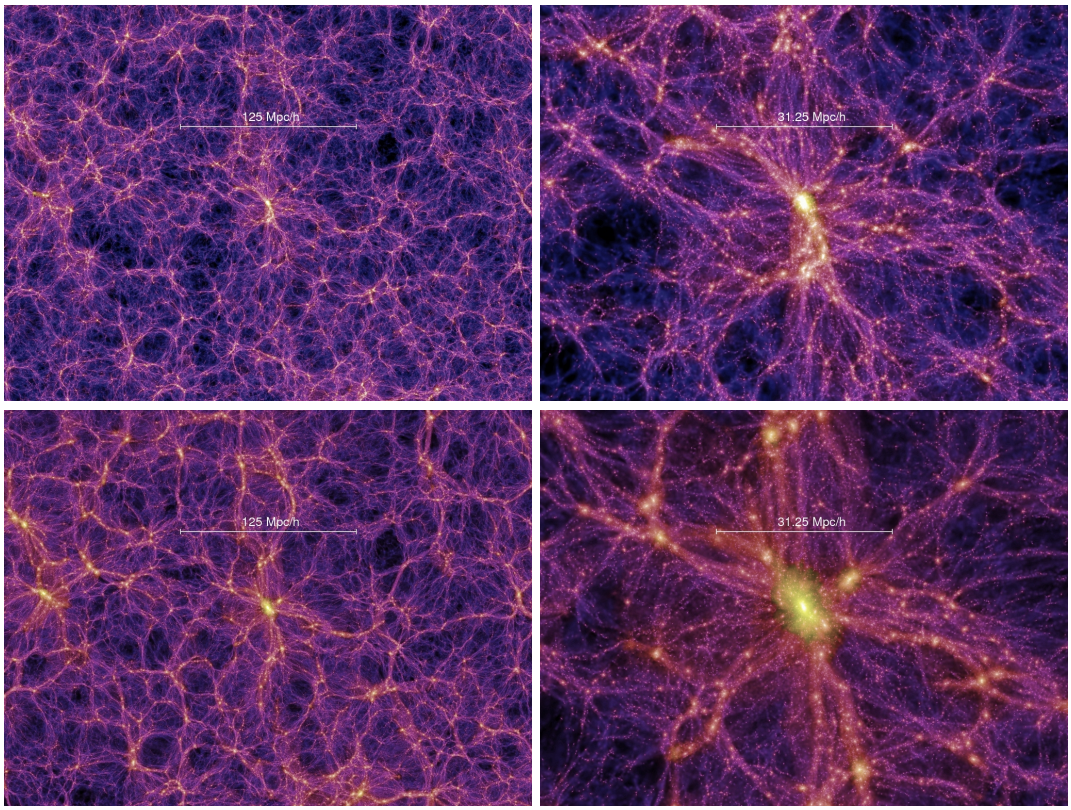

Figure 1.1: The cosmic large-scale structure at $z = 1.4$ (top row) and $z = 0$ (bottom row) with a thickness of 15 Mpc/h from the Millennium Simulation (Springel et al. 2005). Both rows show the same object in different stages of its development. The right panels are a zoom-in on the left panels. The colors represent the density, with bright standing for dense regions, whereas dark regions represent voids. While the density peak (*i.e.* the galaxy cluster) is already in place at $z \sim 1.5$, there is significant mass accretion until today.

contain many galaxies, their contribution to the overall galaxy number count is small: While about 50% of all galaxies live in smaller structures like galaxy groups, only about 5% of the bright galaxies are found in rich clusters.

Galaxy clusters serve as scientific probes for two major research fields of extragalactic astronomy: Structure formation in the universe, and galaxy evolution.

1.1 Scientific aspects of galaxy clusters

1.1.1 Galaxy clusters as probes to constrain cosmology

Marking the highest-density peaks in the distribution of dark matter in the universe, the abundance of galaxy clusters directly traces the cosmological parameters of the now "standard" Λ Cold Dark Matter (ΛCDM) model or "concordance cosmology" relevant

Table 1.1: Important concordance cosmology parameters. Basic parameters are to be measured, derived parameters follow from these. The density of the universe is $\Omega = \Omega_m + \Omega_\Lambda = 1$. Most likely values are given based on the newest 5-year results from WMAP (Dunkley et al. 2008).

Basic parameters		Estimated value
H_0	Hubble parameter ($h = H_0/100\mathrm{kms}^{-1}\,\mathrm{Mpc}^{-1}$)	$71.9 \pm 0.026\,\mathrm{kms}^{-1}\,\mathrm{Mpc}^{-1}$
Ω_m	Total matter density	0.258 ± 0.030
Ω_b	Baryon density	$(0.02273 \pm 0.00062)\,/h^2$

Derived parameters		Estimated value
ρ_0	Critical density for flat universe	$(1-2) \cdot 10^{-26}\mathrm{kg/m}^3$
Ω_Λ	Dark Energy density	0.742 ± 0.030
σ_8	Initial power spectrum normalization	0.796 ± 0.036

for structure formation. "Cold" here means that the rest mass of dark matter particles is much greater than the equivalent mass of their kinetic energy. A listing of the most important concordance cosmology parameters is given in Table 1.1. Statistics about the very existence of galaxy clusters at different redshifts, together with some of their global properties, can probe cosmological parameters. The local cluster mass function, which depends on the matter density Ω_m and the initial density perturbation power spectrum normalization σ_8, can serve as an example. The degeneracy of this dependency in the local case can be broken by measuring the mass function evolution over redshift. After having provided early evidence that the universe is of low matter density, the focus in cosmology-driven galaxy cluster research is now shifting, with the properties and nature of Dark Energy emerging as the most fundamental open questions of concordance cosmology today. Distant galaxy clusters, with their number density and mass function being sensitive to the Dark Energy density Ω_Λ, have been recognized as one of the most promising ways to constrain cosmological parameter space. Together with observations of cosmic microwave background fluctuations and distance measurements from Supernovae Ia, large samples of galaxy clusters allow to constrain cosmological models to a high level of precision. A review of the current state of the art is found in (Schücker 2005).

All of these tests require huge samples to achieve the required statistical significance, and thus extreme sky coverage. Constraining cosmological parameters is thus neither the goal of this thesis nor the goal of the underlying galaxy cluster survey.

1.1.2 Galaxy clusters as laboratories for galaxy evolution

Studies of galaxy clusters across cosmic times are also important tools in the quest to understand the evolution of galaxies. Clusters combine a large number of galaxies of different mass, luminosity, morphology and other properties in one single object. With the compared to the field greatly enhanced galaxy density speeding up the relevant processes and the quite good accounting of intergalactic baryons, their thermal state, and their heavy-element enrichment, clusters can serve as laboratories to study galaxy evolution.

Here, one needs to keep in mind the peculiarities of the cluster environment that possibly alters the evolution of the member galaxies - the dark matter potential trough, and resulting from this both the hot intracluster medium, and the highly increased galaxy density. Due to this, the population mix observed in clusters has some unique features.

One galaxy type, resembling an elliptical galaxy with an extended halo and termed "cD galaxy", is only found in the centers of clusters. cD galaxies resemble elliptical galaxies, with a similar brightness profile, but with a very extended, low-surface brightness halo of stars without clear boundaries in addition. At lower redshifts, their luminosity correlates with the cluster mass, indicating parallel evolution (Lin and Mohr 2004). Compared to the field, clusters contain a lot more of so-called "early type", elliptical and S0 (lenticular) galaxies. Their distribution within the cluster follows the "morphology-density relation" (Dressler 1980), meaning that early-type galaxies predominantly populate the cluster centers, while late-type spiral galaxies are found in the outskirts. Postman et al. (2005) showed that the morphology-density relation is already in place at $z = 1$, albeit with a substantial decrease in the fraction of central S0 galaxies at higher redshifts, going along with a parallel increase in spirals. This results suggests that transformation from spiral into S0 galaxies occurs at $z \lesssim 1.0$ and is driven by environmental effects (see below).

Two formation scenarios have been discussed for elliptical galaxies: monolithic collapse and hierarchical formation. In the monolithic collapse model (Eggen et al. 1962), gas is turned into stars instantaneously by rapid collapse of gas clouds. This burst of star formation is followed by passive evolution with no subsequent star formation. No additional mass is acquired at later times in this model. The fact that red sequence galaxies at the same redshift have almost identical colors (see Figure 1.3) points towards a similar formation age and supports this model. In addition, the colors of passively evolving galaxies usually are consistent with formation redshifts $z_f \geq 3$. The opposite is the case in the hierarchical formation scenario (*e.g.*, White and Rees 1978). Here, smaller dark matter halos form first and merge to more massive halos later. Merging events can substantially change galactic properties like morphology, and can trigger star formation events as well as nuclear activity. In this scenario, mass is acquired throughout the galaxy's lifetime, and massive elliptical galaxies are predicted to appear rather late. But since mergers can change the morphology of galaxies, the ellipticals observed in distant galaxy clusters do not necessarily correspond to the ellipticals observed locally (the so-called progenitor bias, van Dokkum and Franx 1996). To be able to decide between the two scenarios, constraints need to be put on the star formation history as well as the mass assembly history of elliptical galaxies, both of which have been very active areas of recent

research. Eisenhardt et al. (2008) found the $[I - 3.6\mu]$ color evolution of cluster galaxies up to $z \sim 1$ to be consistent with a star burst at $z_f = 3$, followed by passive evolution. At $z > 1$, a wider range of formation times is needed, but still with a preference for higher formation redshifts ($z_f > 3$). The mass assembly of cluster galaxies, thought to occur through mergers in the hierarchical formation picture, is still under debate. The term "dry merger" has been coined for mergers among two passive galaxies that does not spark significant star formation, thus allowing ellipticals to consist almost exclusively of an old stellar population even in an hierarchical formation scenario with its numerous merger events. Some studies show the recent mass growth of passive cluster galaxies to be significant, with a doubling of their mass since $z \sim 1.5$ inferred from the H-band luminosity function evolution (Faßbender 2007), whereas de Propris et al. (1999) conclude from studies of the K-band luminosity function that the assembly of cluster galaxies was largely complete by $z \sim 1$. On the theory side, numerical simulations based on semi-analytical models are now able to make detailed predictions about properties of elliptical galaxies in the hierarchical formation scenario. They predict a star formation activity peak around $z \sim 5$, but only half of the present-day ellipticals' progenitors have $\geq 50\%$ of their final mass assembled by $z \sim 1.5$ (De Lucia et al. 2006).

In addition to the differences in galaxy population based on environment density, there are indications for a redshift evolution of the cluster members. First noticed by Butcher and Oemler (1984), the fraction of blue galaxies in galaxy clusters seems to increase with redshift. This result, later called the "Butcher-Oemler effect", is difficult to measure. It depends strongly on the definition of the blue fraction (see *e.g.* Andreon et al. 2004) and is still under debate (Andreon et al. 2006). Figure 1.2 shows a recent attempt to quantify this effect (Margoniner et al. 2001).

The question whether or not the galaxy luminosity function is independent of environment is also far from being settled. While found to be universal by some studies (Andreon 2004), others find a significant dependence on environment (*e.g.*, Goto et al. 2002).

Undoubted is that galaxies show a bimodality in the color-magnitude diagram, while the strength of this effect again depends on environment. The passively evolving galaxies populate the tight "red sequence" (see Figure 1.3), while star-forming galaxies are found in a broader "blue cloud". These two extremes are separated by the sparsely populated "green valley". The color and slope of the red sequence itself shows evolution with redshift. The slope of the red sequence (redder colors for the more luminous objects) can be explained by a metallicity sequence (Kodama and Arimoto 1997), caused by the fact that the more massive systems can retain more metals from the supernova driven winds of heated gas due to their deeper potential wells with higher binding energy per unit mass. A detailed study of the evolution of the color bimodality for field galaxies at $0.2 \leq z \leq 1.1$ is given by Bell et al. (2004). Long thought to be exclusively formed by passive early-type galaxies, recent studies point towards a population of dusty star-forming galaxies on the red sequence, as *e.g.* found by Wolf et al. (2005) in multi-color observations of the Abell 901 / 902 galaxy cluster system. Without dust obscuration altering the color, the star formation in blue cloud galaxies needs to be shut down to make them transit onto the red sequence. This can *e.g.* happen through a major merger (in this case gas-rich or "wet"). Once the galaxy has arrived on the red sequence, it may

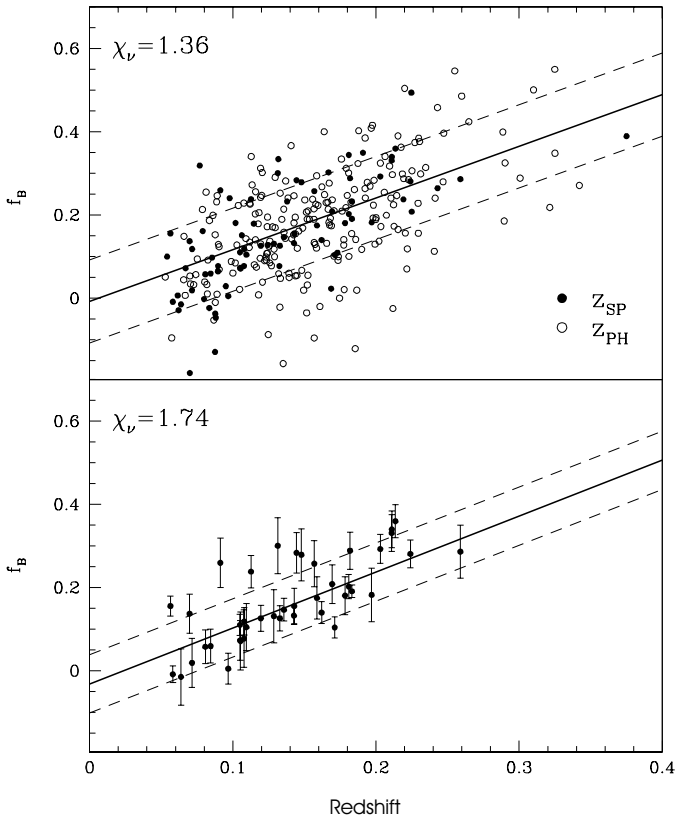

Figure 1.2: Fraction of blue galaxies in the magnitude range between M^*-1 and M^*+2 and within 0.7 Mpc from the center of the cluster. The upper panel shows the entire cluster sample, the lower panel only contains clusters with spectroscopic redshifts and a well-defined color-magnitude relation. Taken from Margoniner et al. (2001).

evolve and brighten there slowly through a series of gas-poor ("dry") mergers. Early quenching, when the merging galaxies are still small, results in fainter red sequence galaxies, whereas in the case of late quenching, the galaxies assemble most of their mass while still blue (Faber et al. 2007). Which of these scenarios is the dominant one is still remaining an open question. Also, the first appearance of the red sequence (Kodama et al. 2007) or the redshift at which the galaxy colors split up to show the bimodality observed at lower redshifts is unclear at present and a topic of active research. Cirasuolo et al. (2007) find this color bimodality to disappear at $z \approx 1.5$. Furthermore, the studies of Cucciati et al. (2006) and Cooper et al. (2007), both based on large spectroscopic surveys, find that at $z \approx 1.5$, the fraction of red galaxies correlates only weakly with overdensity, in contrast to $z < 1$, and that the color-magnitude diagrams of low- and high-density regions become more and more similar with increasing redshift at $z > 1$. Both studies' results argue for a "nurture" scenario (see below) of star formation shutdown, together with a "downsizing" (Cowie et al. 1996 Gavazzi and Scodeggio 1996) scenario, where star formation shifts to lower luminosity systems with increasing time.

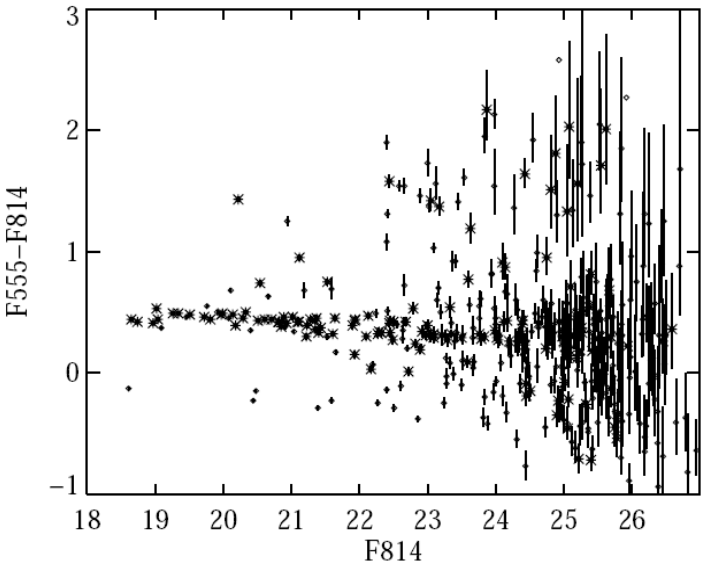

Figure 1.3: The observed color-magnitude diagram for Abell 2390, based on two-filter HST imaging of the cluster core. The asterisks indicated galaxies selected morphologically as early types, and diamonds indicate other galaxies in the image. Errors are 1σ. Taken from Gladders and Yee (2000)

Many different effects have been proposed to explain the observed properties of cluster galaxies, like shape transformation due to frequent high-velocity encounters in the densely populated environment, called "harassment" (*e.g.*, Moore et al. 1996), "ram-pressure stripping" of cold gas from galaxies traveling through the cluster by interaction with the intra-cluster medium, thus suppressing star formation (Gallagher 1978), or "strangulation" of the star formation of galaxies at first encounter with a cluster due to the gas being removed by the cluster gravitational potential (*e.g.*, Balogh and Morris 2000). The formation of huge central cluster galaxies could be stimulated by dynamical friction, causing massive galaxies to spiral to the cluster center on the timescale of a few Gyrs. There, they could form one giant object by "galactic cannibalism" (Hausman and Ostriker 1978), meaning that the most massive galaxy grows by consuming its neighbors. Also being debated is the "pre-processing", a change in the properties of the galaxies on the group scale, before they are entering the cluster environment. This is for example favorable for the merger rate, which is inverse proportional to the velocity difference of the galaxies.

There is a lot of research activity going on to decide which of the proposed effects play the dominant role in shaping the observed cluster population, and whether the differences between the cluster and field populations result from "nurture", meaning that the object properties are a reaction on the environment as for the effects listed above, or from their "nature", *i.e.* differences in structure formation processes between the cluster and field environments.

1.2 Detecting galaxy clusters

To get reliable statistics both for cosmological parameter determination as well as for studies of galaxy evolution, large, homogeneously selected samples of galaxy clusters over the widest redshift range possible are required. Several characteristics of galaxy clusters can be traced as a basis for surveys. The following gives a brief and by no means complete summary of different search methods with example surveys that employ it.

1.2.1 Tracing galactic light

The most straight-forward method, and thus the first one to be applied, is to trace the light emitted by the cluster member galaxies. The first samples were based on galaxy overdensities selected by eye on photographic plates (Abell 1958 Zwicky et al. 1968) of the Palomar Observatory Sky Survey (POSS), a major milestone in astronomical surveys and still in use today in its digitized form (Digitized Sky Survey, DSS). These cluster catalogs are major achievements and contain the brightest and most prominent clusters, but they are limited in redshift and suffer from incompleteness as well as inhomogeneity.

The first automated cluster search again traced clusters as overdensities of galaxies in projection, employing a count-in-cells method on data from the Lick survey (Shectman 1985). With the advent of CCDs, deeper and uniform datasets could be taken. However, the search for more distant galaxy clusters on deep datasets is complicated by fore- and background contamination, bringing the cluster selection methods based solely on selecting projected overdensities to their limit. Several methods have been developed since to add distance information to the data, thus increasing the contrast of the clusters with respect to the field.

The matched-filter algorithm developed for the 5.1 square degree Palomar Distant Cluster Survey (PDCS, Postman et al. 2005) employed positional and photometric data simultaneously to decrease projection effects, using a radial profile filter for cluster detection and a luminosity function as a filter for member selection.

In contrast used to the number counts of the PDCS, the 130 square degree Las Campanas Distant Cluster Survey (LCDS, Gonzalez et al. 2001) traces galaxy clusters as positive brightness fluctuations in the background It was targeted at finding $0.3 \lesssim z \lesssim 1.1$ clusters with shallower data than cluster searches which need to resolve the member galaxies.

The Red-sequence Cluster Survey (RCS, Gladders and Yee 2000) imaged 100 square degrees in two filters (R_c, z'), and selecting galaxy clusters as overdensities of objects whose colors match model calculations of the cluster red sequence. Since the red sequence evolves with redshift, this method successfully avoids projection effects, and the cluster distance can be estimated quite accurately from the red sequence properties in addition. With the filters chosen, this survey can detect clusters out to $z \approx 1.1$.

The wide-field, optical multi-color Sloan Digital Sky Survey (SDSS, York et al. 2000) provides a unique dataset for low-redshift galaxy cluster searches. Several different cluster search methods have been applied to it, one of them being the “Cut-and-enhance

method" (Goto et al. 2002), which uses color cuts to increase the contrast of the clusters, combined with a density enhancement algorithm.

Another cluster search on the SDSS data has been done using the maxBCG cluster finder (Koester et al. 2007), also based on selecting galaxy clusters by their red sequence. As indicated by the name, the brightest cluster galaxy (BCG) serves as a starting point for the subsequent color-selected member search.

In contrast to imaging surveys, spectroscopic surveys can use precise redshifts and thus do not suffer from contamination. Structures can be picked with a simple friends-of-friends algorithm (*e.g.*, Ramella et al. 2002). The major drawbacks are that they require big investments of telescope time, pre-selection of objects and are limited in redshift since spectra with a signal-to-noise ratio sufficient for determination of redshifts are hard to obtain for faint objects.

1.2.2 Tracing the intra-cluster medium

Apart from tracing clusters by their galactic light in the optical and near-infrared, they can also be traced very successfully by the diffuse X-ray emission that the heated intra-cluster gas emits. This detection method does almost not suffer from projection effects, so the presence of diffuse emission in the X-rays together with a photometrically or color-selected overdensity of galaxies is sufficient to confirm the presence of a galaxy cluster (Ostriker et al. 1995). Furthermore, an estimation of the cluster mass is possible through the empirically found correlation of the X-ray luminosity of a cluster with its mass, the $L_X - M$ relationship.

The success of X-ray based cluster searches is largely due to the German-American ROSAT (ROentgenSATellit) mission from 1990-1999. Several surveys have used this observatory to detect galaxy clusters like the ROSAT Deep Cluster Survey (RDCS, Rosati et al. 1998), detecting galaxy clusters out to $z = 1.27$, and others. Several compilations of ROSAT-selected clusters with different properties have been studied, like the ROSAT-ESO Flux-Limited X-ray cluster survey (REFLEX, Böhringer et al. 2004) and the NOrthern ROSAT All-Sky survey (NORAS, Böhringer et al. 2000), both serving as important local comparison samples with $z \lesssim 0.3$ to calibrate for example the luminosity function and the galaxy cluster power spectrum.

The intra-cluster medium (ICM) also influences the photons from the CMB. This so-called Sunyaev-Zel'dovich effect (SZE, Sunyaev and Zeldovich 1980), compton-scattering of the CMB photons on the ICM electrons, shifts the photon energy to a higher level, thus causing a deficit in photons of lower temperature in the direction of the galaxy cluster. The magnitude of this effect is independent of redshift, making it favorable to detect very distant galaxy clusters (*e.g.*, Kneissl et al. 2001). However, this is complicated by the weakness of the signal and possible confusions with other sources of temperature fluctuations in the CMB. To date, no new galaxy clusters have been discovered using the SZE, but the CMB signature of existing systems could be confirmed (see for example (Lancaster et al. 2005) for results on nearby clusters, and (Muchovej et al. 2007) for a study of $z \sim 1$ systems). A review of how the SZE can be used to constrain cosmological parameters is given by Carlstrom et al. (2002).

1.2.3 Tracing the cluster mass

A very different approach is to trace the galaxy clusters by the signal that the gravitational potential of their dark matter halo mass imprints on the appearance of background galaxies. The deflection of their light due to the weak-lensing effect causes subtle changes in their shape. These are hard to measure due to their weakness, but averaged over many objects, the effect becomes visible and permits the creation of a mass profile of the lensing object. To detect gravitational potentials in this way, the presence of a suitable background of more distant galaxies is needed, complicating the process with increasing redshift. Very deep surveys taken under formidable conditions are required to trace clusters out to high redshifts. Taken together with the large areas necessary to be statistically significant for blind searches, this detection method is very costly. A review of the field is given in (Mellier 1999). More recently, Dietrich et al. (2007) compiled a sample of 155 X-ray, optically and shear-selected clusters to study selection biases between the three detection methods. For pre-known clusters, weak lensing is very successfully used to map the dark-matter distribution and obtain mass estimates. For example, Taylor et al. (2004) map the mass distribution of Abell 901 at $z = 0.16$, and Hoekstra (2007) compare weak lensing- and X-ray luminosity-inferred masses for a sample of 20 clusters.

1.3 Recent progress in the identification of distant galaxy clusters

In 2004, the number of known high-redshift galaxy clusters was very small, and practically all of them stemmed from the ROSAT mission; Figure 1.4 illustrates this. The redshift record holding object was at $z = 1.27$, marking the ROSAT redshift horizon. Apart from this, only four clusters at $1 < z \leq 1.27$, and just about ten more at $0.7 < z \leq 1.0$ were known, while the redshift range below $z = 0.7$ was sampled reasonably well.

A large sample of distant galaxy clusters is of special importance to constrain the parameter space both for determining cosmological parameters, as well as to study aspects of galaxy evolution. Thus, a number of surveys was set up to extend the sample of known clusters both to higher redshifts as well as larger numbers.

With the commissioning of the satellite-born XMM-Newton observatory, a number of new X-ray surveys for distant galaxy clusters is being conducted. Some of them are using contiguous data, like the XMM cluster search on the COSMOS field (Finoguenov et al. 2007), the XMM Large-Scale Structure survey (XMM-LSS, Pierre et al. 2006), a 9 square degree contiguous area survey with multi-wavelength data, and the XMM Cluster Survey Romer et al. (XCS, 2001), a serendipitous survey for clusters in all XMM archive fields.

Until recently, the majority of very distant ($z \geq 1.25$) galaxy clusters had been found by XMM-Newton due to their X-ray emission, most of them serendipitously (*e.g.*, Mullis et al. 2005, Barkhouse et al. 2006, Faßbender 2007). This picture is in the process of shifting now, with large-area datasets from imaging surveys involving near-infrared

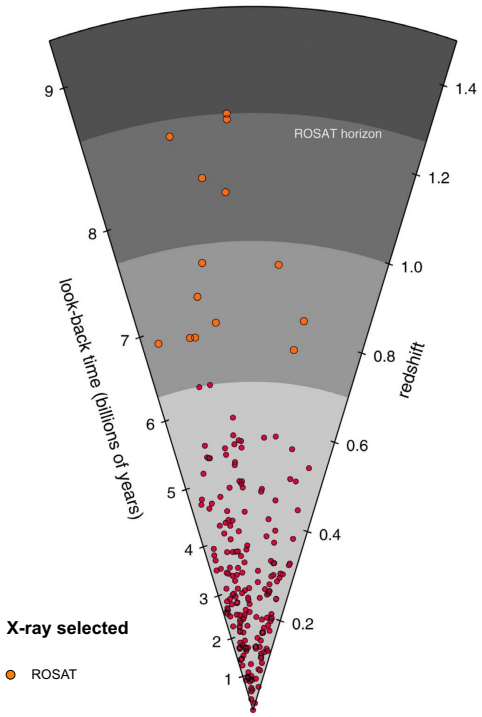

Figure 1.4: Known distant galaxy clusters in 2004. The diagram is only complete above $z = 0.7$; the horizontal distribution of clusters at the same redshift is just for clarity and does not have any special meaning. Adapted from C. Mullis.

coverage becoming available. All projects listed in Section 1.2.1 are limited in redshift to $z \approx 1.1$, where the most prominent spectral feature of early-type galaxies, the 4000 Å break, is redshifted out of the CCD range. The recent availability of wide-field infrared cameras, both ground-based as well as satellite-born, has made it possible to extend the reach of imaging surveys, and a number of new projects is making use of these capabilities.

Several of the recent surveys employ the Spitzer space telescope. The Spitzer Adaption of the Red-sequence Cluster Survey (SpARCS, Wilson et al. 2006) uses a detection method similar to the RCS, but does photometry in the z'-band and the IRAC 3.6μ channel, thus extending the survey redshift range out to $z \sim 2$.
The cluster search of the IRAC shallow survey covers a 7.1 square degree field with Spitzer IRAC as well as ground-based imaging data and spectroscopic redshifts available (Elston et al. 2006). The candidate sample, found by means of tracing galaxy overdensities using photometric redshift information, has been published recently. It includes clusters out to $z \sim 1.7$ (Brodwin et al. 2008), some of them spectroscopically confirmed (Stanford et al. 2005, Eisenhardt et al. 2008).
In the COSMOS field, a cluster search based on a red-sequence selection of objects involving IRAC bands is being conducted, but no candidate sample has been published to date.

In addition, a number of ground-based imaging surveys is currently being conducted. The FLAMINGOS Extragalactic survey (FLAMEX, Elston et al. 2006) covers 7.1 square degrees to a 50% completeness limit of $K_s = 19.2^{\mathrm{mag}}$, aiming at discovering galaxy clusters at $1 \leq z \leq 2$. The data can be used in combination with the Spitzer IRAC Shallow survey.
The cluster search on UKIDSS Ultra Deep survey data has published a first list of candidates out to $z = 1.45$ (van Breukelen et al. 2006), two of which are at $z \geq 1$, based on an overdensity selection employing photo-zs, and tried to confirm the highest-redshift candidate by spectroscopic follow-up observations (van Breukelen et al. 2007). It turned out that the system is a projection of three structures at $z = 1.40$, $z = 1.45$, and $z = 1.48$, the most massive of which at $z = 1.454$ is most likely a cluster of mass $M \sim 10^{14}\,\mathrm{M}_{\odot}$.

A number of high-redshift cluster candidates has also been discovered serendipitously by ground-based imaging, among them a group of red objects around a galaxy with a spectroscopic redshift of $z = 1.51$ by McCarthy et al. (2007), and a photometrically selected candidate at $z = 1.6$ in the GOODS field (Castellano et al. 2007).

Several of the detection methods discussed above employ filters based on assumptions about cluster characteristics to increase the contrast of the clusters with respect to the field, trying to avoid projection effects. While this works well in most cases, there is a danger of biasing the sample towards clusters that match these assumptions. The universal presence of a prominent red sequence for $z \geq 1.5$ clusters, as an example, is still being debated, see Section 1.1.2. The studies by Cucciati et al. (2006) and Cooper et al. (2007) even find that at $z \sim 1.5$, in contrast to $z < 1$, even the most luminous red galaxies do not reside preferentially in high density environments, and that the red fraction of galaxies correlates only weakly with overdensity at $z \gtrsim 1.3$. It is also not known today what the formation epoch of X-ray clusters is or whether there are clusters without significant X-ray emission to be expected, since it is not clear today when and how the ICM forms and virializes. In the light of all this, it is promising to conduct high-redshift cluster searches with a minimum of initial assumptions going into the detection method.

One of the most basic features of a galaxy cluster, one could even say its definition, and thus a detection criterion which fulfills the condition mentioned above, is that a galaxy cluster represents an overdensity of galaxies compared to the surrounding field. This definitely holds at all redshifts. Thus, a survey identifying galaxy clusters by this primary property promises to yield an unbiased sample. One of its most important uses is to probe whether or when the assumptions an which other search methods are based on hold.

These are the underlying ideas of the Heidelberg InfraRed / Optical Cluster Survey (HIROCS, Röser et al. 2004), the project in whose context the work for this thesis has been done. Some of the surveys mentioned above should also be able to select mostly unbiased samples like HIROCS, especially the cluster search on the UKIDSS Ultra Deep survey data, and the IRAC Shallow survey since they both rely on identifying clusters purely as galaxy overdensities. The latter however introduces a preference for passive galaxies with object selection solely in the IRAC 4.5μ band.

The work of this thesis focuses on the COSMOS field, which is included in one of the

four HIROCS fields. With deep, comprehensive datasets in various wavelength regimes available, especially in the X-rays, the COSMOS field is ideally suited for cluster searches and a study of the selection biases introduced by the various search methods.

The outline of this thesis is as follows: In Chapter 2, HIROCS is introduced and described in some detail. Chapter 3 gives an overview of the data in the COSMOS field used for this thesis, both public and proprietary. Chapter 4 deals with data reduction and analysis. The cluster selection algorithm developed for this thesis is explained in Chapter 5, and the results of its characterization using mock sky data are presented. The cluster sample obtained for the COSMOS field is described in Chapter 6; spectroscopic follow-up observations of a group of four cluster candidates are the topic of Chapter 7. The thesis closes with the final discussions and conclusions in Chapter 8.

Throughout this thesis, $H_0 = 70\,\mathrm{km\,sec^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$ are assumed. All magnitudes are given in the Vega system unless quoted otherwise.

2 The Heidelberg InfraRed / Optical Cluster Survey

The Heidelberg InfraRed / Optical Cluster Survey (HIROCS, Röser et al. 2004) is a wide-area multi-band imaging survey designed to provide a homogeneous and unbiased selection of galaxy clusters in the redshift range $0.5 \leq z \leq 1.5$ to study the evolution of the clusters themselves as well as the one of their member populations as a function of redshift. HIROCS will cover 8 square degrees in five filters, one of them being in the near-infrared.

HIROCS is part of the large extragalactic key program MANOS (MPI for Astronomy Near-infrared / Optical Surveys) which has been launched in 2002 to exploit the newly commissioned wide-field imaging abilities of the Calar Alto observatory. This program consists of two parts, one of which is HIROCS or MANOS-wide. The companion survey MANOS-deep, also known as COMBO-17+4, is a deep imaging survey in 21 bands covering approximately 0.77 square degrees. COMBO-17+4 is an extension of the COMBO-17 survey (Wolf et al. 2003), adding deep imaging in four near-infrared bands to study star formation rates of field galaxies out to redshift $z = 2$.

2.1 Survey goal and design

As has been discussed in the introduction, many galaxy cluster surveys trace clusters by secondary properties like X-ray emission from the hot intra-cluster gas or the presence of member galaxies living on the red sequence. To minimize biases in the cluster selection and to provide a homogeneous sample of galaxy clusters going out to redshifts where it is not yet clear to which extent the assumptions made by other surveys still hold, HIROCS searches for galaxy clusters purely as overdensities of galaxies in 3D (RA, DEC, z) space. RA and DEC can be measured from the projection onto the celestial plane, whereas the distance information comes from the photometric redshifts estimated by the multi-color classification based on imaging in multiple bands. One of the major goals of the survey is to probe the evolution of the cluster red sequence at $z > 1$ and the evolution of the ICM at high redshifts by follow-up X-ray observations of the HIROCS clusters.

The following parts of this chapter have been adapted from (Röser et al. 2004) and the MANOS proposal (MANOS-Proposal).

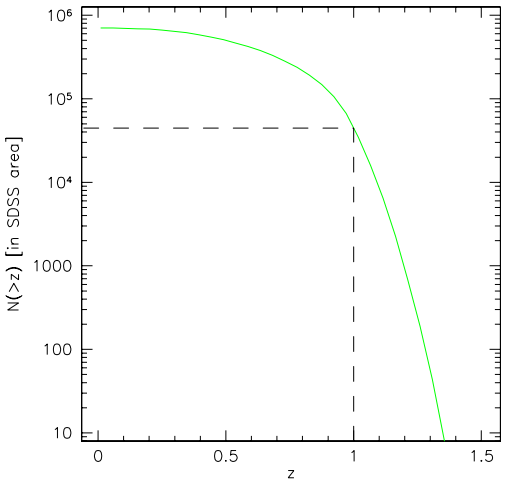

Figure 2.1: Cumulative redshift distribution $N(>z)$ of galaxy clusters with $M \geq 5 \times 10^{13} M_{\odot}$ over the whole SDSS survey area, taken from Bartelmann and White (2002). Clusters are detected at 2σ significance based on surface brightness enhancements in summed $r' + i' + z'$ SDSS data. Per square degree, the estimated number of high-redshift, $z \geq 1$ clusters is about 5.

2.1.1 Survey area

The survey area covered by HIROCS is motivated by the expected space density of high-redshift galaxy clusters. Bartelmann and White (2002) have calculated the number of clusters detectable in the SDSS due to to their surface brightness enhancement, see Fig. 2.1, resulting in an average of 5 high-redshift ($z \geq 1$) galaxy clusters per square degree. While the Bartelmann & White calculations are based on 2σ detections of the clusters and stacking of images which enhances the depth of the data, the HIROCS data, which go significantly deeper than the SDSS, promise to yield considerably more candidates. Thus, the Bartelmann & White result is taken as a lower limit for the expected HIROCS cluster yield. Based on these estimations, HIROCS needs to cover an area of the order of 10 square degrees for the cluster sample to be statistically significant also in the highest-redshift bin (about 50 galaxy clusters above redshift of one).

2.1.2 Filter set

The filter set to be used has to be optimized taking into account two boundary conditions: It has to be as small as possible to avoid excessive observational needs while still allowing for photometric redshifts precise enough to detect galaxy clusters in the desired redshift range.

Rich clusters consist to a large fraction of early type galaxies which were already in place at high redshifts (*e.g.*, Eisenhardt et al. 2008), thus the filter set needs to permit tracing the signatures of early type galaxies in the desired redshift range. The most prominent feature in the spectrum of elliptical galaxies is the 4000 Å break. To get

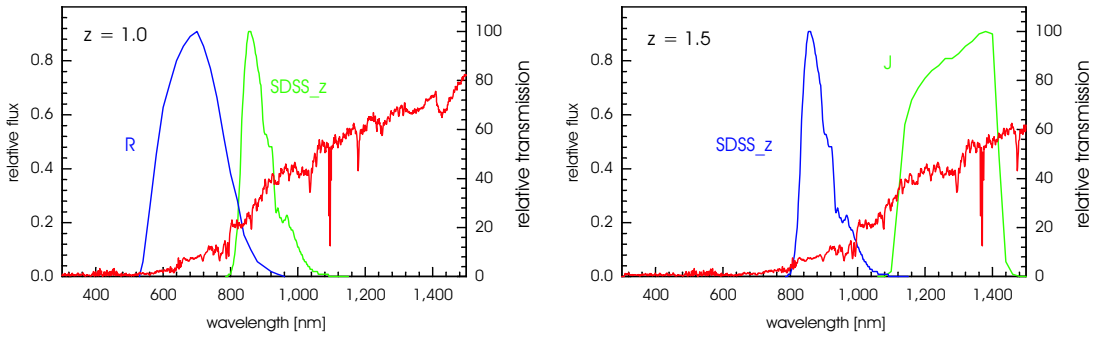

Figure 2.2: Spectrum of an elliptical galaxy at $z = 1.0$ (left panel) and at $z = 1.5$ (right panel). While at $z = 1.0$ the SDSS z filter at 913.4 nm can be used to bracket the 4000 Å break on the longer wavelength side, an infrared filter is mandatory for higher redshifts. Here the J filter at 1242 nm is shown as an example.

a precise photometric redshift estimation, this 4000 Å break needs to be bracketed by two filters. Since it is redshifted out of the range of CCDs at $z \approx 1$ (see Figure 2.2), having at least one infrared filter is mandatory to detect galaxy clusters by means of photometric redshifts above this limit.

The optimal filter set was chosen based on tests with subsets of the CADIS (Wolf et al. 2001b) filter set. CADIS uses four broad bands (B, R, J and K') together with 13 medium band filters. In the context of this survey, Wolf et al. (2001a) have developed a method to classify objects and estimate redshifts for quasars and galaxies based on multi-color imaging data, which was then used on the CADIS data set. For testing purposes, an additional intermediate Y (sometimes also called zJ) filter between z and J was introduced. The flux for CADIS objects in this additional band was estimated by convolving the best-fitting template object spectrum from the multi-color classification with the instrument response function using the Y filter.

For various subsets of the CADIS filter set, all objects brighter than $I = 23^{\mathrm{mag}}$ were re-classified. The results, mainly the photometric redshifts, were compared to the original CADIS classification (Figure 2.3).

While the scatter at redshifts above unity becomes very large without the J-band, inclusion of the Y-band does not significantly improve the photometric redshifts. The minimum number of filters still resulting in reliable photometric redshifts turns out to be five. The final HIROCS filter set consists of B, R, i, z, and H. CADIS $I1$ and CADIS $I2$ had to be dropped since the LAICA camera used for HIROCS is not equipped with these filters. Instead, SDSS $- i$ and SDSS $- z$, being equivalent in wavelength coverage, have been included. This also simplifies magnitude calibration and mosaicing by enabling comparison with the SDSS catalog. Even though the J-band is sufficient for photometric redshift estimations up to $z = 1.5$, the H-band was finally chosen over the J-band to possibly enable the search for clusters even above this redshift.

In the COSMOS field, HIROCS uses public data in B_j, r^+, i^+, and z^+(the original optical HIROCS filters), plus the u^*-, g^+-, V_j-, and NB816-band. Shallow Ks data is

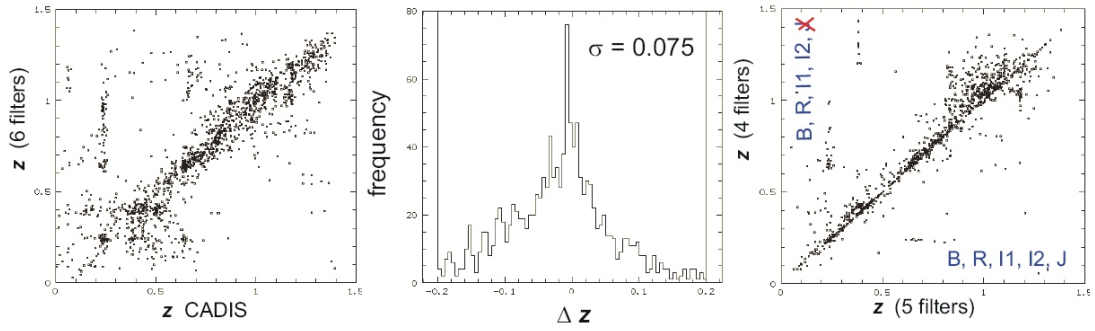

Figure 2.3: Comparison of multi-color classification results for subsets of the CADIS filter set and the full filter set used to select the HIROCS filter set. Left panel: The photometric redshifts from CADIS compared to those derived from a six-filter (B, R, CADIS $I1$, CADIS $I2$, Y, and J) subset. Central panel: Uncertainty of the photometric redshifts using this set of six filters. Right panel: While it can be shown that the Y-band is redundant and can be dropped, photometric redshifts above unity become very uncertain if the J-band is also omitted.

also available, and the H-band is added as proprietary HIROCS data. More details about the COSMOS field and data are given in Chapter 3.

2.1.3 Target fields

For extragalactic surveys, large field sizes are desirable not only to increase the sample size, but also to trace the effects of large-scale structure (LSS). The rectangular shape of the HIROCS fields has been chosen with this effect in mind, with a projected size large enough to also span voids. In addition, using disjunct fields helps to counter cosmic variance. Additionally, the fields should not contain bright foreground stars, which would take up a substantial area of the exposures due to being bloated as a result of the survey depth.

HIROCS covers 8 square degrees in four fields: The MUNICS S2F1-, the CADIS 16h-, a 22h field unique to HIROCS and the COSMOS 10h field (see Table 2.1). Two of the HIROCS fields include fields from the CADIS and MUNICS (Drory et al. 2001) surveys. The original fields are much smaller than the areas covered by HIROCS at the same position, but comprehensive data including slit spectra up to $z \sim 1$ are available here. The spectra will be used to check the reliability of photometric redshifts, analogous to the method employed in Section 4.7.1. Details about the COSMOS field are given in Chapter 3.

The four fields span the entire range in RA to enable observations throughout the whole year. With the exception of the CADIS-16h field, all fields are chosen to be equatorial to permit the usage of observatories on both hemispheres. Details about the selection of the HIROCS-22h field can be found in (Falter 2006). Originally, HIROCS was intended to cover a total of 11 square degrees in four fields, requiring a total observation time of 73.7 clear nights. In 2007, the decision was made to limit the survey area to 8 square

Table 2.1: The HIROCS target fields. RA and DEC are given in the J2000 system. Column "b" gives the galactic latitude. Column "Size" lists the dimensions of the fields. The galactic extinction E($B-V$) is shown in column 6.

Field name	RA (J2000)	DEC (J2000)	b	Size	E($B-V$)
MUNICS-S2F1	$03^{\rm h}06^{\rm m}12^{\rm s}.0$	$-00°20'35''$	$-48°.2$	$2° \times 1°$	$0.080^{\rm mag}$
COSMOS-10h	$10^{\rm h}00^{\rm m}28^{\rm s}.6$	$02°12'21''$	$42°.1$	$1.4° \times 1.4°$	$0.018^{\rm mag}$
CADIS-16h	$16^{\rm h}24^{\rm m}32^{\rm s}.0$	$55°44'32''$	$42°.1$	$2° \times 1°$	$0.006^{\rm mag}$
HIROCS-22h	$21^{\rm h}59^{\rm m}58^{\rm s}.7$	$02°23'14''$	$-39°.2$	$2° \times 1°$	$0.065^{\rm mag}$

degrees to finish within a reasonable time frame without substantially worsening the SNR: $1° \times 2°$ in the MUNICS-3h-, the CADIS-16h-, and the HIROCS-22h field, plus the 2 square degrees of the COSMOS-10h field. The special status of the COSMOS field is discussed in Chapter 3. The layout of the HIROCS fields is shown in Appendix E. It differs from the one given in (Falter 2006), and shows the new field sizes resulting from cutting the survey area to 8 square degrees.

2.1.4 Limiting magnitudes

The target depths for the different filters were again calculated based on elliptical galaxies, assumed to be the characteristic population of galaxy clusters even at high redshifts. The target depth was chosen to enable the detection of an elliptical galaxy at redshift $z = 1.5$, $1^{\rm mag}$ fainter than L^* on a 5σ level under a seeing of $1.5''$. The model spectrum of an elliptical galaxy at $z = 1.5$ from the template library of the multi-color classification (see Section 2.2.1.2) was used to calculate the magnitude limits. The integration times for the HIROCS filter set are shown in Table 2.2.

The total integration time needed to cover one square degree is 160 ksec or 7.6 clear nights including overhead. In the COSMOS field, the required integration times are smaller due to the publicly available data (see Chapter 3). In total, HIROCS needs 50.9 clear nights at the Calar Alto 3.5 m telescope to be completed.

2.1.5 Instrumentation

Optical data in B, R, i, and z are collected using the wide-field imager LAICA in the prime focus of the Calar Alto 3.5 m telescope. LAICA is equipped with four $4k \times 4k$ CCDs, arranged with gaps slightly smaller than the size of the CCDs ($15'$) between the chips. The total field of view (FoV) is about $45' \times 45'$ with a pixel size of $0.225''$. Due to the chip layout, four shots are needed to cover a contiguous area of one square degree on the sky.

In addition, the Wide-Field-Imager (WFI) at the 2.2 m ESO/MPG telescope on La Silla is used to collect B- and R-band data. In contrast to LAICA, the eight $2k \times 4k$ CCDs of WFI are arranged in a quasi-contiguous mode with gaps between the chips at most $23''$ wide. The FoV of WFI is also about one quarter of a square degree with a pixel size of $0.238''$.

Table 2.2: Exposure times and limiting magnitudes for the HIROCS filter set. Column "M*" gives the absolute rest-frame magnitudes for an elliptical galaxy at the knee of the Schechter function and redshift $z = 1.5$. K corrections are listed in column "K-corr". The apparent magnitude of the object is given in column "m*". The observation limit (column "obs. lim.") is chosen to be 1^{mag} fainter than the L^* galaxy. The integration time Δt is needed to reach the signal to noise ratio SNR for each filter when using the 3.5 m telescope on Calar Alto. The total integration time per square degree is given in column "$\Delta t/\square^\circ$". Note that the B-band integration times given in this table are different from the one given in Falter (2006).

Filter	M*	K-corr	m*	obs. lim.	Δt	SNR	$\Delta t/\square^\circ$
B	-22.00	1.06	24.2	25.2	10.0	12.2	40 ksec
R	-23.09	1.73	23.8	24.8	4.0	12.3	12 ksec
i	-23.45	1.42	23.1	24.1	4.0	6.8	12 ksec
z	-23.79	1.10	22.4	23.4	10.0	5.1	40 ksec
H	-25.02	-0.26	19.9	20.9	3.0	5.1	48 ksec
						sum	160 ksec
						eff.	68.8 h
							7.6 clear nights / $\square^\circ$

For data collection in the H-band, HIROCS makes use of the near-infrared wide-field imager OMEGA2000 in the prime focus of the Calar Alto 3.5 m telescope. OMEGA2000 is equipped with a Rockwell HAWAII2 $2k \times 2k$ detector. With its $0.45''$ pixels, OMEGA2000 can image $15.4' \times 15.4'$ in one exposure. 16 shots are needed to cover one square degree.

2.1.6 Current status

Data acquisition has been nearly completed by the end of 2007, and data reduction and analysis are underway. Preliminary analysis of one square degree in the MUNICS-3h field yielded a rich cluster candidate at $z \approx 0.7$, one candidate at $z \approx 1.25$, and one candidate at $z \approx 1.35$. The analysis of the COSMOS field could be completed first due to publicly available data and will be discussed in this thesis. The first follow-up observations of a $z \approx 0.85$ cluster candidate group in the COSMOS field have been completed, see Chapter 8. Furthermore, additional imaging in supplementary near-infrared bands has been obtained for the highest-redshift candidate sample in the COSMOS field, comprising 12 clusters with $1.2 \leq z \leq 1.55$. These data will be used to get more secure lower bounds for the photometric redshifts and probe deeper into the cluster luminosity function, however, the analysis is beyond the scope of this thesis.

2.2 Methodology

2.2.1 Multi-color classification of objects

Surveys targeted at studying the evolution of galaxies and structures composed of galaxies like clusters need to accurately determine the object class as well as the redshift for a large number of objects having very different properties. While spectra allow for the most accurate object classification and precise redshifts, observational costs are excessive for spectroscopy of a large number of faint objects, even when using multi-slit spectrometers of fiber multiplexing. In addition, this method requires pre-selection of objects. A different approach is to classify objects based on multi-color imaging data. As part of this PhD thesis, Wolf (1999) has developed a multi-color object classification scheme in the context of the CADIS survey, which has later also very successfully been applied to the COMBO-17 survey (Wolf et al. 2004). Based on template libraries, this classification software determines the object class for a variety of astrophysical objects, and estimates photometric redshifts for galaxies and quasi-stellar objects (QSOs). The experience from the CADIS and COMBO-17 surveys has shown that multi-color classification of objects based on data taken with a filter set tailored to the specific survey needs, choosing from broad-, medium-, and narrow-band filters, serves most survey purposes well enough at dramatically lower costs compared to spectroscopy, and without the need for pre-selection of objects.

2.2.1.1 Method

In principle, a multi-color imaging survey derives very low-resolution spectra of the objects. The survey is done in n filters, thus photometry samples the spectral energy distribution (SED) of objects in n bands and yields n flux measurements together with the according magnitudes for all objects in the catalog. These flux measurements define $n-1$ independent colors. The multi-color classification matches these observed colors to colors derived from template libraries containing the SEDs of astrophysical objects. For this method to work well, the template libraries need to cover a widely complete range of object SEDs (see Section 2.2.1.2 for a description of the libraries used with HIROCS). Each object is being assigned probabilities to belong to the class of stellar objects, galaxies or QSOs. For extragalactic objects, the redshift and its error are estimated using a minimum error variance method. Details about the mathematical methods used can be found in (Wolf et al. 2001a). An example of a best-fitting SED from the HIROCS survey in the COSMOS field is shown in Figure 2.4.

2.2.1.2 Template color libraries

Template libraries which accurately model the spectra of astrophysical objects are a vital ingredient for the multi-color classification. These template libraries should cover all major families of objects to be encountered in the survey, ranging from stellar objects to galaxies and QSOs, over the complete redshifted wavelength range. Each object group has its own spectral features which make identification based on multi-color imaging data possible.

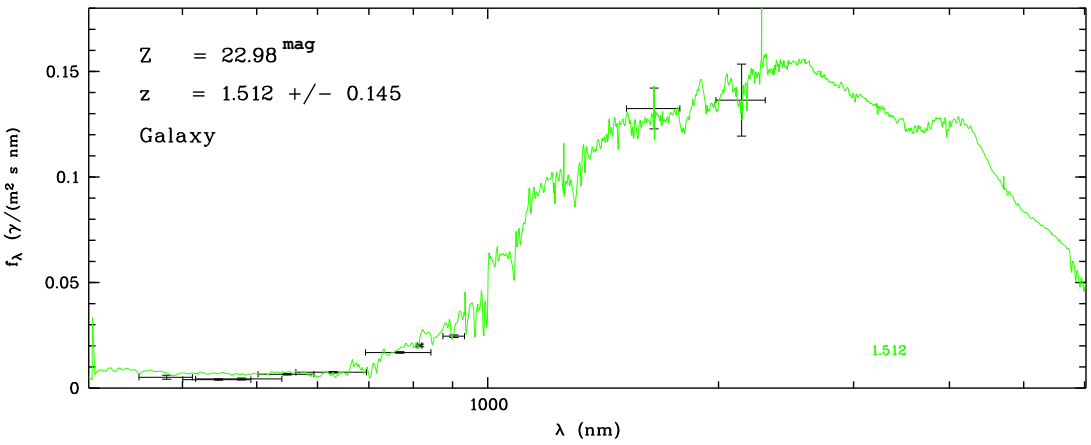

Figure 2.4: Best-fitting template galaxy spectrum (green line) for an object in the COSMOS field with flux measurements in u^*, B_j, g, V_j, r', i', $NB816$, z', H, and Ks, with the errors indicated. The estimated photometric redshift of the galaxy is $z_{\mathrm{phot}} = 1.512$.

The template libraries used in HIROCS are:

- The stellar library:
  This library is based on the 1150 Å – 25000 Å UK spectra of the Pickles catalog (Pickles 1998). It contains 96 empirical stellar spectra of the spectral types F, G, K, and M, covering luminosity classes I to V.
- The white dwarf / blue horizontal branch library:
  The spectra in this template library are based on theoretical models provided by D. Köster. Effective temperatures are covered ranging from 6 000 to 40 000 K, surface gravities in the range $\log g = [6, 9]$. The low gravity values are consistent with models of Subdwarf B-type (sdB) stars or blue horizontal branch (BHB) stars.
- The galaxy library:
  This is the heart of the multi-color classification for HIROCS since the survey is primarily searching for high-redshift cluster galaxies. The galaxy library is based on the PEGASE population synthesis code (Fioc and Rocca-Volmerange 1997). The template library spectra have been calculated on a two-dimensional grid covering 60 ages and six extinction levels. The star formation history is an initial burst with an exponential decline ($\tau = 1$ billion years) for all templates, the ages of the burst range from 50 million to 15 billion years. Kroupa initial mass functions, initial metallicity of 0.01, and no extinction were assumed. Extinction is applied afterwards as a screen following the SMC law defined by Pei (1992), using six extinction levels: $\mathrm{E}\,(B-V) = [0.0, 0.1, \ldots 0.5]$. The template library covers redshifts out to 6.7 with a resolution of 0.005 on a $\log(1+z)$ scale (417 steps in total), with Lyman forest absorption included.
- The QSO library:
  The QSO library is based on the SDSS template spectrum. The SDSS template is

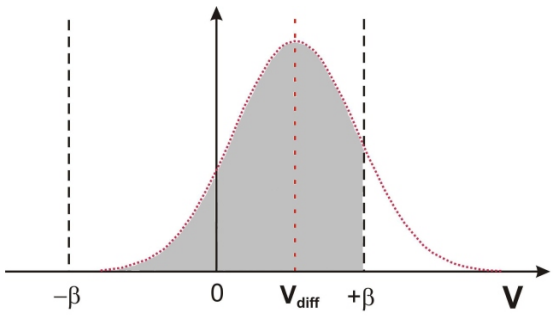

Figure 2.5: Example redshift probability distribution function of an object within the aperture around the central object with a fraction of it being within the velocity cut. The shaded part would thus contribute to the local density of the central object. Taken from Falter (2006).

varied in intensity, added to a power law continuum and multiplied by a redshift-dependent throughput function. The latter is modeling the absorption function of Hydrogen bluewards of the Lyman alpha line. Redshifts cover the range $0.504 \leq z \leq 5.96$ on a $\log(1+z)$ scale with a step size of 0.01 (155 steps in total). There are 20 different values covering the spectral index of the continuum and eight values of intensity relative to the mean template for the intensity of the emission line shape. As has been shown by COMBO-17, QSOs cannot be classified reliably using just broad-band filters. So far, the QSO library is not used in HIROCS, but medium-band data will become publicly available in the COSMOS field in the future and will enable reliable classification of QSOs in this field.

2.2.2 The cluster search algorithm

As has been discussed in the introduction, a variety of galaxy cluster search methods have been developed to date. When looking for high-redshift galaxy clusters, imaging methods have to introduce some way to increase the contrast of the clusters with respect to the contamination from fore- and background objects, for example by introducing color cuts to isolate the red sequence of cluster galaxies. However, any search method making use of intrinsic cluster properties will also introduce selection effects to the sample of clusters detected. The cluster search method which promises to suffer the least from selection effects is to search for clusters of galaxies purely as spatial overdensities of galaxies. The major drawback of this search method is that it is more prone to false positive detections because of projection effects or LSS.

HIROCS uses a cluster detection method based on searching for peaks in the local object density distribution. The goal of this search method is to isolate objects living in overdense regions, which then are selected as cluster candidates.

The local density for each object is computed by repeating the following procedure for each galaxy in the field:

- Based on the photometric redshift estimated by the multi-color classification pro-

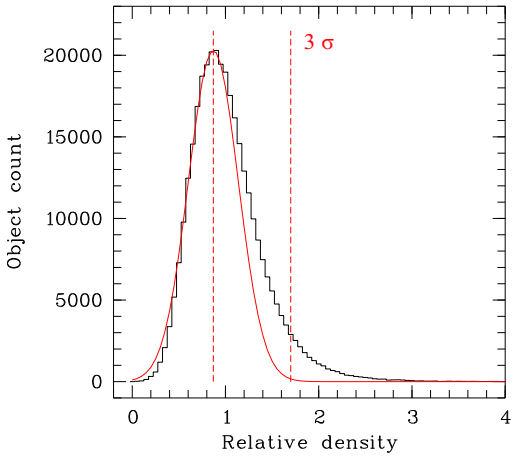

Figure 2.6: Distribution of local densities for an example field. A Gaussian is fitted to the rising flank of the distribution. Overdense objects with local densities above the 3σ cut trace the peaks in object density where the cluster candidates are to be found.

cedure, the projected size of an Abell radius (or a fraction thereof) on the image around the object is calculated.

- A cut in redshift [z_{low}, z_{high}] around this central object is defined based on a cut in velocity space. For each object within the aperture, its redshift distribution function, which is assumed to be Gaussian and based on the photometric redshift and its error, permits to calculate the probability of this object to live within the specified redshift range by integrating its area under the Gaussian between the redshift cuts (see Figure 2.5).
- These fractions of the redshift distribution functions for all objects within the aperture are summed and normalized to the average density of galaxies over the whole field at the redshift of the central object. The result, a single number which is a measure for the local density, is assigned to the central object.

The results are rather insensitive to the limits in velocity space. Currently, these limits are defined as $\pm 6000\, km/s$ (see Falter 2006).

Using a histogram of the local densities of all objects in the field, the overdensity cut for the cluster search is defined. For this, a Gaussian is fitted to the rising flank of the histogram, as shown for an example field in Figure 2.6. All objects having local densities of more than 3σ above the mean density in the field are used to search for galaxy clusters. The effect of the density cut on the example field is shown in Figure 2.7. The left panel shows the [x, y] distribution of galaxies with local density > 1.0, i. e. above the average field density. When applying the 3σ overdensity cut (right panel), cluster candidates become apparent as concentrations of the overdense objects.

This method for cluster detection has been successfully tested with CADIS and COMBO-17 data, see (Falter 2006). In the context of this thesis, a cluster selection algorithm based on friends-of-friends selection of the overdense galaxies has been developed, and

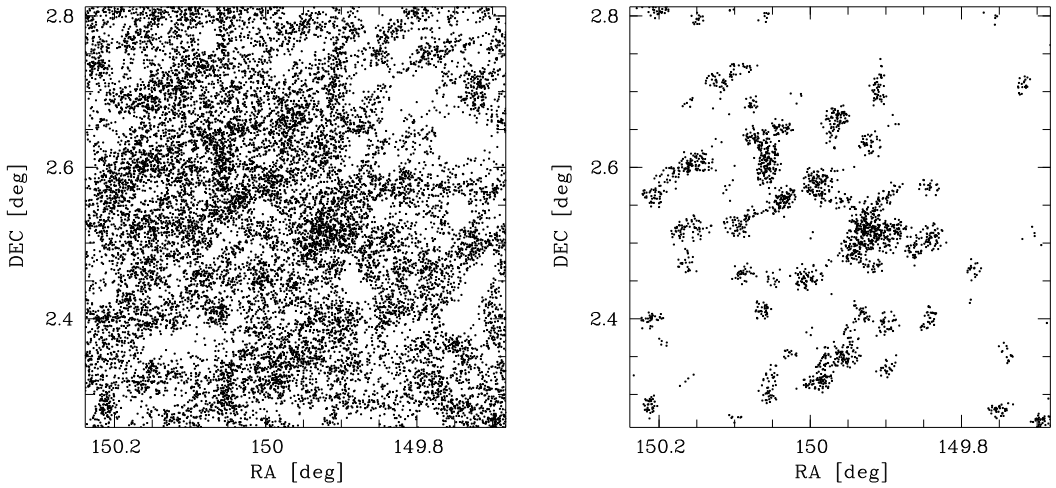

Figure 2.7: Effect of the overdensity cut on $z^+ \leq 25^{\mathrm{mag}}$ objects in a part of the COSMOS field north-west of the field center (compare to Figure 6.1). Left panel: Objects above the mean field density at $0.5 \leq z \leq 1$. Right panel: The same field after applying the 3σ overdensity cut.

the cluster selection function has been evaluated with mock sky data from galaxy evolution simulations (see Chapter 5).

3 The COSMOS survey, field and data

The HIROCS 10h field includes a large part of the COSMOS field and makes use of the extensive public data available from the COSMOS survey. In addition, proprietary H-band data has been taken in this field, permitting reliable photometric redshifts out to $z \approx 1.5$. This has been necessary because so far no deep near-infrared data has been provided by the COSMOS team. Initial efforts of HIROCS to contribute its H-band data to the COSMOS collaboration, thus becoming a partner, had to be discontinued because it could not be guaranteed that the H-band data would meet the COSMOS seeing requirements.

3.1 The COSMOS survey

The Cosmic Evolution Survey (COSMOS, Scoville et al. 2007c) is a large panchromatic extragalactic survey covering a single field on the sky with an area of about two square degrees. The extensive dataset which has been taken using all major observatories in the world covers wavebands from the X-ray to the radio regime will be used to address among others the following science goals:

- The assembly of galaxies, clusters and cold dark matter on scales up to $\geq 2 \cdot 10^{14}\,\mathrm{M}_{\odot}$
- The evolution of galaxy morphology, galactic merger rates, and star formation as a function of LSS and environment
- Reconstruction of the dark matter distributions and content using weak gravitational lensing at $z < 1.5$
- Evolution of active galactic nuclei (AGN) and the dependence of black-hole growth on galaxy morphology and environment
- The mass and luminosity distributions of the earliest galaxies, AGN, and intergalactic gas at $z = 3 - 6$ and their clustering

Subprojects of COSMOS detect galaxy clusters based on X-ray emission and red sequence detections using Spitzer IRAC data. The goal of HIROCS is to establish an orthogonal sample of galaxy clusters for this field using its own detection method. Comparison between these samples will enable to assess selection biases between the various cluster search methods, with important implications for future cluster surveys.

3.2 Public data

The following paragraphs give an overview of the public images and catalogs available from the COSMOS collaboration together with their corresponding science program, but cover only the datasets used by HIROCS. An overview of all COSMOS datasets available is given in (Scoville et al. 2007c). The data used for this thesis are taken from the first official data release (DR1) in May 2007. An overview of the optical and near-IR data used for the first official COSMOS catalog is given in (Capak et al. 2007).

3.2.1 Ground-based optical and near-infrared data

Deep optical data taken with Suprimecam on Subaru are available in the broad bands B_j, g^+, V_j, r^+, i^+, z^+ and the narrow band NB816 (Takahashi et al. 2007). Near-UV data in u^* has been acquired using the Mega-Prime cam on CFHT. A range of other, less deep optical data in similar or duplicate bands is also available, but was not used for HIROCS. Limiting magnitudes as given in (Capak et al. 2007) are listed in Table 3.1.

Of the ground-based near-infrared data that are being taken for COSMOS, the DR1 contains only Ks band data obtained with FLAMINOGS at the 4 m telescope at Kitt Peak National Observatory and ISPI at the CTIO 4 m telescope. After initial astrometric problems, these images are now usable without additional corrections (Version 3.0). The Ks data do not reach a depth comparable to the deepest optical COSMOS data taken with Subaru (see Table 3.1).

A detailed description of all the ground-based optical and near-IR data and the catalog derived from them is given in (Capak et al. 2007). The DR1 catalog is i^+-band selected with a limiting magnitude of $i^+_{AB} \leq 25^{\mathrm{mag}}$ and contains only objects in the central two square degrees of the COSMOS field, where multi-wavelength coverage is uniform. Due to the lack of deep near-IR data, the photometric redshifts are reliable only below $z \approx 1.1$. Details about the photometric redshift code used for the catalog and the photo-zs derived are given in (Mobasher et al. 2007).

Figure 3.1 shows the normalized filter transmission profiles for the public COSMOS optical and near-infrared wavebands together with the transmission profile for the proprietary HIROCS H-band.

The data in all filters (used by HIROCS) have been combined to a mosaic, resampled to a pixel size of $0.15''$ and point spread function (PSF) homogenized to get a uniform PSF full width at half maximum (FWHM) across the whole field (but different for each band). Table 3.1 lists the PSF FWHM of the data as given in (Capak et al. 2007). All images come in units of nJy/pixel. Information about the RMS are provided in separate images, also in units of nJy/pixel. Unusable areas are marked by the value -inf.

To provide easier data handling, all optical and near-IR mosaics have been divided into 144 tiles of 4096×4096 pixels with an overlap of 48 pixels on each side (see Figure 3.2).

In order to be able to use the images with the HIROCS reduction and photometry software, the tiles for each filter were reassembled to four quadrants. A single mosaic covering the whole field could not be used due to memory restrictions of the computer

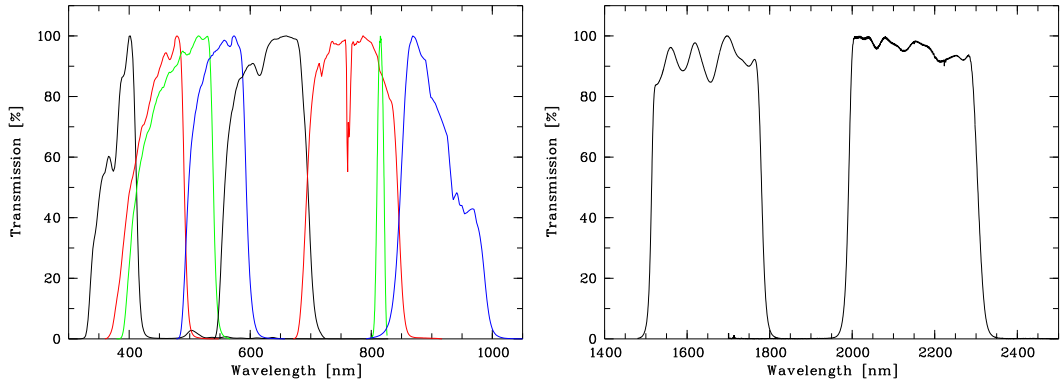

Figure 3.1: Filter transmission profiles for the COSMOS public optical data set in u^*, B_j, g^+, V_j, r^+, i^+, NB816, and z^+ (left panel) and the near-infrared data set with the public Ks band and the proprietary HIROCS H-band which is described in Section 3.3 (right panel). The profiles are normalized to a maximum transmission of 100% and take into account the atmosphere, the telescope, the camera optics, the filter, and the detector.

Table 3.1: PSF FWHM and depth of the COSMOS data release one images in all the optical and near-IR filters used by HIROCS as given by the COSMOS collaboration in (Capak et al. 2007). The depth of the data is at 5σ level given in AB magnitudes as measured for point sources on the unsmoothed images in an aperture of $3''$. The saturation limit for point sources is also given in AB magnitudes.

	u^*	B_j	g^+	V_j	r^+
PSF FWHM	$0.90''$	$0.95''$	$1.58''$	$1.33''$	$1.05''$
Depth [mag]	26.4	27.3	27.0	26.6	26.8
Saturation [mag]	15.8	18.7	18.2	18.7	18.7

	i^+	NB816	z^+	Ks
PSF FWHM	$0.95''$	$1.51''$	$1.15''$	$1.50''$
Depth [mag]	26.2	25.7	25.2	21.6
Saturation [mag]	21.8	16.9	18.7	10

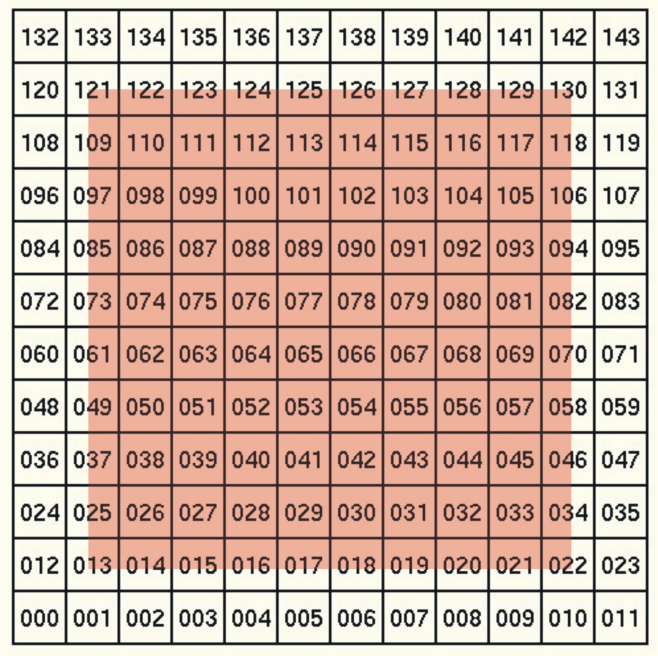

Figure 3.2: Tile layout of the public COSMOS optical and near-IR data. The area to be covered by HIROCS H-band data is indicated in red (see also Appendix E).

used for data reduction. Each quadrant is an image of $18\,000 \times 18\,000$ pixels, covering an area of $2\,700'' \times 2\,700''$. The pixel size remains unchanged at $0.15''$. At the borders, the four quadrants overlap by $60''$. These quadrants do not cover the whole area available in the 144 original tiles, but an area somewhat larger than the one covered by HIROCS H-band data.

3.2.2 Hubble Space Telescope ACS data

The Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS) data in the COSMOS field cover about 1.64 square degrees in the F814W filter, which corresponds to the i-band (Scoville et al. 2007a). Details about the reduction process are given in (Koekemoer et al. 2007). During acquisition of the ACS data, both the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) and the Wide Field Planetary Camera 2 (WFPC2) were used as parallel instruments to obtain data at 1.6μ covering 5% of the field as well as UV data covering about 40% of the field.

The ACS images were drizzled to a resolution of $0.05''$ per pixel, flux calibrated, astrometrically aligned, and combined to a mosaic. Weight map images are also available. In contrast to the optical and near-IR data, the ACS images come in units of counts. For convenience, the full mosaic is split up into 49 smaller tiles. Bad areas or blank pixels are marked by a weight map value of zero.

HIROCS uses these ACS data to obtain morphologies of the cluster candidate member galaxies. Since the images are not used to do photometry and thus do not have to be

used with the reduction and photometry pipeline, the ACS tiles are not recombined to the four quadrants like their optical and near-IR counterparts.

3.2.3 S-COSMOS Spitzer IRAC data

The S-COSMOS Spitzer survey (Sanders et al. 2007) will carry out a uniform deep survey of the full 2 square degree COSMOS field in all seven Spitzer bands (IRAC 3.6μ, 4.5μ, 5.6μ, and 8.0μ plus MIPS 24.0μ, 70.0μ, and 160μ. In DR1, the complete IRAC data plus a small MIPS test area have been made publicly available. HIROCS uses the IRAC 3.6μ, 4.5μ, and 5.6μ data. The 8.0μ data were included initially, but abandoned later (see Section 4.3).

Due to the lower resolution of IRAC compared to the optical and near-IR instruments, the IRAC images are provided as a single color corrected mosaic for each IRAC channel with a pixel scale of $0.6''$ and in units of MJ/sr. To convert the images into units of μJ/pixel, a conversion factor of 8.4616 has to be applied. Coverage maps in units of MJ/sr as well as uncertainty maps are also provided.

Unlike the optical and near-IR images, the IRAC mosaics fit into the memory of the computer used for reduction and photometry as a whole. Thus, object detection and photometry are done on the whole mosaic at once.

3.2.4 XMM X-ray data: XMM-COSMOS

The COSMOS field will be covered by a 2.1 square degree XMM mosaic with a total observing time of 1.4 Ms and a depth of 10^{-15}ergs cm^{-2}s^{-1} (Hasinger et al. 2007). The data of the first public release comprise the first 36 XMM pointings which completely cover the area of the COSMOS field at 57% of the final depth in all three energy bands: $0.5-2.0$ keV, $2.0-4.5$ keV, and $4.5-10.0$ keV. Apart from the background subtracted images, there are also background subtracted and exposure corrected images in units of MOS1 counts/second, the corresponding error files, exposure maps in units of effective MOS1 seconds, and wavelet reconstruction images in units of MOS1 counts/second available. All images have a resolution of $4''$ per pixel.

Finoguenov et al. (2007) give a list of clusters in the COSMOS field based on detections in the $0.5-2$ keV band, which this thesis will use to compare the HIROCS cluster candidates with. Initially detected by diffuse emission in the X-ray images, clusters are selected only if a peak in the early-type galaxy distribution based on the optical / near-IR catalog is found at the same position. As shown by Ostriker et al. (1995), a combination of diffuse X-ray emission plus a photometric overdensity of galaxies at the same position are sufficient to prove that this galaxy overdensity is a virialized structure. Wavelet filtering is used here to enhance the contrast of the member galaxies with respect to the field. The redshift of the candidates is determined from the photometric redshifts of the member galaxies. The X-ray selected cluster list, limited in redshift to $z < 1.3$ by the photometric data, comprises 72 clusters with redshifts $0.08 \leq z \leq 1.25$.

3.2.5 VLA radio data: VLA-COSMOS

The Very Large Array (VLA) was used to cover the entire 2 square degree COSMOS field at 20 cm with a resolution of $\sim 1.7''$ and a sensitivity of $\sim 8\,\mu$Jy ($1\,\sigma$) in the VLA-COSMOS survey (Schinnerer et al. 2007). Observations were done for 240 hours in the VLA A-array configuration and for 24 hours in the VLA C-array configuration.

Flux calibrated, astrometrically correct images are available in SIN as well as TAN projection. HIROCS makes use of the radio source catalog compiled from these observations to identify radio sources in the cluster candidates. The radio source catalog contains 3 643 reliably measured sources down to $50\,\mu$Jy.

3.2.6 zCOSMOS spectroscopic data

zCOSMOS is a spectroscopic survey in the COSMOS field using VIMOS at the VLT (Lilly et al. 2007). The survey consists of two parts: zCOSMOS-bright, obtaining spectra for a magnitude-limited sample of about 20 000 $0.1 \leq z \leq 1.2$ galaxies with $I_{AB} < 22.5^{\mathrm{mag}}$ over the whole 1.64 square degrees of the COSMOS field with ACS F814W coverage, and zCOSMOS-deep, a survey of about 10 000 galaxies in the central 1.0 square degree of the COSMOS field selected by color criteria to have redshifts $1.4 \leq z \leq 3.0$.

The first zCOSMOS public data release (zCOSMOS DR1) contains 1264 zCOSMOS-bright spectra. Of these, 910 have a quality flag of 3 or 4 (very reliable), permitting their usage as a comparison sample for HIROCS photometric redshifts. Cross-correlation with the HIROCS object table shows that out of these 910 objects, 862 objects can be identified with HIROCS objects. This comparison sample covers a redshift range of $0 \leq z \leq 1.33$, with the majority of objects being at $z < 1.0$. Unfortunately, most of the zCOSMOS DR1 objects lie in areas where there is no H-band data at all, or it is very shallow (see Figure 3.3 and Figure 3.4 for information on the depth of the H-band mosaic).

3.2.7 Other spectroscopic data

One zCOSMOS-bright and one zCOSMOS-deep mask have been reduced by S. Noll to provide a spectroscopic comparison sample before the official zCOSMOS data had been released (Gabasch et al. 2008). Of these 551 objects, 152 are identified as galaxies with reliable redshifts. Due to coordinate uncertainties, only 55 zCOSMOS-bright objects with $0.01 \leq z \leq 1.22$, and 16 zCOSMOS-deep objects with $2.06 \leq z \leq 2.63$ could be unambiguously identified with objects from the HIROCS catalog.

In addition, the sample of 336 Quasars candidates in the COSMOS field compiled by Prescott et al. (2006) also contains 184 galaxies. Of these, 139 objects with $0.01 \leq z \leq 0.91$ are in common with the HIROCS object catalog and are included in the comparison sample.

In total, reliable spectroscopic redshifts are available for an additional 214 objects from these two datasets. The total spectroscopic comparison sample used for this thesis thus comprises 1072 objects.

Table 3.2: Overview of the H-band data obtained to date for the COSMOS field.

Observing run	Date	# of science images	seeing
230X	ALHAMBRA team, 02/2005	256	0.8 - 3.8
230	23.03.2005 - 02.04.2005	375	0.9 - 2.2
232	03.04.2005 - 12.04.2005	537	0.7 - 2.1

3.3 Proprietary HIROCS H-band data

The HIROCS H-band data are taken with OMEGA2000 at the Calar Alto 3.5 m telescope in a joint effort of the large extragalactic survey projects ALHAMBRA (Moles et al. 2008), MUNICS-deep (Goranova et al., in preparation) and HIROCS. The COSMOS field will be covered by 25 OMEGA2000 pointings of $15.4' \times 15.4'$ for a total area of about 1.6 square degrees with a targeted depth of 21 mag (5σ, Vega) and seeing $< 1.6''$.

The H-band data available to date have been taken in three different observation campaigns in February, March and April 2005. For each images, 20 exposure of 3 seconds each are co-added in memory, resulting in images with one minute integration time. Assuming a seeing of $1.5''$, which is the upper acceptance limit for HIROCS, 50 frames are required per pointing to reach the desired depth.

Currently, the H-band data cover 1.3 square degrees (21 pointings) with a non-uniform depth (see Figure 3.3 and Figure 3.4 for the mosaiced sum image of the data and a weight map with depth information, respectively). Data quality in these 21 pointings is very inhomogeneous. Some exposure series are heavily affected by clouds, others fall short of the goal in the number of exposures taken, resulting in a depth shallower than envisaged. Of the 21 pointings with H-band coverage, 11 reach a limiting magnitude of at least $H = 21.3^{\mathrm{mag}}$ (3σ), deep enough to estimate photometric redshifts above $z \approx 1.1$. Thus, the area covered by them is suitable for searching high-redshift galaxy clusters. An overview of the observation campaigns is given in Table 3.2, with the distributions of the seeing shown in Figure 3.5.

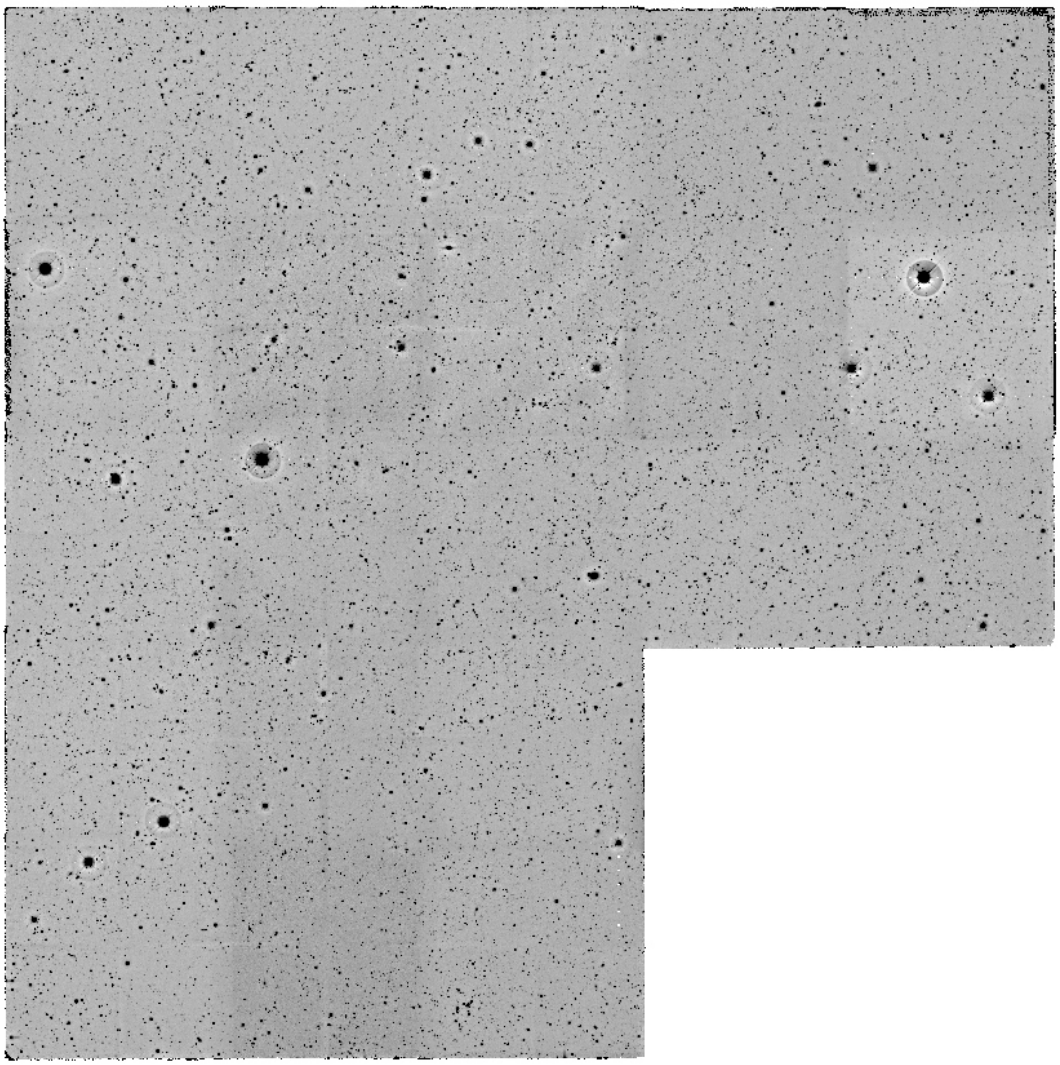

Figure 3.3: H-band mosaic of the COSMOS field. See Figure 3.4 for pointing numbers. In the SW corner (white area), no H-band data is available to date. The varying depth of the pointings is already visible in this mosaic, but is even more obvious in the weight map image (Figure 3.4). In pointings 1c and 2c, the background is not flat because passing clouds obscured the observations.

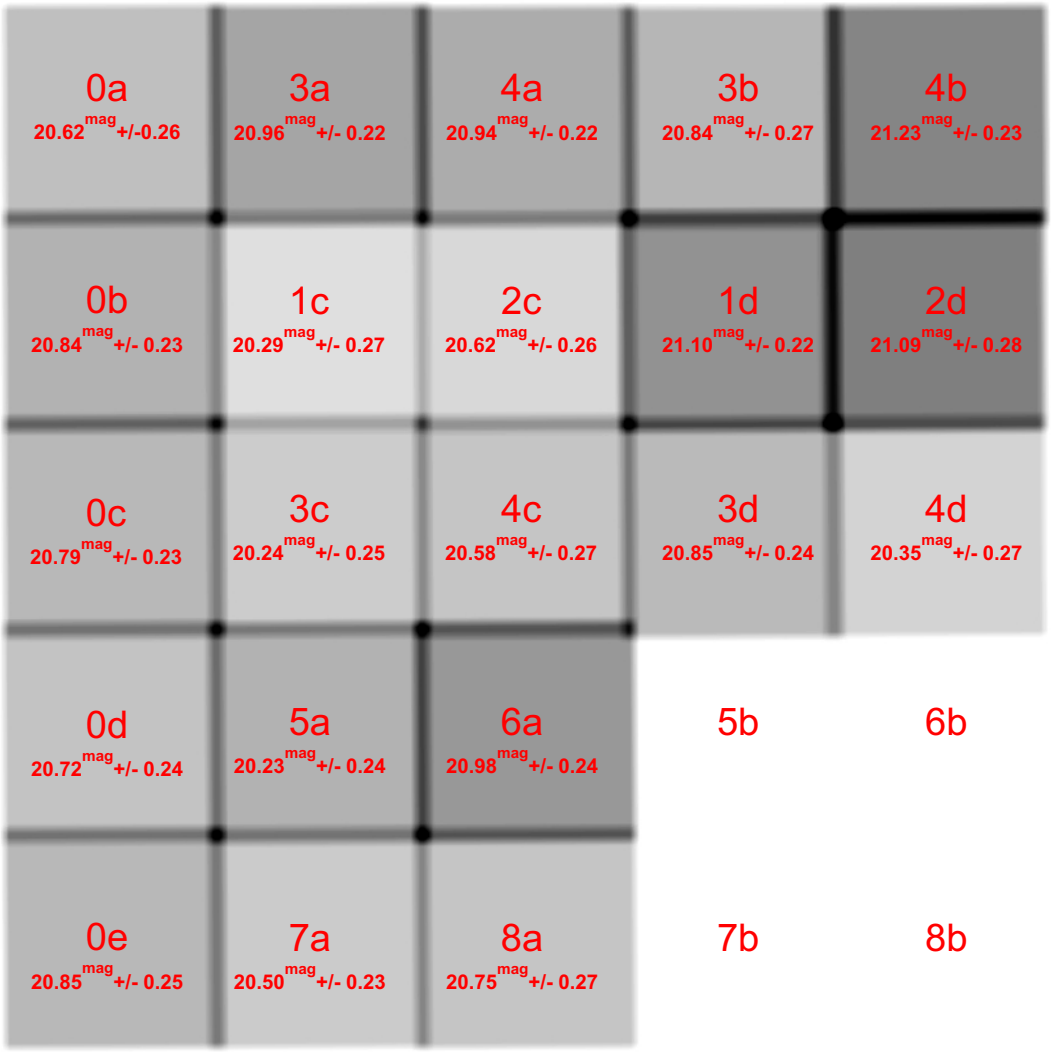

Figure 3.4: Weight map image of the COSMOS H-band mosaic shown in Figure 3.3. Pointing numbers and limiting magnitude (5σ) are indicated. Darker areas correspond to deeper pointings. In the overlap regions between the pointings, images from all adjacent pointings are summed up, resulting in stripes which go deeper than single pointings alone.

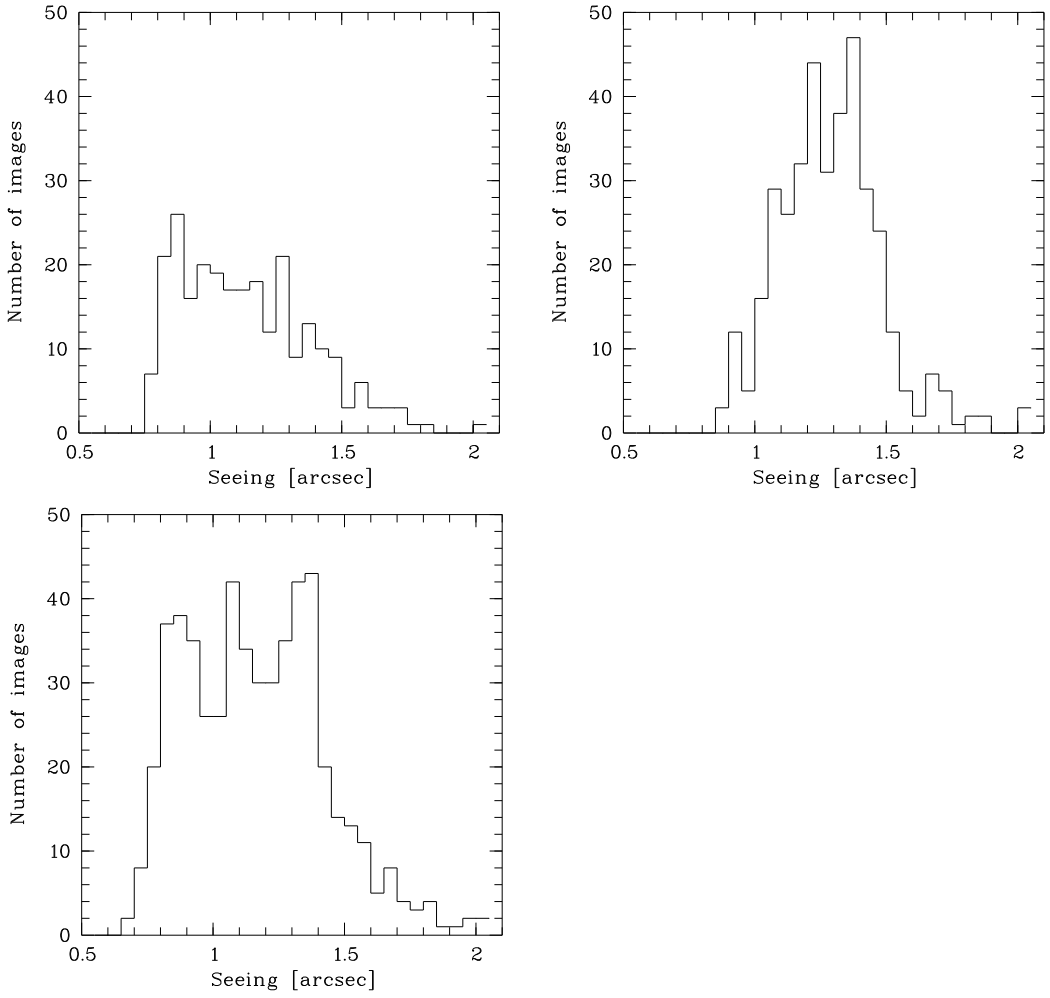

Figure 3.5: Distribution of seeing on the single images of the observing runs 230X (top left), 230 (top right), and 232 (bottom).

4 Data analysis

4.1 Integration of COSMOS public optical and near-IR data into the MPIAphot software framework

As described in Section 3.2.1, the public COSMOS DR1 optical and near-IR images in u^*, B_j, g^+, V_J, r^+, i^+, NB816, z^+, and Ks come as a set of 144 tiles with 4096×4096 pixels and a pixel size of $0.15''$ both for the science data and the RMS maps. To be able to integrate the data more easily with the existing HIROCS data reduction and analysis pipeline MPIAPHOT, these tiles are being reassembled to larger mosaics. In principle, one single mosaic for each filter would be favorable. Instead, with images of $18\,000 \times 18\,000$ pixels being the maximum to fit into the memory of the computer used for data analysis, a layout using four quadrants with the maximum image size, corresponding to an area on the sky of $2700'' \times 2700''$, is adopted. The total area covered by these quadrants is bigger than the area covered by the H-band data, but smaller than the area covered by all 144 tiles.

Care has to be taken of the image levels. The COSMOS images, both science and RMS, are in units of nJy. Bad pixels or areas without coverage are indicated by pixel values of -inf. Since the mean background value is zero, negative pixel values occur due to background noise in the science images. Since negative values cause problems with the MPIAPHOT photometry program EVALUATE, an offset has to be applied to the science images.

The usual way to measure the flux error of an object using MPIAPHOT is to derive it using error propagation on photometric measurements of this object and the background contribution on a series of images. For the COSMOS data, the object flux is measured on the science images, while the error of the flux is measured on variance images created by squaring the RMS images. Details about this are given in Section 4.5.1.

In order to prepare the data to be used with MPIAPHOT, the following steps are taken:

The RMS images are squared to convert them into variance images. The flag for bad pixels, initially set to -inf, is replaced with the internal MPIAPHOT broken pixel flag. Finally, all descriptors[1] needed by MPIAPHOT are written.

The science images are scaled by a factor of 10 000, and an offset of 100 000 is applied to ensure all values are positive. All -inf values (representing *e.g.* saturated pixels) are replaced with the MPIAPHOT broken pixel flag. The information about broken pixels from the RMS images is copied to the science frames in addition. All descriptors needed by MPIAPHOT are written in the final step.

[1] A descriptor is a variable associated with the image, for example the exposure time. In astronomical image formats, descriptors are stored as parts of the images.

4.2 Reduction of proprietary H-band data

The H-band data taken with OMEGA2000 at Calar Alto Observatory stem from three observation campaigns in February, March and April 2005.

Besides the science frames, sets of calibration images are available for each campaign, including:

- Series of dark exposures with increasing integration times
- Series of dome flatfields with increasing integration times
- Series of dusk and dawn sky flatfields

These calibration images are used in the process of reducing the science frames.

In total, 1104 one-minute H-band exposures in 21 pointings are available for this analysis.

4.2.1 Creation of bad pixel masks

Bad pixels have count levels which are not well-related to the input signal. Their output cannot be corrected, thus they have to be excluded from further analysis. Possible defects include:

- "Hot pixels" which always give a high readout value.
- "Dark pixels" whose readout value is always zero.
- Pixels with a non-linear photon response curve.
- Pixels with a non-linear dark current.

Non-correctable errors can also be caused by the readout electronics. This results in a "bad column" in the direction of a channel. All these defective pixels, the so-called "bad pixels", have to be determined and stored in a bad pixel mask. While some pixels are constantly bad over time, others may change properties between campaigns. For example, changes in temperature, especially when warming up the camera after dismounting it from the telescope and cooling after remounting it for the next campaign, can cause pixel defects to change. Because of this possible change in behavior, one bad pixel mask (BPM) has to be created individually for each campaign.

Good pixels show a linear increase of their dark current, as well as their count level when illuminated, with integration time. Series of dark exposures and evenly illuminated flatfields, both with increasing integration time, are used to test this. For both exposure series, a linear fit over integration time is calculated for each pixel of the detector. The fit parameters are then used to define bad pixels. A bad pixel has either

- a response curve slope of ≤ 0
- a RMS off by more than 5σ (10σ for the flatfield series; scattered light prevents the use of more conservative values here) from the mean RMS of the detector

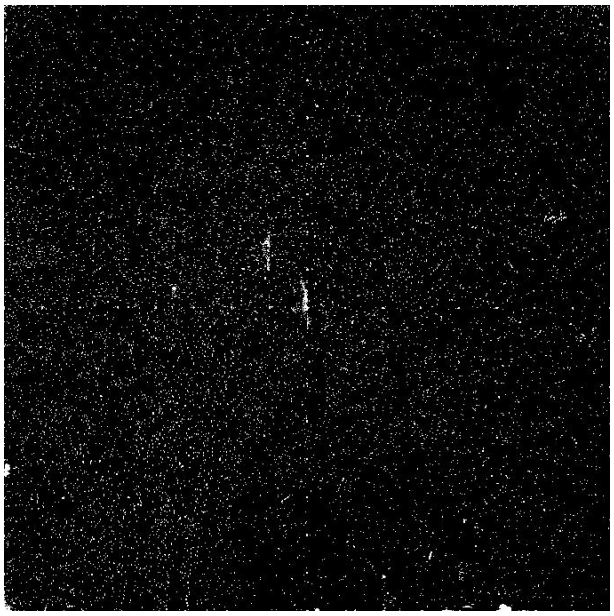

Figure 4.1: Example of a bad pixel masks for OMEGA2000 containing about 20 000 pixels flagged as "bad".

- a zeropoint off by more than 5σ (7σ for the flatfield series, again due to scattered light) from the mean zeropoint of all pixels.

Care has to be taken here of the pixels close to the amplifiers on the chip. These pixels get heated more than others and would show as defective, even though their behavior is correctable. They have to be treated with extra caution to not be included as bad pixels; the added signal is removed later in the sky subtraction step.

In total, the bad pixels masks for this instrument contain roughly 20 000 (0.5%) bad pixels. An example BPM for the OMEGA2000 camera is shown in Figure 4.1.

4.2.2 Creation of flatfields

Flatfield images model the imperfect response of the entire optical system (telescope, mirrors, lenses and instrument with the detector). They are taken using even illumination of the telescope. The resulting flatfield image is not flat, but uneven due to a number of effects:

- Vignetting
- Dust on lenses or filters
- Global quantum efficiency variations of the detector
- Fixed pattern noise: Pixel-to-pixel sensitivity variations in the detector

The effects included in a flatfield are multiplicative, in contrast to *e.g.* the additive component of stray light. Due to this, division of the science images by the flatfields corrects for the effects listed above.

There are several different ways of obtaining flatfields:

- Using a screen inside the dome, illuminated by lamps. These are called "dome flatfields". In the case of the Calar Alto 3.5 m telescope, dome flatfields cannot be used to correct science images because of stray light and dome structure visible in the images.
- Using the evenly illuminated sky at dusk or dawn, resulting in the so-called twilight flatfields. This is only possible without clouds, and care has to be taken of brighter stars showing up in the flatfields, for example by dithering between the exposures.
- They can also be extracted from the background of regular science exposures. These "science flatfields" need a large number of dithered input frames from different pointings to be able to exclude bright objects and in order to reach a sufficient signal to noise ratio. Sometimes they are called "sky flatfields", but here the term "science flatfield" will be used to prevent confusion with the twilight flatfields, which are also called sky flatfields by some. Science flatfields provide the best correction, if they have been created cautiously.

The response of the OMEGA2000 detector varies considerably on the pixel to pixel noise level between different observing campaigns. Because of this, flatfields have to be created for each campaign individually. In the process of the creation of master flatfields for the H-band data, various flatfield types are being used.

First, a master dome flat is created by subtracting the sum of all [dome flatfields taken with the lamp off] from the sum of all [dome flatfields taken with the lamp on]. This automatically corrects for the dark current. The sum image is normalized with respect to the mean level in its central part, resulting in the master dome flatfield. As mentioned above, this dome flatfield cannot be used to correct science images. However, it is playing an important role as a "catalyst" in the creation of the twilight flatfields, as described in the next paragraph.

The near-infrared twilight flatfields have count levels and overall structure from image to image. Unlike the dome flatfields, they cannot simply be summed up. Instead, a clipped averaging procedure is applied. To fully remove objects, a small-area filter is applied, and outlier values due to stars or cosmics are removed.

There are two problems with simply applying this procedure to near-infrared images:

- The shape of the flatfield itself is uneven due to spatial sky background variations. Bright or faint parts which are too far off from the average values could be clipped away.
- The high pixel-to-pixel variation artificially increases the noise, making the removal of object halos with the outlier filter impossible.

To avoid the latter problem, all twilight flatfield images are divided by the master dome flatfield. The resulting images are "flat" to first order, with both the overall distortion and the high pixel-to-pixel variation removed, but still show effects which can not be corrected with the dome flatfield alone. In the next step, these images are averaged using the clipped averaging procedure. The resulting averaged image is multiplied by

the master dome flatfield to cancel its effect out. Finally, normalization with respect to the central part of the image gives the master twilight flatfield.

While the master twilight flatfield is already suited a lot better than the master dome flatfield to correct the science image, the best flatfield is created out of the science images themselves. To obtain a sufficient SNR of the flatfield, the number of input images for the creation of the science flatfield has to be as large as possible. They have to stem from different pointings to enable reliable removal of objects in the images, because the amplitude of the dither pattern is too small to completely remove remnants of bright objects. Due to its superiority over the master dome flatfield, the master twilight flatfield created before is chosen to serve as the catalyst here. The algorithm for creating the science flatfield is as follows:

- The input images are taken from p different pointings. The number of images i is the same for all pointings.
- Model the overall shape of the background for each input science image and store it in a separate [back] file, with the mean image level subtracted. The difference between the input image and the background is stored in [flat] files: input = [flat] + [back] + mean image level.
- Divide each [flat] image by the catalyst (the master twilight flatfield).
- Repeat the following steps i times (as often as there are images per pointing):
  - Take a group of p [flat] images, one image per pointing. Compute a clipped average [average_i] of them.
  - Multiply the average image with the catalyst to cancel its effect out.
  - Sum the corresponding [back] images to obtain a background profile of the average image: [average_i_back].
- In the end, compute a clipped average of all [average_i] frames. Add the sum of all [average_i_back].
- Normalization of the resulting image gives the master science flatfield.

Tests show that when the image resulting from the division of a raw frame by the master science flatfield has noise levels which are $20\% - 30\%$ above the value expected from photon noise statistics. While not perfect, this value is acceptable for a master flatfield. As an example, the master science flatfield for the observation campaign in April 2005 is shown in Figure 4.2.

4.2.3 Pipeline reduction of the images

The reduction of the H-band images is done with the OMEGA2000 data reduction pipeline developed by René Faßbender (Faßbender 2007, and references therein). In its latest version, the OMEGA2000 pipeline features

- Flatfield correction
- Sky modeling and subtraction

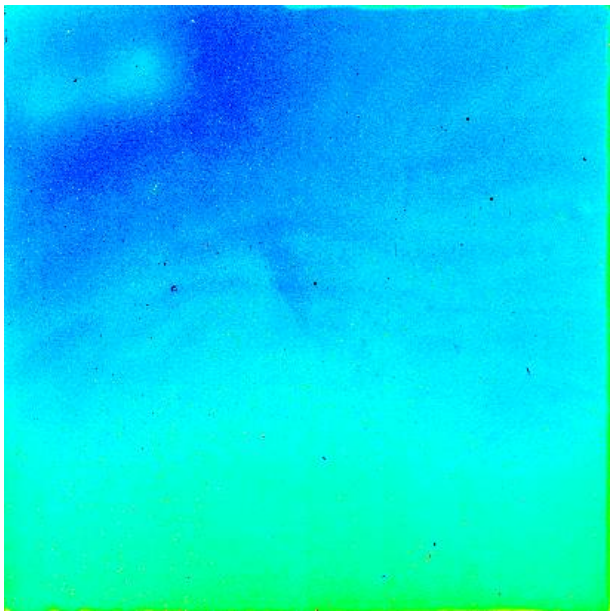

Figure 4.2: Master science flatfield for OMEGA2000 from the observation campaign in April 2005. The non-uniform sensitivity of the detector is readily visible. The green areas of the detector are about 126% the sensitivity of its central part, while the sensitivity of the dark blue regions is only about 86% of the central sensitivity.

- Bad pixel correction
- Object mask creation
- Summation with cosmic hit correction using fractional pixel offsets

The images are reduced separately for each campaign, using the according master science flatfield and the bad pixel mask.

Infrared images have very high background levels that show strong spatial and temporal variations. The uncertainties in the background can be larger than the flux of astronomical objects, making an accurate modeling of the sky contribution to the object flux essential.

To prevent the background variations from disturbing the sky modeling of the OMEGA2000 reduction pipeline, all images are background flattened before entering the pipeline. Before this, the images are divided by the flatfield, so spatial changes in the sky background are the only sources of variations. The shape of the background is fitted by a polynomial on a x, y grid and subtracted from the image. The flux in the image is preserved by adding the mean count level of the background after subtraction of the model. In the final step, the images are multiplied by the flatfield again, thus canceling its effect out.

In the OMEGA2000 pipeline, the sky signal is extracted from a series of seven dithered images bracketing the respective image. These images are stacked in pixel space, so each pixel column in the resulting data cube stems from the same pixel of the detector.

Because of the small offsets in telescope pointing direction between the single images due to the dither pattern, the same pixel on the detector will target different parts of the sky in subsequent images. This means that most of the pixels in one column of the data cube will show the sky signal even if object flux is included in some of them. To exclude the object flux possibly showing up as outlier values in these pixel columns, the sky signal of the respective pixel is taken to be the median of the values of each pixel column.

Pipeline reduction is done in three steps. The first run of the pipeline is done with standard parameters, resulting in a series of flatfield and bad pixel corrected, sky subtracted images. In the vicinity of bright, extended objects larger than the dither pattern, the median-based sky model is distorted. Thus, objects above a SNR cut are masked out using an object mask which is created in the second step. In the third step, this object mask is used to obtain optimal sky subtraction for the final reduction of the images.

In principle, the reduction pipeline also corrects the bad pixels by replacing them with a value obtained through linear interpolation between the count values of neighboring pixels. This method is not ideal to correct for bad pixels inside of objects. For the COSMOS field H-band data, another correction scheme was adopted: The bad pixels are treated just like cosmic ray hits, as is explained in the next section. In preparation for this, all bad pixels in the pipeline output images are replaced with a value close to the saturation value using the position information of the BPM.

4.2.4 Astrometry

All pipeline output images are resampled to a pixel size of $0.3''$ and stored in sky coordinate gnomonic projection using the field center as reference and stars from 2MASS (Skrutskie et al. 2006) to determine the transformation matrix. In the case of OMEGA2000, no undistortion step is needed because the instrument does not have a measurable optical distortion.

4.2.5 Cosmic ray hit removal

The OMEGA2000 reduction pipeline is also capable of correcting cosmic ray hits ("cosmics") by replacing them with a median value determined from a series of images. In contrast to this, bad pixels in the OMEGA2000 pipeline are corrected by replacing the affected pixels with interpolated values from the neighboring pixels of the same image, thus altering their flux. In principle, bad pixels could be treated in exactly the same way as cosmics. Because of this, cosmic ray hits are corrected independently from the reduction pipeline using MPIAPHOT routines. In preparation to this step, the values of bad pixels are boosted to resemble cosmics using the information stored in the BPM. This way, they are identified and treated just like cosmics and will not be mentioned separately below.

In MPIAPHOT, cosmics are removed from the single images by comparing them with a median image. Outlier values are replaced by the median count level at the corresponding position on the sky. To be able to correct cosmics on the edges of the single

pointings, where the number of images from the respective pointings is not large enough to compute a reliable median, the median image is a mosaic of all pointings in the field. This way, the median can be computed for all edges inside the field due to the overlap between the pointings.

For creation of the median mosaic, the very inhomogeneous seeing values of the single images have to be taken into account. Stellar objects in images with a very good seeing could be wrongly treated like cosmic ray hits if compared to a median image with a bad seeing. In case of the COSMOS H-band data, not only is the seeing of different pointings very different, but also the seeing values of single images from one pointing can vary substantially. To control this effect, all images are divided into two groups with a seeing cut at $1.2''$, and two separate median mosaics are created.

This gives rise to two issues that need to be taken care of. First, the images of some pointings with a strong intrinsic variation in seeing are split up. Some images go into the $< 1.2''$ seeing bin, the rest falls into the $> 1.2''$ seeing bin. It can be the case that one or neither of the two seeing bins contains enough images for this pointing to give a SNR high enough to allow for computing a reliable median. Second, the pointings which are in the same seeing bin do not necessarily lie adjacent to each other, resulting in a patchy median image with holes. Cosmic removal is not possible on the edges of the median image, due to the same reason as for the edges of single pointings. If edge areas in addition to the unavoidable ones at the field borders are introduced, the area without cosmic correction is increased unnecessarily.

Both of these issues can be avoided by adding a sufficient number of images from the other seeing bin. They need to fill up the holes between the pointings to allow for cosmic correction on the fringes, and add enough images to perform the computation of the median in each pointing that otherwise does not contain a sufficient number of frames.

With the median mosaics in place, cosmics are corrected in the single images by comparing their pixel values with the median value at the same astronomical position, using $\kappa - \sigma$-clipping to determine the outliers and replacing them by the median value. The surroundings of a corrected pixel are checked in a second iteration using a lower κ value, and also corrected if necessary.

In the final step, cosmic ray hits are corrected in the original frames using the information from the image in gnomonic projection.

4.2.6 The full mosaic and subsums for photometry

For the object detection in the H-band, a mosaic of the whole field is created as a weighted sum of the gnomonically projected images. Weighting of the image is done according to the formula

$$\frac{1}{\text{transmission} \cdot \text{backgr.noise}^2 \cdot \text{PSF}^2}. \tag{4.1}$$

A corresponding weight map is also produced in this step. The H-band mosaic and its weight map are shown in Figure 3.3 and Figure 3.4, respectively. The inhomogeneous

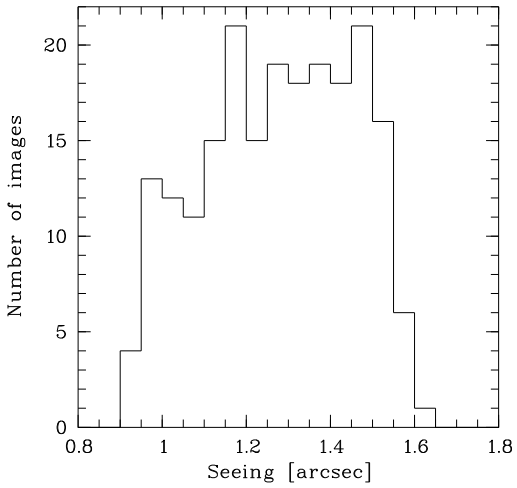

Figure 4.3: Seeing distribution of the H-band subsums in the COSMOS field.

depth of the data is visible already in the mosaic, with shallower pointings showing increased background noise. Some of the pointings also suffer from clouds obscuring the observations, with not even the elaborate background correction procedures described above being able to fully correct their influence. Data depth is shown as shades of gray in the weight map image, with darker areas on the image corresponding to deeper data. The layout of the pointings is very apparent. The overlap regions between the single pointings, where images from all adjacent pointings contribute, go deeper than the areas where images from only one pointing are summed up.

Ideally, photometry should be done on the single images to obtain better error measurements. Due to the short exposure time in the H-band, there are too many in the whole field to reasonably handle all of them. For this analysis, a hybrid approach is chosen: photometry is done on subsums of five images each.

The OMEGA2000 pipeline software is used for co-addition of the single exposures. Exploiting its capability to use fractional pixel offsets and weights during the summation process results in increased depth of the subsums compared to straightforward summation of the images in pixel space. But even with the increase in image sharpness compared to summation using integer pixel offsets, the PSF of the subsums is somewhat larger than the PSF of the single images used to create them. If the PSF of the input images is close to the HIROCS limit of $1.5''$, it can be the case that the resulting subsum has a PSF above this threshold and has to be excluded from further analysis. The 211 H-band subsums in the COSMOS field have PSFs ranging from $0.93''$ to $1.62''$. Two of them have to be excluded because their PSF is above the threshold, leaving 209 subsums for the photometry. The seeing distribution of the subsums is shown in Figure 4.3.

In principle, the number of cosmics and bad pixels in a sum of n images is about n times the number of bad pixels in one of the input images (the exact number is slightly different due to the slightly reduced area of the subsum compared to a single images which is a result of the dither pattern). However, because of the procedure used in

this analysis to correct the bad pixels, namely to replace them with the median pixel value at the same astronomical position, their error compared to the "true" value is very small. Furthermore, also because of the dither pattern, each bad pixel contributes only about $1/n$ of the flux on the subsum, with the exact value depending on the relative weights of the n pixels. Overall, it is safe to assume that the bad pixels do not alter the results of the photometry in any significant way. Because of this, no separate cosmic and bad pixel frames are calculated for the subsums.

4.3 Integration of the Spitzer IRAC data

For this thesis, only the 3.6μ, 4.5μ and 5.8μ IRAC data are used; the 8.0μ band was abandoned after initial tests because of little information gain relevant for the cluster search, but at the price of a common PSF (see Section 4.5.1) broadened by $0.2''$, resulting in a lower limiting magnitude for all other bands. In contrast to the other COSMOS mosaics, there are no pixels flagged as being unusable or saturated in the science or uncertainty images.

In a first step, the astrometry of the IRAC mosaics is checked against the 2MASS point source catalog . The MPIAphot find/obj routine is used to extract the brightest sources of the IRAC mosaics, resulting in 174 bright objects in common with the 2MASS catalog. The RMS of the scatter is about $0.23''$ for both the 3.6μ and the 4.5μ mosaics, comparable to the position scatter RMS of the H-band data and limited by the accuracy of 2MASS.

The determination of the PSF for the IRAC images turns out to be more complicated. The find/obj routine is also capable of fitting the PSF by stacking cutout images for the brightest objects in the field. On the resulting image, the PSF is determined very accurately, making use of the excellent SNR the summation of the bright objects provides.

In a first test, the PSF for the 3.6μ image is calculated to be $2.4''$. The corresponding cutout stack image has a somewhat asymmetrical appearance (Figure 4.4, left panel). Closer inspection shows that this distortion and the resulting broadening of the PSF is due to the brightest stars being saturated, but not flagged as such (Figure 4.4, right panel). If they are excluded from the analysis, the PSF FWHM is measured to be $1.8''$, which is the value given by the IRAC manual. For the 4.5μ mosaic, the PSF is determined to $1.75''$ when taking into account the saturation effects, and the width of the 5.8μ PSF is $1.95''$.

As preparation for the following photometry, subimages matching the size of the optical and near infrared COSMOS images are cut from the IRAC 3.6μ, 4.5μ, and 5.8μ mosaics. The subimages are 7268×7268 pixels large with a pixel size of $0.6''$. The total area of one subimage is about 1.47 square degrees. Next, these images are converted to units of nJy by applying the conversion factor of 8461.1. Analogous to the other COSMOS mosaics, an offset is added to the science image to avoid negative values in the background caused by noise, and the uncertainty frames are squared to provide variance information. After these steps, the images can be integrated smoothly with the other data and analyzed with the MPIAPHOT package.

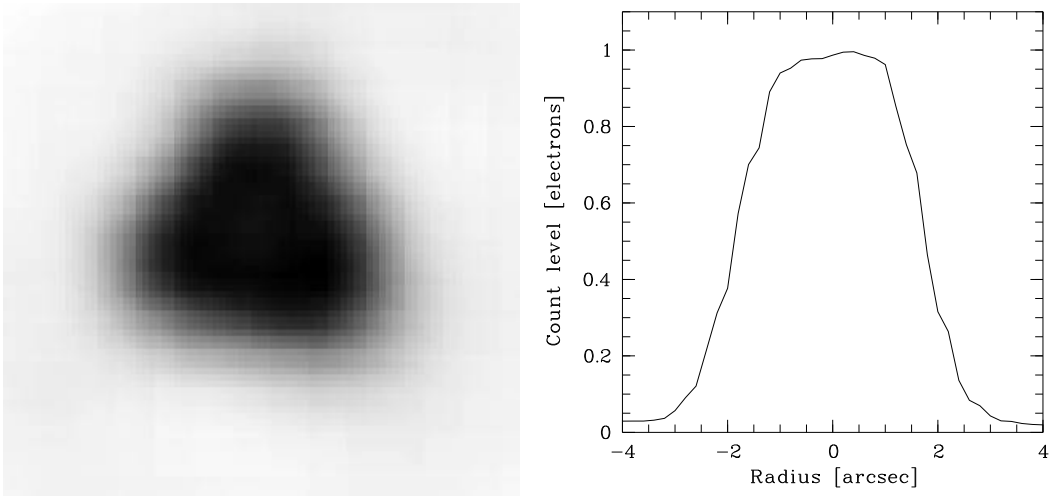

Figure 4.4: Left panel: Cutout stack images for brightest objects in the IRAC 3.6μ band. Right panel: Corresponding profile.

4.4 Combination of public and proprietary data

4.4.1 Creation of a master catalog

A master catalog should ideally contain all objects in the field down to a well-defined detection limit, and no spurious objects in addition to them. For multi-color datasets, which are possibly inhomogeneous both in depth and PSF, the most common way to build a master catalog is to select objects on the sum image of the deepest band, as is done for example in (Capak et al. 2007). This method of creating a master catalog provides a well-defined selection function, however it suffers from several drawbacks: First, the need for a tradeoff between depth and source separation could arise if the deepest image is not the one with the best PSF at the same time. In the special case of the COSMOS dataset, the deepest band, i^+, is also the one with the best PSF, so this is not an issue here. Second, the usage of only one band introduces a color selection of objects in the catalog. Objects with special characteristics, like faint objects with strong emission lines outside the band used for selection, or dust-reddened objects which only show up in the near-infrared, will be missed by this method, depriving the multi-color dataset of one of its advantages. Adding sum images from several bands together and selecting objects on the resulting frame is a procedure sometimes used to avoid this drawback. The optimal method however is not adding object flux, but object SNR (Szalay et al. 1998); this is also used in MPIAPHOT.

First, objects are detected on the sum image of each band in gnomonic projection, which should contain as much integration time as possible, down to a well-defined SNR using the SEXTRACTOR software (Bertin and Arnouts 1996). The sum images are mosaics combining pointings with possibly different seeing properties, so the SEXTRACTOR smoothing function is applied with a kernel based on the seeing, which is averaged over the whole image, during object detection. In addition, a weight map is used to avoid problems arising from varying depth of the pointings. The resulting object catalog still

contains some spurious objects in the vicinity of bright stars, along diffraction spikes of stars, and also some objects with extreme PSF FWHM. These are flagged in the object table half automated / half manually. Automated cleaning involves the objects around the brightest stars, spurious objects with extreme PSF FWHM and those moderately bright stars which were assigned a wrong PSF. Spurious objects along diffraction spikes and some leftover objects that were not found by the automated object cleaning have to be flagged manually. The SExtractor setup files used for this analysis are given in Appendix C.

In the second step, the object catalogs of the different bands are combined. Here, the MPIAphot method differs slightly from the one described by Szalay et al. (1998). An artificial, noise-free image representing the whole field is created. Then, all unflagged objects from the different SExtractor lists created in the step before are inserted as Gaussians into this image at the corresponding object position, with the strength proportional to the SNR of the detection, and shape parameters according to the SExtractor results. If an object is detected in more than one band, the corresponding Gaussians in the artificial image overlap, resulting in an even more prominent object there. The position of this combined object is very stable against small uncertainties in object position in the input catalogs. The master catalog is created by running SExtractor on the artificial image with a special setup for noise-free data. Because of the special image and object properties, the recovery rate for the artificial objects in this run is 100%, with no additional spurious detections.

The resulting master catalog meets the goal of containing objects detected in different bands very well, however there is a point to keep in mind arising from a peculiarity of the combined public COSMOS / proprietary H-band dataset used for this thesis. Some of the optical bands go extremely deep, resulting in bright stars being bloated and covering large areas on the images. In the less deep bands, where the bright stars cover smaller areas, real objects are detected around them which are obscured by the bright star halo in the deep bands, resulting in distorted photometry for these objects. Based on their photometric properties, these objects can be excluded from the object table used for searching the galaxy clusters after the multi-color classification.

The HIROCS master object catalog for the COSMOS field is created from object catalogs extracted from the u^*, B_j, V_j, r^+, i^+, z^+, and H. Some of the available bands are not included in this process for various reasons. The Ks data is too shallow, so no additional objects to the ones already detected in H are to be expected. In addition, the seeing in Ks is bad, thus its inclusion would only broaden the objects on the master image and make object separation more difficult. g^+ was left out for a similar reason. While it goes very deep, its PSF FWHM is among the broadest in this dataset. Because of the very similar wavelength coverage of B_j, V_j and g^+, the gain in additional objects would be very little, while due to the high SNR of the g^+-band and its bad seeing the objects on the artificial master image would be broadened substantially. Finally, NB816 was left out because it is a very narrow filter not going deeper than the broad filters covering the same wavelength regime, so at most only a handful of very special objects not relevant for the cluster search would be added to the catalog. After removal of fringe objects which would have false photometric measurements because they do not completely lie on the mosaic, the HIROCS master catalog for the COSMOS field

contains 653 785 objects, covering an area of $5240'' \times 5240''$ or 2.1 square degrees on the sky. The mean positional accuracy is about $0.1''$.

4.4.1.1 Selection of point sources

A subcatalog containing only point sources is needed mainly for two purposes: The selection of stars for the photometric calibration of the single bands (see Section 4.6.1), and for the stellar main sequence shifting technique used to fine tune the photometric zero points (see Section 4.6.2.1). For these applications, the point source catalog does not need to be comprehensive, but should not contain too many contamination objects.

In the COSMOS field, point sources are selected from the master catalog following a rather simple procedure, based on their PSF properties in B_j, i^+ and z^+. B_j and i^+ are the two optical bands with the smallest PSF of $0.95''$ each. In addition, these two bands are the deepest of the whole dataset, thus ensuring a sufficient depth of the point source catalog. To include red objects which are required for the main sequence shifting of colors involving H and Ks, point sources are also selected based on the z^+ band.

The PSF determination is done by running the MPIAPHOT photometry program EVALUATE in a special setup for PSF fitting (see Appendix D.2 for the configuration file) on the images in the bands listed above with the master object catalog as input. EVALUATE assigns PSF properties to each object in the catalog based on fitting an elliptical Gaussian to the image. In the final steps, objects are marked as point sources if their PSF properties are consistent with those of a stellar object in one of the bands. Tests show that for almost all objects in the point source catalog, this is the case in all three filters used for their selection, as is to be expected for well detected objects. This test confirms that the point source catalog does not contain a lot of contamination objects.

4.4.2 Creation of a Spitzer IRAC 4.5μ selected object catalog

The red sequence galaxies of distant galaxy clusters primarily shine in the near infrared. Since the IRAC bands are not part of the HIROCS object table (see Section 4.7.1.2) and the H and Ks bands are quite shallow compared to the optical and IRAC bands, a second object catalog selected on the 4.5μ IRAC band is used to detect red sequence galaxies at high redshifts, thus providing additional information about the high-redshift cluster candidates.

As for the selection of the master catalog, SEXTRACTOR is used to extract the objects from the mosaic image. The setup files used for the 4.5μ image is given in Appendix C.2. The intermediate step involving the artificial noise free image is not needed in this case because the catalog is selected based on a single band only. The 4.5μ selected catalog contains 86 018 objects on an area of $4360'' \times 4360''$ or 1.47 square degrees on the sky. Allowing coordinate uncertainties of $0.4''$ in x and y, corresponding to 2σ of the scatter RMS of the 4.5μ object detection with respect to the 2MASS catalog, 66 711 of these 86 018 objects are in common, while 19 307 are unique 4.5μ detections.

4.5 Photometry

4.5.1 Photometry in the different bands

The MPIAPHOT program EVALUATE is used to derive instrumental magnitudes as weighted sums over the image area. This program is optimized to give good colors, not total magnitudes. The width of the weighting function is adjusted such that the measurements refer to a common Gaussian "aperture" for all images: $\sigma^2_{\rm common} = \sigma^2_{\rm weight} + \sigma^2_{\rm seeing}$, equivalent to the beam in aperture synthesis radio astronomy (Röser and Meisenheimer 1991). Images with narrow seeing FWHM are analyzed with a broad weighting function, and vice versa. This procedure ensures that object colors are correct even if the images were taken under different seeing conditions. The Gaussian aperture has to be somewhat larger than the FWHM of the PSF of the images with the worst seeing to avoid that the weighting function becomes too narrow. An extremely narrow weighting function would result in incorrect color measurements even for small positional uncertainties which could arise *e.g.* from different object shapes in the various wavebands due to morphology, like prominent, off-center star formation region visible preferably in the bluer bands. Because of this, the images with the worst seeing determine the width of $\sigma_{\rm common}$, requiring to apply an upper limit for the seeing of the input data ($1.6''$ for this analysis).

For a reliable error determination, photometry is preferably done on the single science exposures, or on subsums thereof in case the total number of single images is too large to handle, as is the case for the COSMOS H-band data. In case of the public COSMOS DR1 data, only full mosaics are available, with additional variance images providing information about the errors. Here, the photometry scheme is slightly different. Object flux is measured on the science mosaics, whereas the errors are measured on the variance mosaics using an identical photometry setup. Both measurements are combined afterwards.

Photometry for the COSMOS field is done three times:

1. Using the HIROCS master catalog, without the IRAC bands
2. Using the HIROCS master catalog, including the IRAC 3.6μ and 4.5μ bands
3. Using the 4.5μ selected catalog, including the IRAC 3.6μ and 4.5μ bands

4.5.1.1 Photometry without IRAC bands

In the case of photometry for the HIROCS master catalog without the IRAC bands, the upper limit in seeing has been chosen to be $1.6''$. Here, the width of the Gaussian aperture described above is chosen to exceed the worst seeing FWHM by about half the pixel size of OMEGA2000, the camera with the lowest resolution used for HIROCS, resulting in a value of $1.8''$.

4.5.1.2 Photometry including IRAC 3.6μ and 4.5μ bands

Photometry including the IRAC 3.6μ and 4.5μ bands is essentially the same as without, the only difference being that the width of the Gaussian aperture has to be increased to a value of $2.3''$ because of the broader PSF of the IRAC images and their larger pixels.

4.6 Creation of the flux table

The flux table contains all photometric measurements on the single images and is the input for the multi-color classification. It is created in a two-step process. First, flux tables for the individual filters are created. These are combined in the second step to obtain the final flux table containing all flux measurements.

For each band, the best image from each pointing is chosen as the "normalization image". The mosaic correction is applied only to these normalization images. Bright stars from the 2MASS catalog are used as reference, and the magnitude offsets between the catalog values and their measurements on the normalization images are used to bring all normalization images to the same level, necessary for combining the pointings to a single mosaic. Note that at this step, this is only a relative, not the absolute flux calibration. All measurements from the other images of the respective pointing are calibrated with respect to the measurements on the normalization image. For this relative calibration, a set of normalization stars is chosen on each image. Averaging over their flux fractions with respect to the normalization image using a $\kappa - \sigma$ clipping method determines the relative calibration factors.

For the public optical COSMOS data, the steps listed above are simplified because the images already come as flux calibrated mosaics. Since the four quadrants in each band are cutouts from the same mosaic, they need not be mosaic corrected, and with only one image per quadrant and filter, no relative calibration with respect to a normalization image is needed. The procedure used to obtain the absolute flux calibration is described in the following subsection.

The IRAC images can be treated in the same way as the optical COSMOS data as they too are flux calibrated mosaics. Data handling is even further simplified because of their larger pixel size, the IRAC mosaics can be held in memory as a single image and need not be split into quadrants.

4.6.1 Absolute calibration of the data

MPIAPHOT requires standard stars with known flux across all bands to calibrate the instrumental magnitudes derived by EVALUATE. For HIROCS, the standard way is to obtain optical spectra of two stars in each of the four corners of a field for this purpose. Their spectra are then folded with the instrumental response curves for the filters used, giving the total magnitudes of these stars in those bands. Comparing these total magnitudes to the instrumental magnitudes derived by EVALUATE allows to compute calibration factors. Averaged over all standard stars in the field, these can

then be used in a next step to scale the fluxes of all objects in the field to their absolute values.

Calibration of the COSMOS public optical data The saturation magnitudes for the optical bands of the public COSMOS dataset are as low as 21.8 in i^+ (see Table 3.1). This means that stars that are bright enough for obtaining spectra precise enough for calibration would all be saturated in these bands. As a consequence, the standard HIROCS procedure for calibration cannot be used with this dataset. On the other hand, the COSMOS images come pre-calibrated in units of nJy / pixel. Assuming that this calibration is correct, a self-calibration method resembling the standard HIROCS calibration can be applied.

For this, a set of "artificial" standard stars is created. Artificial in this case means that their spectrum is not measured, but constructed as a simple polygon from the flux measurements on the calibrated COSMOS data. These standard stars are selected from the point source catalog to be stellar objects without close neighbors and well-measured in all the optical bands. Their flux is then measured using a special EVALUATE setup, using a fixed PSF with the same FWHM as the seeing. The normalization of the weighted Gaussian aperture used by EVALUATE is chosen such that the flux value obtained with this setup is the total flux for point sources. With these total flux values, "artificial", low-resolution spectra standard stars are constructed manually. Five per quadrant, or 20 in total for the whole field, allow for a consistency check of the calibration factors based on them. In the first step, all quadrants are calibrated individually. The calibration factors obtained for each quadrant in this step are checked for consistency. In a second step, the four calibration factors are combined using an error-weighted mean, and the calibration is repeated using the averaged calibration factors, thus ensuring a uniform total calibration across the field.

Calibration of the IRAC data The calibration of the IRAC data is done in the same way as the calibration for the optical COSMOS data. The IRAC images also come pre-calibrated, so again a set of stars is chosen to self-calibrate the instrumental EVALUATE flux measured using the Gaussian aperture with the total flux of the calibration stars. As described in Section 4.3, care has to be taken when dealing with the brightest stars on the IRAC images as they are saturated. A set of three unsaturated calibration stars is chosen for the IRAC mosaics. Because of the complicated shape of the IRAC PSF, the way to measure the total flux is different from the measurements on the optical images. Here, a fixed aperture of $3''$ is chosen, and a correction factor is applied to scale the flux value obtained to the total flux value. The correction factors for the $3''$ aperture are interpolated linearly between the ones given in the COSMOS IRAC data description for aperture sizes of $2.9''$ and $4.1''$. They are 0.905 and 0.904 for the 3.6μ and the 4.5μ band, respectively. Like for the calibration of the optical data, artificial standard star spectra are created using these total flux values. The calibration factors obtained through the three standard stars are checked for consistency and averaged for the whole mosaic.

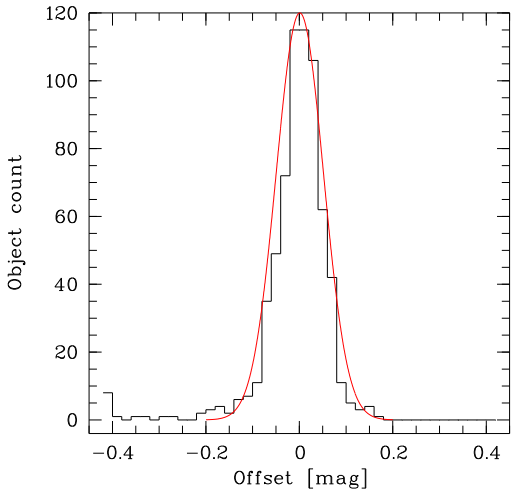

Figure 4.5: Magnitude offsets between the HIROCS H-band measurements for the COSMOS field and $H < 14.5^{\mathrm{mag}}$ point sources from 2MASS for the whole area of the mosaic where H-band data is available, regardless of its depth. The zeropoint offset is -0.0014^{mag}, with a scatter of $\sigma = 0.053$.

Calibration of the near-infrared data For the near-infrared data in H and Ks, the calibration procedure is again similar to the one for the optical data, the difference being that for these bands the artificial standard stars are created not using flux measurements on the images, but catalog magnitudes from 2MASS. The standard stars, 16 for H and 12 for Ks, are again selected to be stellar objects well-measured on the images, without close neighbors, and bright enough to only have small errors in the 2MASS catalog.

The flux calibration then follows the same recipe as for the optical bands. After the flux tables for H and Ks are calibrated, the 2MASS catalog can again be used for a consistency check. For the H-band flux tables, the magnitude offsets to all 2MASS objects with $H < 14.5^{\mathrm{mag}}$ in the field are calculated. The resulting Gaussian distribution of the offsets can be used to fine-tune the calibration. The same is done for the Ks-band with $K < 14.5^{\mathrm{mag}}$ 2MASS objects. After fine tuning, the magnitude zeropoint of the near-infrared bands is accurate to better than 0.01^{mag} and within the accuracy of the 2MASS catalog. An example for the magnitude offsets in H after fine tuning is shown in Figure 4.5.

4.6.2 Flux table colors

The object's colors which are the input for the multi-color classification are computed from the flux measurements stored in the flux table. As described in Section 2.2.1, the multi-color classification is done in an n-dimensional space, where n is the number of independent colors (the number of filters is $n + 1$).

The multi-color classification uses magnitudes (and thus the colors derived from them) in the asinh system (Lupton et al. 1999). The "traditional" color systems are statistically badly behaved when calculating colors of objects close to the detection limit. Computing

colors from faint magnitudes is identical to dividing two uncertain flux measurements. They are close to the singularity of the logarithm for zero flux, thus small differences in the flux measurements can lead to arbitrary color values. By avoiding this singularity, the asinh magnitude system is, in contrast to the other magnitude systems, well-behaved when computing colors of faint objects.

4.6.2.1 Zeropoint fine tuning

Even after the absolute calibration procedures described above, there is still room for small fine tuning in the object colors. This fine tuning is done using a stellar main sequence color shifting technique which has become a standard procedure for multi-color classification (*e.g.*, Wolf et al. 2001b). For this, a stellar main sequence template library is overplotted on the point source catalog measurements from the flux table in color-color space. Small calibration offsets are added to the objects' measured colors until the measurements exactly match the stellar template library's colors. Subsequently, these color offsets are applied to all object colors in the flux table before the multi-color classification is run.

Since extragalactic surveys are interested in exact classification of extragalactic objects, while the color calibration is done for stellar main sequence objects, this method only works if the stellar main sequence template library's colors are consistent with the galaxy template library's colors. A number of stellar template libraries exist in the literature. Some are based on theoretical models, while others are based on spectroscopic measurements. For this thesis, tests were done with the Pickles library (see Section 2.2.1.2) which is based on spectroscopic measurements extrapolated into the near-infrared out to the K-band, and a model-based stellar template library covering wavelengths from 10 nm to 1000 μ (Hereafter referred to as the "Decin library"; Decin, priv. comm., see also (Gustafsson et al. 2008)).

Comparison of the template spectra shows that even for the same stellar type, the SEDs from the two libraries can be quite different. Figure 4.6 shows the SEDs for main sequence stars of the spectral types F0V, G0V, K0V, and M0V from the Decin library overplotted on those from the Pickles library. While both libraries usually agree reasonably well in the optical part of the SEDs, the near-infrared parts tend to differ more, the extreme case being the SED of an M0V star (lower right panel of Figure 4.6). While the Decin library SED, based on stellar atmosphere model calculations, shows features out to the highest wavelengths, the Pickles library SED is a blackbody interpolation beyond the reach of optical spectroscopy at $\sim$ 1000 nm, with quite some disagreement in the overall shape of the SED. This of course also results in different colors for the same stellar types. Figure 4.7 shows some of the stellar main sequence color-color diagrams used to fine tune the calibration. Here, the colors are shifted such that the colors expected from the Pickles library stellar main sequence (plotted in green) matches the data (plotted in black). Expected colors from the Decin stellar library are overplotted in red. The differences in colors between the two template libraries is most prominent for the latest stellar types, getting more pronounced with the regarded colors becoming redder, as for example for M0V stars in the bottom right panel of Figure 4.7.

Since the Pickles library only goes out to 2500 nm, shifts for the IRAC colors can only

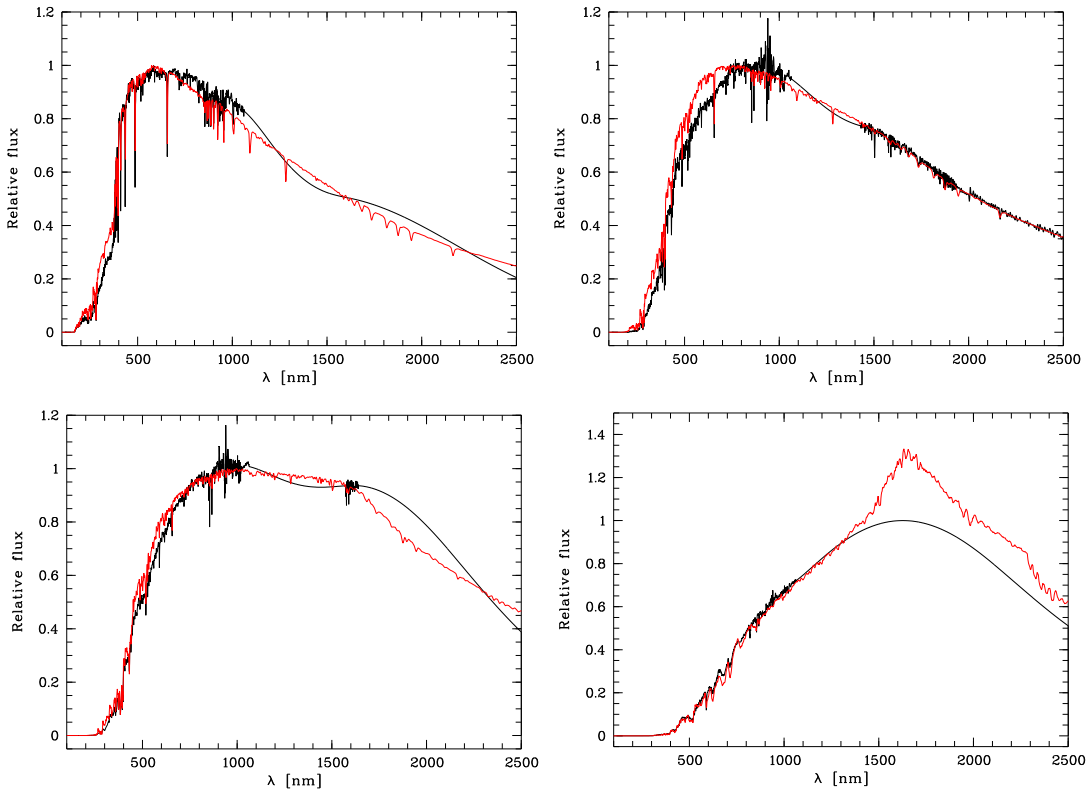

Figure 4.6: Comparison of the SEDs from the Decin library (red line) and the Pickles library (black line) for main sequence stars. Top left panel: F0V; top right panel: G0V; bottom left panel: K0V; bottom right panel: M0V

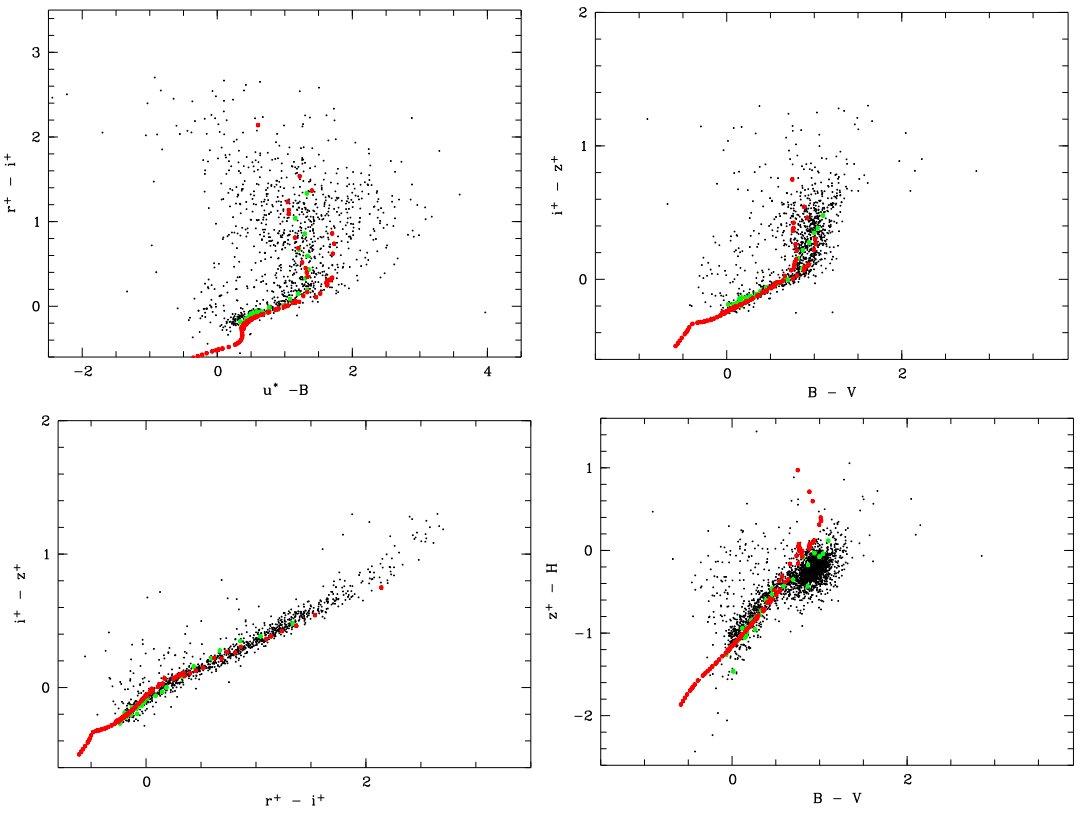

Figure 4.7: Color-color plots of stellar objects used to fine tune the color calibration. Overplotted on point sources in the COSMOS field (black dots) are the expected colors from the stellar template libraries (Pickles: green dots; Decin library: red dots). In the bottom right panel, only the main sequence stars from the Pickles library are shown. Color shifts are applied such that the Pickles library main sequence objects match the observed stellar main sequence objects.

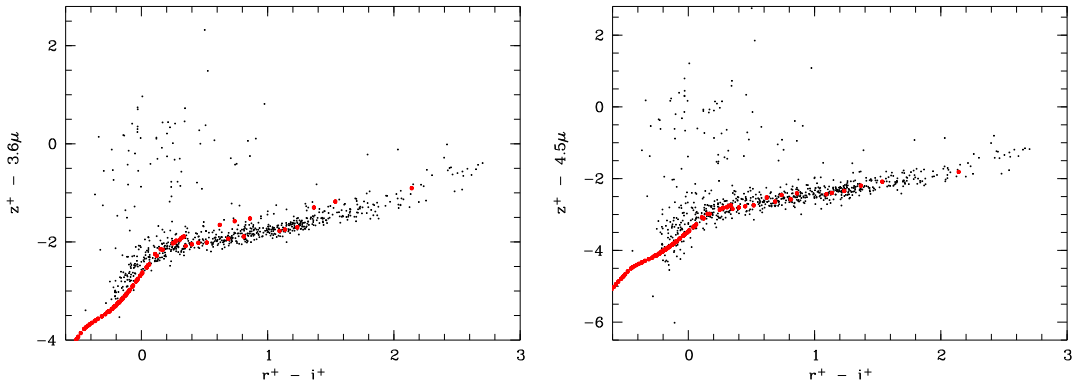

Figure 4.8: Color-color plots of the expected colors from the Decin library (red dots), overplotted on point sources in the COSMOS field (black dots) for colors involving IRAC 3.6μ (left panel) and 4.5μ (right panel) bands. Color shifts are applied such that the model library main sequence objects match the observed stellar main sequence.

be determined using the Decin library. Figure 4.8 shows the expected $(z^+ - 3.6\mu)$ vs. $(r^+ - i^+)$ and $(z^+ - 4.5\mu)$ vs. $(r^+ - i^+)$ colors of the Decin library overplotted on point sources in the COSMOS field, with shifts applied such that the template main sequence matches the observed stellar main sequence. Again, these color shifts are subject to verification with photometric redshifts from the classification of the spectroscopic comparison sample.

Since this color shifting technique provides only relative calibration, it cannot be used to tell which stellar template library is "better" in absolute terms. However, the differences in color calibration also result in different classification of the objects in the field, and especially of the objects from the spectroscopic comparison sample. Comparing the quality of the photometric redshifts with the spectroscopic redshifts (see Section 4.7.1), it can be tested which stellar template library's colors match the colors of the galaxy template library better. From the two stellar template libraries tested for this thesis, the colors obtained from the Pickles library result in more precise photometric redshifts for the spectroscopic comparison sample than those obtain from the Decin library, leading to the conclusion that the Pickles library is more consistent with the PEGASE galaxy template library than the Decin library. As a consequence, the Pickles library is used to fine-tune the calibration of the COSMOS dataset.

4.7 Object classification

With the colors determined accurately, the flux table is ready for the multi-color classification. In this step, the most plausible template objects from the template libraries of stars, White Dwarfs and galaxies are matched to each object in the flux table in an n-dimensional color space. Based on this, the likelihood of the catalog object to belong to either one of these classes is calculated, and a final classification is assigned (Wolf et al. 2001a).

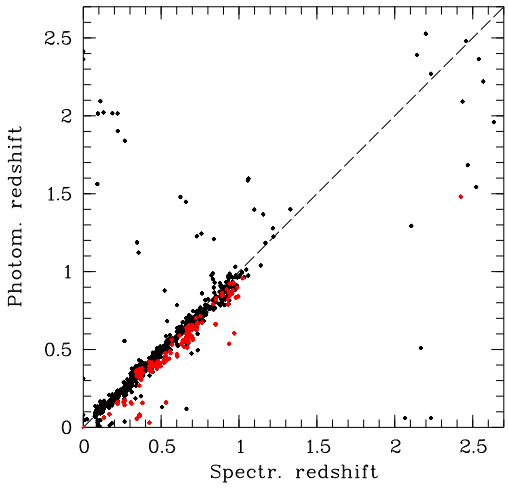

Figure 4.9: Plot of the photometric vs. spectroscopic redshifts for the comparison sample of 1072 objects. Here, the multi-color classification uses the bands u^*, B_j, g^+, V_j, r^+, i^+, NB816, z^+, H, and Ks. Galaxies classified as old and red are marked in red. A systematic offset of these galaxies in the spectroscopic redshift range $0.35 \leq z \lesssim 1$ is noticeable.

4.7.1 Comparison with the spectroscopic sample

The photometric redshift is one of the most important results of the multi-color classification for galaxies, thus a sample of galaxies with spectroscopically determined redshifts (see Sections 3.2.6 and 3.2.7) is used for testing the accuracy of the photometric redshifts and serves as a consistency check for the color shifts determined using the stellar main sequence shifting technique described above. Because of the big differences in SED shape between the various types of galaxies, it is safe to assume that if the photometric redshift of a galaxy is estimated accurately by the classification code, the galaxy type should also be determined reasonably accurate.

4.7.1.1 Classification without IRAC bands

The following comparisons use the HIROCS master table classified using the bands u^*, B_j, g^+, V_j, r^+, i^+, NB816, z^+, H, and Ks as reference. Color shifts as determined by the main sequence matching technique using the Pickles library are applied. Figure 4.9 shows a comparison of the resulting photometric redshifts with the spectroscopic redshifts for the comparison sample. The overall quality of the photometric redshifts is pleasing, except for the galaxies that are classified as old and red, which are marked in red. These galaxies show a systematic offset compared to the rest of the sample in the spectroscopic redshift range $0.35 \leq z \lesssim 1.1$. Above this redshift, no conclusion is possible because of the lack of objects. Dust absorption does not cause problems here, since all of these objects are classified as virtually dust free, as to be expected for red and dead galaxies with the last significant episode of star formation long ago in the past.

If the galaxy templates perfectly modeled the observed galaxies, the photometric redshifts of the comparison sample should automatically be excellent, except perhaps a few outliers, caused by measurement uncertainties in combination with degeneracies between high-z and low-z spectra which could only be broken with additional filters. Since the offset in redshift of old and dust free objects is not random and restricted to certain redshifts, a systematic offset between the measured object colors and the colors from the galaxy template library must be present. This assumption can be tested: If one fixes the redshift of these objects to the correct value, the color difference between the observed object and the template must become evident. In practice, $z_{\mathrm{phot}} = z_{\mathrm{spec}}$ is ensured by classifying the objects from the spectroscopic comparison sample using only templates of the correct redshift z_{spec}. The difference in colors between the best-matching templates at the correct redshifts and the observed galaxies is plotted in Figure 4.10 as a function of redshift for the various types of galaxies. For the younger objects, plotted as black dots, no pronounced systematic offset between measured and template color is visible in the optical colors except for $(u^* - B_j)$. This is very different for the old and red galaxies which are plotted in red. In practically any of the observed colors, a trend with redshift is visible compared to the templates. Like for the young objects, this is especially pronounced in $(u^* - B_j)$.

The picture is a little different when looking at the infrared colors shown in Figure 4.11. There are indications for a growing divergence between the young and old templates with increasing redshift, especially for $(z^+ - 3.6\mu)$. In both diagrams, the offset for the old galaxies shows an upwards trend, while the offset for the younger templates is getting increasingly negative. However, no final conclusion can be drawn because of the lack of objects at redshifts > 1.1. The few high-redshift, $z > 2$ objects seem to indicate that for $(z^+ - H)$ the trend towards negative offset values for the young galaxies does not continue. This seems plausible since a well-established wavelength range of the spectrum is redshifted into the H-band regime at these redshifts. Another thing to note from these two diagrams is that the scatter in the offsets is strongly increased compared to the optical colors. For the H-band, the scatter is to a large extent due to the shallowness of the data, especially in the regions where the spectroscopic comparison sample is taken from. The IRAC images go a lot deeper than the H-band, and a large part of the scatter in $(z^+ - 3.6\mu)$ is due to uncertainties in the templates in this wavelength regime. Figure 4.12 showing the SED of a passive galaxy at $z \approx 1$ illustrates this problem. While the shape of the SED is very well fitted in the red optical bands, the best-fitting template is too red in the infrared, causing an offset in all bands redwards of z. This is probably due to a difference in the star formation history between the templates and the real objects.

There is also a mismatch in the $(u^* - B_j)$ colors of the templates and the observed objects. The reason becomes apparent when looking at the example in Figure 4.13, showing the template SED in green with the measurements and their errors overplotted in black. While this is an extreme case with a difference of 1.2^{mag} between the observed $(u^* - B_j)$ color and the expected color from the template, it illustrates the general difference in shape present for almost all of the old, red objects. While the template SED is horizontal in the UV, the u^* measurements of the object are in agreement with a SED declining in the UV.

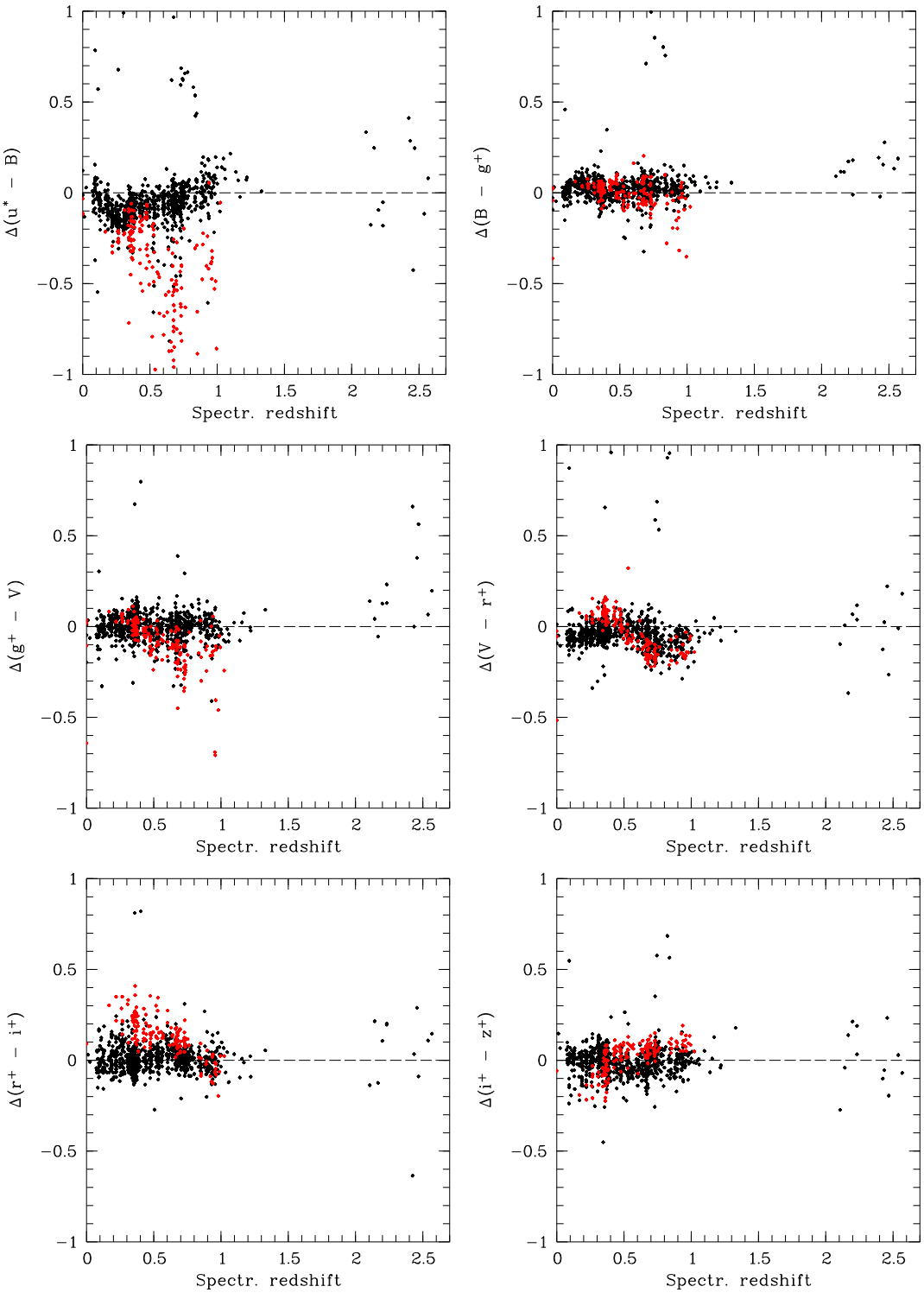

Figure 4.10: Offsets $c_{\mathrm{measured}} - c_{\mathrm{template}}$ between the observed optical colors and the template colors as a function of redshift for the objects of the spectroscopic comparison sample. Objects classified as old and red are plotted in red. Top left panel: $(u^* - B_j)$; top right panel: $(B_j - g)$; middle left panel: $(g - V_j)$; middle right panel: $(V_j - r^+)$; bottom left panel: $(r^+ - i^+)$; bottom right panel: $(i^+ - z^+)$.

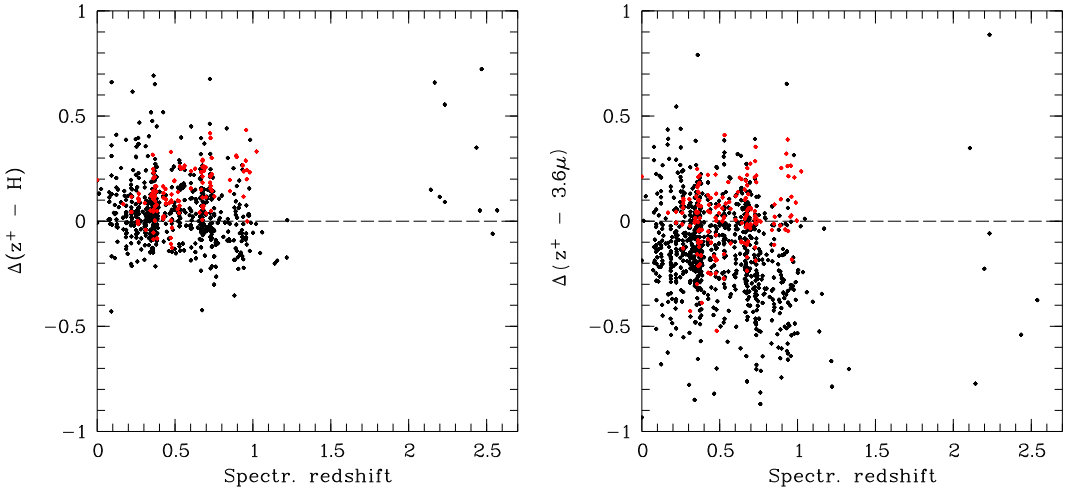

Figure 4.11: Offsets $c_{\mathrm{measured}} - c_{\mathrm{template}}$ between the observed near-infrared colors and the template colors as a function of redshift. Colors are the same as in Figure 4.10. Left panel: $(z^+ - H)$; right panel: $(z^+ - 3.6\mu)$.

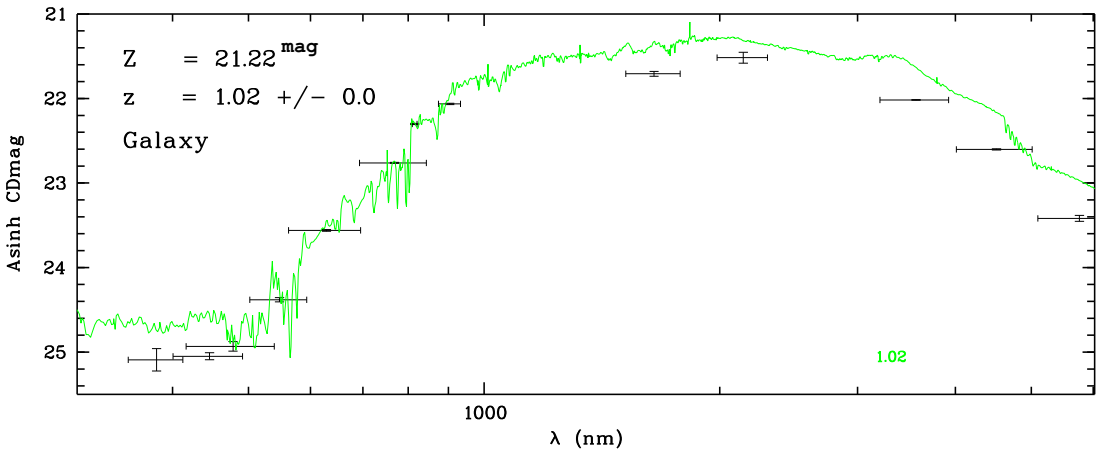

Figure 4.12: Example of a passive galaxy with mismatching overall shape of the best-fitting SED. The template object's SED is too red in the near-infrared and too blue in the near-UV.

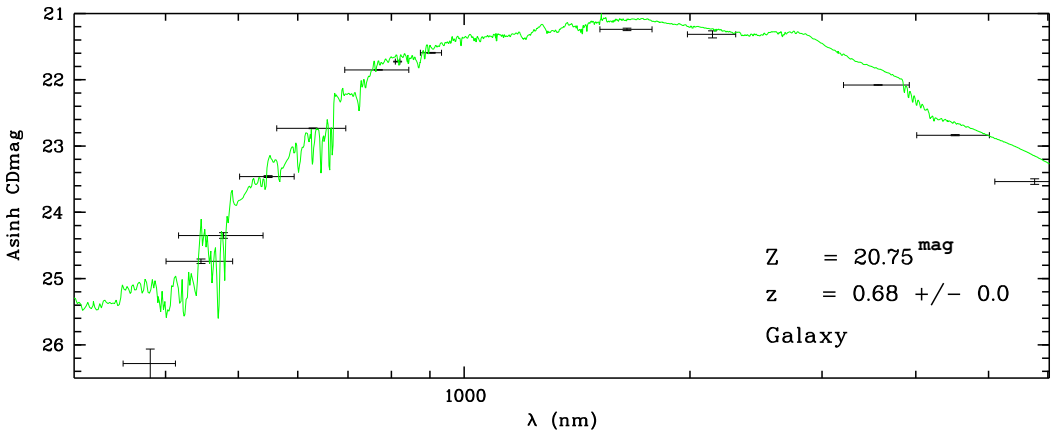

Figure 4.13: Example of an object with mismatching $(u^* - B_j)$ observed / template colors. The template SED is plotted in green, with the measurements and their errors overplotted in black. The measurement in u^* is much fainter than the value expected from the template model.

To check if the redshift offset of passive, red galaxies is unique to the COSMOS data, the COMBO-17 Chandra Deep Field South spectroscopic comparison sample is classified using only the five broad band COMBO-17 filters U, B, V, R, and I (Wolf et al. 2004). The plot of the resulting photometric redshifts vs. spectroscopic redshifts, shown in Figure 4.14, exhibits some of the characteristics of Figure 4.9. Because I is the reddest filter used here, the quality of the classification worsens already below $z = 1$, in contrast to the COSMOS data. The offset in photometric redshift for the old, dust free galaxies seems to be present especially around $z \approx 0.6$. This offset is gone if the classification is done using the full set of five broad and twelve medium band filters, leading to the conclusion that with the set of templates used for this thesis, at least some medium band filters would be necessary to get correct photometric redshifts for the old, dust free galaxies. In the COSMOS field, they are expected to be releases within COSMOS DR2 in late 2008, and can be integrated into the HIROCS data then.

4.7.1.2 Classification including IRAC bands

Next, the HIROCS master table including IRAC 3.6μ and 4.5μ bands is tested. The color shifts for the non-IRAC bands are the same as in the previous analysis. Shifts for colors involving the IRAC bands are determined by the stellar main sequence matching technique using the Decin template library. The result is shown in Figure 4.15. The quality of the photometric redshifts is clearly inferior to the classification without the IRAC bands. Evidently, the photometric redshifts are a lot more uncertain, and the number of total misclassifications is strongly increased. However, one has to note that when using the stellar main sequence matching technique as shown in Figure 4.8, color shifts for the IRAC bands are mainly determined from the position of the M dwarfs, which lie rightwards of the knee in the distribution of data points. The behavior of their SED in the infrared depends on many parameters like surface gravity which are not well known, it is thus very uncertain and still under debate. As noted already above, the

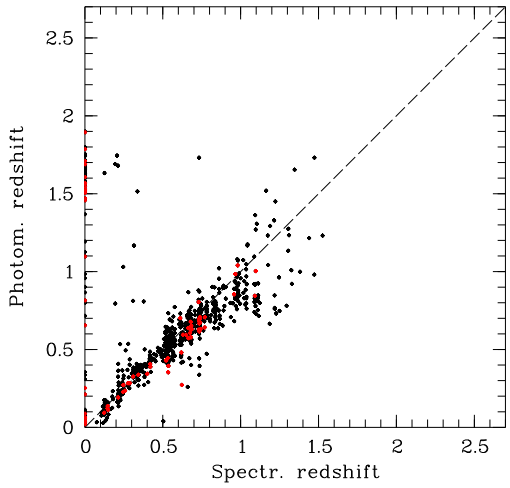

Figure 4.14: Photometric vs. spectroscopic redshift for COMBO-17 objects from the Chandra Deep Field South (Wolf et al. 2004). Here, only the broad band filters U, B, V, R, and I are used for the multi-color classification.

Decin stellar library seems to not match well with the PEGASE galaxy templates in the IRAC wavelength regime.

The question can be raised if color shifts other than those determined via the Decin stellar library may lead to better photometric redshifts. Indeed, the relative calibration of the IRAC bands can be changed to yield much better photometric redshifts, as shown in Figure 4.16.

Compared to the classification without the IRAC bands, the overall shape has not changed much. The number of outliers seems to be slightly reduced compared to Figure 4.9, but closer inspection shows that many of the former redshift outliers are not being assigned the correct redshift when including the IRAC bands, instead they are unclassified or classified as stellar objects and thus do not appear in the redshift plot.

One would expect that the IRAC bands help with the classification especially of high-redshift objects. The comparison with the spectroscopic sample does not show significant improvement in the photometric redshifts, though one has to note that no objects with spectroscopic redshifts $1.3 \leq z \leq 2$ are included in the sample, which would benefit to a larger degree from the IRAC bands. Unfortunately, the color calibration for the IRAC bands is very uncertain. The color shifts necessary to obtain the optimal photometric redshifts are 0.4^{mag} different from those determined by the stellar main sequence fitting technique, and cannot be motivated with anything other than the improvement in the photometric redshifts. This is probably due to uncertainties in the infrared part of the SEDs of late type M0V stars, making calibration uncertain. In addition, experiences with IRAC data for the COMBO-17 survey (Meisenheimer, priv. comm.) show that colors involving IRAC bands give rise to problems with the multi-color classification, probably caused by galaxy templates not matching the data in the IRAC wavelength regime. As a result, the IRAC bands are not used for the multi-color classification at this point, until better galaxy templates become available.

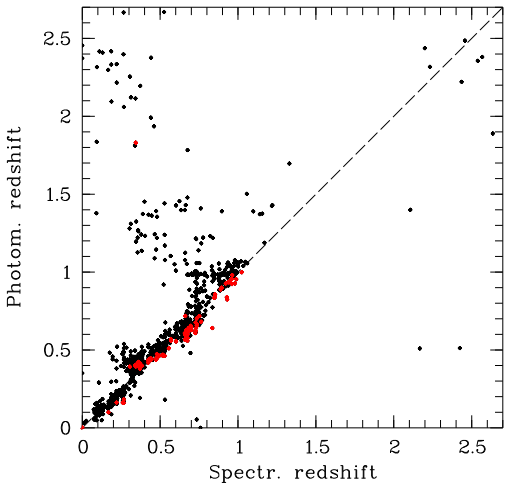

Figure 4.15: Photometric redshifts vs. spectroscopic redshifts for the comparison sample of 1072 objects. Here, in addition to the bands u^*, B_j, g^+, V_j, r^+, i^+, NB816, z^+, H, and Ks, also the IRAC 3.6μ and 4.5μ bands are used, with the shifts determined by the stellar main sequence matching method. As in Figure 4.9, Galaxies classified as old and dust free are marked in red.

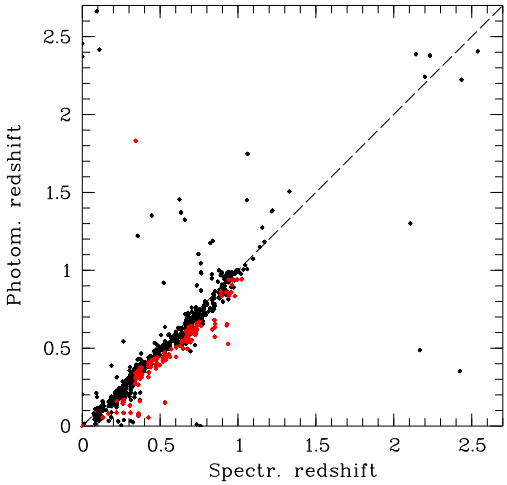

Figure 4.16: Plot of photometric redshifts vs. spectroscopic redshifts for a classification using the bands u^*, B_j, g^+, V_j, r^+, i^+, NB816, z^+, H, Ks, IRAC 3.6μ, and 4.5μ. Note that the colors involving IRAC bands are shifted 0.4^{mag} different from those in Figure 4.15. As in Figure 4.9, Galaxies classified as old and dust free are marked in red. Again, a systematic offset of these galaxies in the redshift range $0.35 \leq z \leq 1$ is noticeable.

Table 4.1: Color shifts used for the COSMOS bands. The color is given in the left column, the zeropoint shift used in the final flux table classification is given in column "Zeropoint shift".

Color	Zeropoint shift
$u^* - B_j$	-0.01
$B_j - g^+$	-0.05
$g^+ - V_j$	-0.18
$V_j - r^+$	-0.05
$r^+ - i^+$	-0.09
$i^+ -$ NB816	-0.03
$i^+ - z^+$	-0.12
$z^+ - H$	-0.20
$z^+ - Ks$	-0.36

4.7.2 Final classification

The HIROCS catalog for the COSMOS field is classified using u^*, B_j, g^+, V_j, r^+, i^+, NB816, z^+, H, and Ks. The final classification of the flux table is done using the color shifts determined with the Pickles library. Table 4.1 lists the color shifts used for the COSMOS bands. Ideally, all colors would be computed using adjacent filters. For this data set, two exceptions to this are necessary: The NB816 filter is used as a supplement, thus it is only connected to the i^+ band and not used to bridge the i^+ and z^+ bands. In the near-infrared, the H-band does not cover the field completely, so $(H - Ks)$ cannot be computed for all objects. Because of this, also the Ks-band is connected to the z^+ using $(z^+ - Ks)$. The offset between the optical and the near-infrared bands is quite large and probably due to the different calibration methods used (self-calibration with the data itself vs. calibration using 2MASS).

The resulting photometric redshifts are shown in Figure 4.9. While being already quite good, the overall quality of the photometric redshifts can be further improved by correcting the photometric redshift for the old, red objects in the redshift range $0.35 \leq z \leq 1.2$. These show a well defined systematic offset compared to the spectroscopic redshifts which can be modeled and removed manually. For this, a quadratic function is determined as a fit to the offset of the photometric redshifts. Using this, they can be corrected to be

$$z_{\mathrm{phot,\,corrected}} = z_{\mathrm{phot}} - \left(0.26 \cdot z_{\mathrm{phot}}^2 - 0.38 \cdot z_{\mathrm{phot}} + 0.09\right),\ 0.35 \leq z_{\mathrm{phot}} \leq 1.2.$$

The resulting plot of photometric vs. spectroscopic redshifts is shown in the left panel of Figure 4.17, together with a distribution of $\Delta(z)$ in the right panel. The overall accuracy of the photometric redshifts is very good, with a negligible zeropoint error and a scatter of $\sigma = 0.023$ over the entire redshift range. The number of catastrophic outliers with $\Delta(z) > 3\sigma$ is 8%. One should note that the numbers given here are most likely lower limits for the accuracy of the photometric redshifts in those parts of the field where deeper H-band data is available. Most members of the spectroscopic

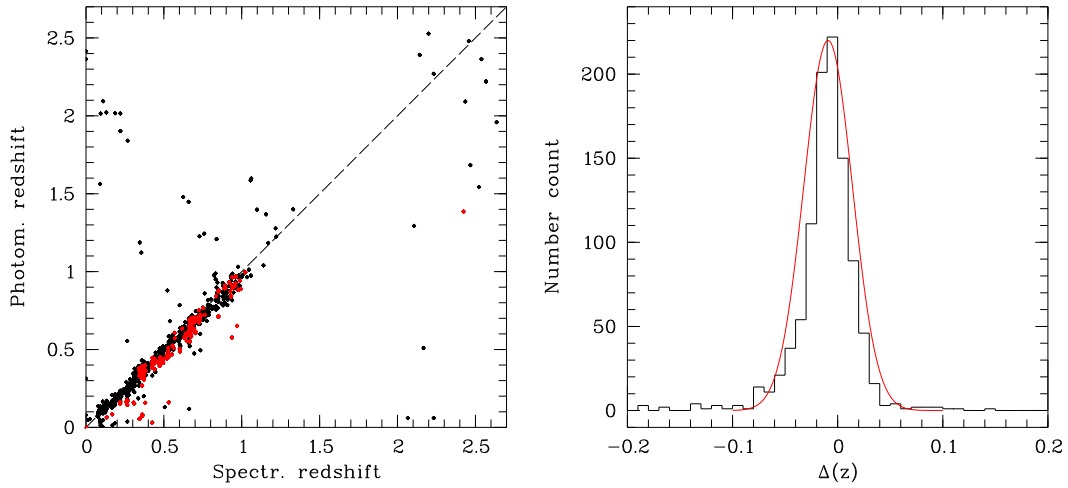

Figure 4.17: Left panel: The same plot as in Figure 4.9, but this time the photometric redshifts of the galaxies classified as old and dust free have been manually corrected between $z = 0.35$ and $z = 1.2$. Right panel: Distribution of $\Delta(z) = \frac{z_{\mathrm{spec}} - z_{\mathrm{phot}}}{1+z_{\mathrm{spec}}}$ for the objects in the left panel. The offset is 0.01, the scatter is $\sigma = 0.023$. The number of catastrophic outliers with $\Delta(z) > 3\sigma$ is 8%.

comparison sample lie in areas not covered by the H-band yet, or where the H-band data is shallow. The search for the highest-redshift clusters is restricted to areas where the H-band goes sufficiently deep. The redshift regime of $z < 1$, where clusters are searched on the complete field, is not affected by biases in the redshift distribution. Catastrophic outliers are present here, but only in the sense that objects are either scattered to very high ($z > 1.8$) or very low ($z < 0.2$) redshifts. While this effect can affect real structures by decreasing the number of member galaxies, lowering their contrast with respect to the field and making them harder to detect, it cannot produce spurious structures at $0.45 < z < 1$, the target redshift range of the cluster search on the whole field.

The redshift distribution for the COSMOS field is shown in Figure 4.18. The left panel shows the histogram for all objects in the catalog, regardless of redshift error or brightness, whereas the right panel shows objects selected to have a SNR of at least 5 in z^+. When excluding the faint objects with unreliable photometry, the peak around $z = 2$ decreases in height, showing that the majority of these objects are misclassifications. At redshifts below one the shape of the histogram looks as expected due to the results from *e.g.* the VIMOS VLT Deep Survey (VVDS, Le Fèvre et al. 2005). This selection will be used for the $z \leq 1$ cluster search on the whole COSMOS field. Figure 4.19 again shows the redshift distribution, this time for the subcatalog used to search for the distant galaxy clusters, see Section 6.2. Using only areas which have a depth of at least $H = 21.4$ (3σ) improves the shape of the redshift distribution substantially above $z = 1$ by significantly decreasing the number of catastrophic outliers, again illustrating the need of near-infrared data for reliable photometric redshifts in this redshift regime. It should be noted that this catalog is constrained to objects with $z < 2$.

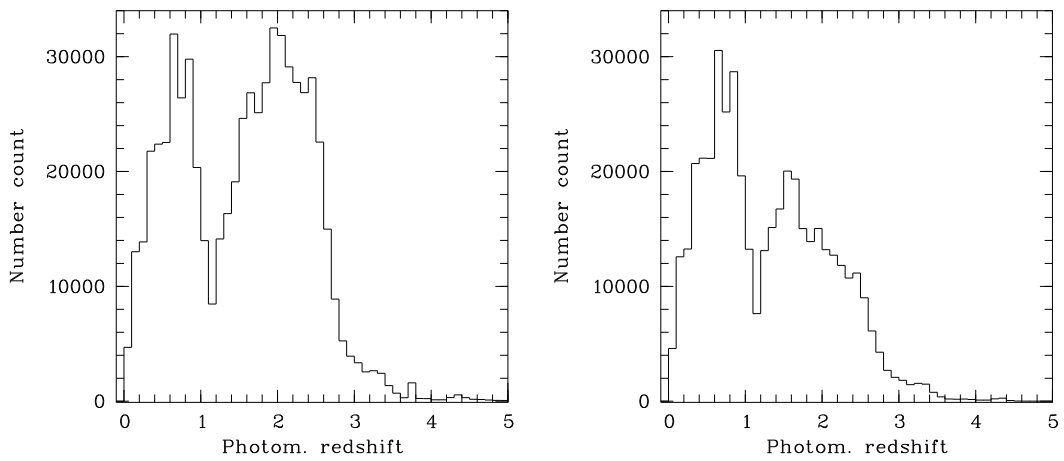

Figure 4.18: Redshift distribution for objects in the COSMOS field. Left panel: Histogram for all objects in the catalog. Right panel: Objects selected to have a SNR of at least 5 in z^+.

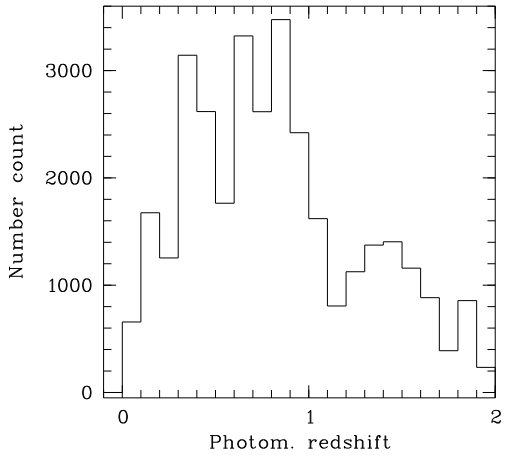

Figure 4.19: Redshift distribution for the subcatalog on areas with a depth of at least $H = 21.4$ (3σ) and photometric redshifts $z < 2$.

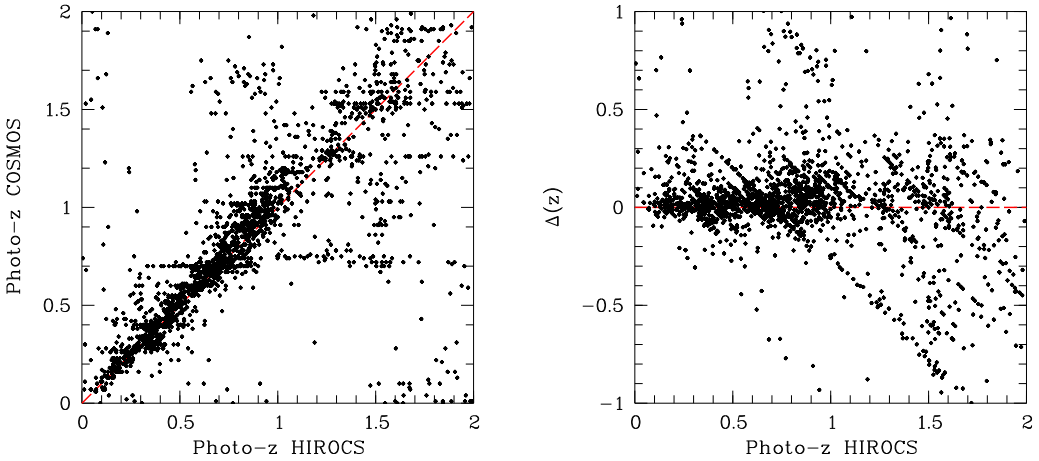

Figure 4.20: Comparison between the photometric redshifts of HIROCS and those from the official COSMOS catalog below $z = 2$. The right panel shows the redshift difference $\Delta(z) = z_{\mathrm{COSMOS}} - z_{\mathrm{HIROCS}}$ in the y-axis. Objects from the two catalogs are matched with a positional tolerance of $0.2''$. For clarity, only every 100th object is plotted. It should be noted that the objects plotted here have magnitudes as low as $z^+ = 25$.

In all redshift distributions, a peak at $z \approx 0.7$ is readily visible. It is caused by the prominent LSS studied by Guzzo et al. (2007), the most massive structure in the COSMOS field, with thousands of member galaxies.

Figure 4.20 shows a comparison of the HIROCS photometric redshifts with the photometric redshifts from the official COSMOS catalog below $z = 2$. Objects between the two catalogs are matched with a positional tolerance of $0.2''$. The agreement between the two samples below $z = 0.7$ is quite good for photometric redshifts, with minimal systematic offsets and only few objects with ambiguous redshifts $z \approx 0$ / $z > 1.5$. In the redshift range $0.7 \leq z \leq 1$, there is a number of objects which get significantly higher photo-zs in the respectively other catalog. Furthermore, the COSMOS photo-zs are systematically higher in this redshift range (see also Section 7.4 for the special case of a $z \approx 0.84$ galaxy cluster). Some degeneracies of the COSMOS photometric redshifts with respect to the HIROCS photo-zs are noticeable, the most prominent one at $z_{\mathrm{COSMOS}} \approx 0.7$. Above $z = 1$, the scatter and the number of objects with totally different photo-zs in the two catalogs increase significantly due to the absence of deep near-infrared photometry in the COSMOS catalog.

There exist a number of redshift degeneracies between the HIROCS and the COSMOS photo-zs showing up as horizontal and diagonal stripes in the left and right panels of Figure 4.20, respectively. They result from objects having different photometric redshifts in HIROCS being compressed to a single redshifts by the COSMOS code. Which of the two solutions (if any) is correct cannot be decided at present; additional spectroscopic redshifts are required for this.

4.7.3 Calculation of rest frame magnitudes and colors

Rest frame magnitudes are computed by convolving a redshifted filter curve with the best-fitting template SED, normalizing with the total magnitude in the nearest observed band. As the program EVALUATE, which is used to derive the photometry, yields instrumental magnitudes, the SEXTRACTOR mag_best measurements obtained during catalog creation are used as total flux measurements for SED normalization.

For the catalog of the whole COSMOS field, total magnitudes are obtained in the i^+-band image, which has the narrowest PSF of all optical bands while being one of the deepest images at the same time. In addition, it traces object flux in a wavelength regime which is crucial for the cluster galaxies studied here. Due to the i^+-band images being saturated at 20^{mag} for extended sources and 21.8^{mag} for point sources, some of the brighter objects in the catalog do not have rest frame magnitudes assigned. However, none of the cluster members are affected by this.

In the areas with deep H-band coverage used for the cluster search above $z = 1$, z^+-band measurements are used to provide total magnitudes for some objects that do not have i^+-band SEXTRACTOR magnitudes, most of them being objects with a reliable SEXTRACTOR detection only in the H-band.

5 Cluster selection algorithm and -function

5.1 The HIROCS cluster selection algorithm

One of the essential parts of a photometric redshift cluster survey is the cluster selection function. The cluster survey should reliably isolate and pick the galaxy clusters with as little overall contamination by non-member galaxies as possible. As discussed in the introduction, high-redshift cluster surveys need some way to select galaxies according to their redshifts, thus enhancing the contrast with respect to fore- and background objects. There is a number of ways to do this, for example selecting objects with colors in agreement with red sequence galaxies at the respective redshifts, or the use of photometric redshifts, which have the advantage of not adding priors (and thus possible selection biases) about the properties of distant galaxy clusters.

Unlike for spectroscopic redshift surveys, it is not possible to select structures with a simple friend-of-friends (FoF) algorithm in photometric redshift surveys. Because of the comparably large uncertainty of the photometric redshifts, all objects in the field would be selected as belonging to one structure (Botzler et al. 2004). One solution is to do FoF searches in redshift slices, thus avoiding runaway effects. In a next step, all structures found in the different slices have to be tested if they are in fact belonging to the same overall structure, and combined if this is the case. An implementation of this algorithm, the so-called "extended friends-of-friends" algorithm, has been successfully tested with data from the MUNICS survey (Botzler et al. 2004).

As described in Section 2.2.2, HIROCS takes a different approach by computing the local density for each object. Galaxy clusters form the density peaks among the distribution of galaxies, so the goal is to select all overdense objects belonging to the same structure. Plotting the positions of objects with local density above the cutoff threshold shows that the majority of overdense objects forms structures which are well-separated when taking into account both projection on the sky as well as redshift. However, because of possible overlap in projection between structures at different redshifts, selection purely based on the position on the sky is not optimal as it could select objects from structures at different redshifts. Using the additional redshift information, no runaway effects are to be expected when using a FoF algorithm to select groups of overdense objects. Dividing the data into redshift slices introduces unnecessary complication, structures could be artificially split up, falling into more than one slice and would have to be combined later.

The cluster selection program implemented for this thesis uses the following algorithm based on FoF to select structures of overdense objects:

First, the object table is sorted with respect to decreasing local density. In the following steps, only objects above the pre-defined overdensity cut are considered:

- The object with the highest overdensity value is picked as starting point.
- The starting object's photometric redshift is not necessarily the mean redshift of the structure. The mean structure redshift is estimated as follows:
  - In a 300 kpc search radius around the most overdense object, the three most overdense objects (including the starting object) within a redshift range of ± 0.1 around the redshift of the most overdense object are selected.
  - Their photometric redshifts are averaged, resulting in the estimated mean redshift of the structure.
- The search cuts in redshifts space are defined to be $\pm 2\sigma$ around the mean redshift. The σ value is taken from the comparison between the spectroscopic and photometric redshifts of the comparison sample.
- All connected overdense objects of the structure are searched with a FOF algorithm. Again, the search radius between galaxies is chosen to be 300 kpc. The friends are marked in the object table with the structure ID.
- Once all members of the structure are found, the algorithm continues from beginning, this time searching around the most overdense object that is not yet marked as belonging to a structure.
- As a last step before the output of the results, a cutoff using a pre-defined minimum number of objects per structure can be applied. This way the inclusion of structures with low significance in the output catalog is avoided.

Because no assumptions about the shape of the structures and the distribution of the member galaxies enter into this algorithm, it is capable of reliably selecting also structures which are shaped asymmetrically or do not exhibit a strong central concentration. However, due to the FOF selection of member galaxies, it is possible that two structures which are close in redshift space as well as in the projection on the sky are not separated, as is shown in the example of the $z = 0.73$ large scale structure in Figure 5.1. Furthermore, it is possible that in this case not all member galaxies are selected. If the structures' redshifts are not as close together as in the example, the combination of redshift separation, intrinsic scatter and the $\pm 2\sigma$ redshift boundaries might cause the high and low redshift tails to be truncated, thus reducing the number of selected member galaxies.

The cluster selection algorithm can be configured to only select structures within a specified redshift range by limiting the selection of starting objects to overdense galaxies within these boundaries. It may happen that massive structures containing many members slightly outside this redshift range are selected, because some of their overdense members are included in the search range due to the scatter of the photometric redshifts, serving as starting points for the algorithm. In this case, nearly all of their member objects are selected since the estimate of the average redshift of the structure is the mean of the overdense object selected initially (with its redshift being within the boundaries), and the two most overdense galaxies within 300 kpc in projection on the

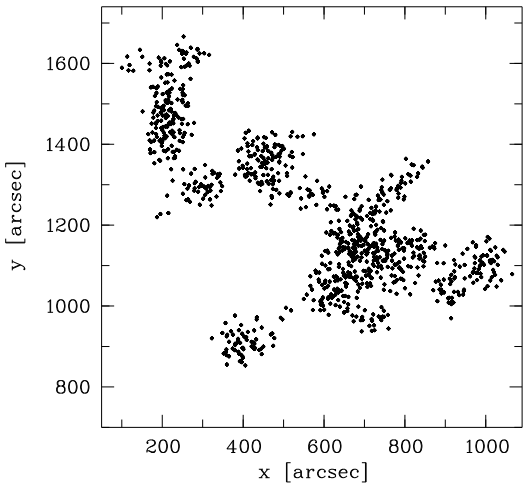

Figure 5.1: Large scale structure at $z = 0.73$ in the COSMOS field which is not separated by the cluster selection algorithm. All subcomponents are close together in redshift space, and also in the projection on the sky. There are "bridge" objects whose membership to a single structure is ambiguous, causing the FOF algorithm to connect the structures.

sky and close in redshift space (± 0.1, this time not taking the boundaries into account). The estimated structure redshift might be slightly off because of the bias of the initially selected object, but object selection will still be nearly complete because the redshift boundaries do not affect the subsequent selection of members.

The HIROCS cluster selection based on the local density calculation and overdense object selection using the algorithm described above has been characterized with mock sky data based on simulations. The results are presented in the following section.

5.2 The HIROCS cluster selection function

5.2.1 General remarks

All cluster detection algorithms need to be tested against data with well-defined, pre-known properties to assess their capabilities and quantify their performance like the recovery rate and its dependence on cluster properties, the false detection rate etc.

The HIROCS 3D overdensity cluster search algorithm has been analyzed by Falter (2006). Artificial clusters inserted into real multi-color data from the COMBO-17 survey were used to determine the recovery rate with respect to the number of members. Its susceptibility for false-positive detections has been tested using randomizations of the object catalog with respect to the projected positions on the sky with the magnitudes and redshifts unaltered, as well as with randomizations of the redshifts with the positions left unchanged, analogous to the works of Kim et al. (2002) and Goto et al. (2002). In both tests, no structures above 3σ overdensity were found, pointing towards the robustness of the detection algorithm against spurious detections.

These tests proved the reliability of the cluster detection algorithm and the feasibility of using it to conduct a multi-color search for distant galaxy clusters. However, due to the nature of the data used for these tests, the cluster selection could be tested more thoroughly if the following points were taken care of:

- These tests could probe recovery rates based on the number of members of the artificial clusters only. No assessment of the recovery rate with respect to the cluster masses has been possible.
- The artificial clusters were assumed to be isolated in space, with a radial King profile (King 1968) for the distribution of all the member galaxies in one test case, and a King profile for the distribution of the passive members together with the blue cloud members being distributed evenly across the cluster in a second test case, respectively. While the second scenario is more realistic, it still assumes the clusters to be isolated and does neither account for the correct cosmological context like the LSS the clusters are embedded in, nor a possible shape evolution of the galaxy clusters with redshift.
- The population mix of the cluster galaxies was somewhat static. A parameter study has been performed to assess the dependence of the recovery rate on the population mix under the assumption that the blue cloud members are not centrally concentrated, while the passive galaxies are. This did not take into account the evolution of the population mix with its distribution across the cluster with redshift. The number of objects of the different populations and thus different colors per cluster is affected by the magnitude cut of the catalog, thus evolutionary effects play a role in the performance of the cluster search.
- False detection rate tests based on randomization could not take into account projection effects, the major cause of false detection in cluster searches based on photometric redshifts.

All these shortcomings can be addressed by using catalogs derived from mock skies based on simulations of the evolution of dark matter halos, the LSS and the embedded galaxies as a function of redshift as the testing dataset. Mock catalogs from simulations provide a way of testing the algorithm against more parameters and with more realistic data: The clusters are embedded in LSS; they have realistic shapes which evolve with redshift; their galaxy population shows evolution effects, and false detection rates can be analyzed taking into account projection effects. Moreover, the simulations provide dark matter halo masses for the clusters, thus allowing for an assessment of the recovery rate based on cluster mass.

However, the quantitative results based on the simulations are only precise if the underlying model provides a realistic picture of the universe and takes into account all the relevant effects. While the dark matter simulations model the properties of the LSS over a huge range in redshift remarkably well, even very basic mechanisms of galaxy evolution are not fully understood to date, complicating the creation of galaxy evolution models. Known problems of contemporary models include incorrect galaxy colors, resulting for example in a wrong sign of the red sequence slope and the absence of a green valley, near-infrared magnitude issues due to dust absorption models (see Section 5.2.3). Thus, the results based on the mock catalogs have to be taken with a grain of salt.

5.2.2 Properties of the mock sky

For this thesis, a mock sky created by Obreschkow (in preparation) is used as a basis for creating a mock catalog with properties resembling those of the HIROCS COSMOS data. It makes use of the publicly available catalog simulating galaxy evolution described in (De Lucia and Blaizot 2007), which uses a modified version of the Croton et al. (2006) model. Extending the dark matter halo structure of the Millennium Simulation (Springel et al. 2005), semi-analytical models were used to simulate the galaxies populating the dark matter halos and trace their evolution across cosmic timescales. Cooling flows as well as AGN feedback to suppress the cooling flows are included in these simulations. The cosmological parameters used are $\Omega_\mathrm{m} = 0.25$, $\Omega_\mathrm{b} = 0.045$, $h = 0.73$, $\Omega_\Lambda = 0.75$, $n = 1$ and $\sigma_8 = 0.9$. Parent dark matter halos are identified using a friends-of-friends algorithm and decomposed into a set of disjoint substructures; the main halo of a structure typically comprises 90% of its mass. To compute the photometric properties of the model galaxies, the stellar population synthesis model from Bruzual and Charlot (2003) is used. More details about the galaxy evolution model can be found in (Croton et al. 2006).

The mock sky is created by calculating the light cone of a present-day observer, reaching backwards in the simulation data. The underlying Millennium simulation spans a comoving volume of $500\,\mathrm{Mpc}/h$, corresponding to the comoving distance out to $z \sim 0.17$. Since the mock sky should extend out to $z = 2$, replications of the simulation have to be used, exploiting the periodicity of the simulations. In order to cover the volume needed for this study, eight repetitions of the simulation volume are required, meaning that some of the galaxies are included in the catalog up to eight times. However, because each of these repetitions is at a different redshift, these galaxies are in a different state of their evolution in each of the repetitions. In projected area, the mock sky covers about 51 square degrees.

The details about the construction of the mock sky will be found in (Obreschkow in preparation). The method applied is similar to the one described in (Kitzbichler and White 2007), which can be used as a general reference and description of the process of creating a mock sky from simulated data.

For each galaxy, the mock sky catalog provides:

- The projected coordinates on the sky
- Its cosmological redshift
- Its apparent redshift, taking into account Doppler effects
- Its cosmological distance in Mpc
- The virial mass of the dark matter subhalo the galaxy is the center of
- The virial mass of the most massive dark matter halo the galaxy's subhalo is associated with
- Absolute rest-frame magnitudes in B, V, R, I, and K, without dust absorption
- Absolute observer-frame magnitudes in the SDSS-bands u, g, r, i, and z, as well as in the 2MASS-bands J, H, and K, all of them with dust absorption taken into account.

As an estimate for the total virial mass of a halo, the virial masses of all its members above the brightness cut applied (see below) are summed. This is biased low, since some subhalos might not be accounted for. The bias should not be very pronounced though, since the missing galaxies are the fainter ones, predominantly living in smaller halos and thus not adding much to the total mass budget.

5.2.3 Modeling the survey data

The mock catalog provides exact values for the redshifts of objects. For a realistic analysis of the performance of the cluster search algorithm, the properties of the real-world data, *i.e.* the limited accuracy of photometric redshifts, have to be modeled and imposed on the mock catalog.

In a first step, the absolute magnitudes have to be converted to apparent ones. With the data provided in the mock catalog, this task is an easy one. Magnitudes in the SDSS as well as the 2MASS bands are already provided in the observer frame, so no redshift dependent K-corrections are required. The effects of dust absorption are also already contained in the magnitudes, so the only task that is left is to calculate the apparent magnitude based on the luminosity distance. The magnitude correction is computed as

$$\Delta M = 5 \cdot (\log_{10} D_{\mathrm{lum}} - 1)\,, \tag{5.1}$$

with the luminosity distance D_{lum} given in parsecs.

Unfortunately, there is an issue with the observed H-band magnitudes in the mock catalog, probably due to the dust absorption model. Resulting from this, the number counts in the near-infrared do not match real data (see Kitzbichler and White 2007). The simulated z-band magnitudes do not suffer from the effects of the dust absorption model, and are thus the underlying quantity for selecting the mock catalog (see Section 5.2.4.1).

5.2.3.1 Photometric redshifts

Next, two datasets are being created:

1. A comparison sample with precise redshifts for the reference cluster search. Here, the cosmological redshifts are used. This sample will henceforth be called the "spectroscopic mock sky".
2. A dataset matching the survey data with modeled photometric redshifts. This sample will henceforth be called the "photometric mock sky".

For the second dataset, the scatter of the photometric redshifts and their errors have to be modeled. For this, the apparent redshifts are randomized with a Gaussian distribution. The standard deviation of the scatter in $\Delta z = \frac{z_{\mathrm{phot}} - z}{1+z}$ is chosen to match the value derived from comparison of the photometric and the spectroscopic redshifts of the comparison sample, see Section 4.7.2. The scatter measured there is $\sigma = 0.023$; here the value $\sigma = 0.025$ is adopted. No significant dependence of the accuracy of the

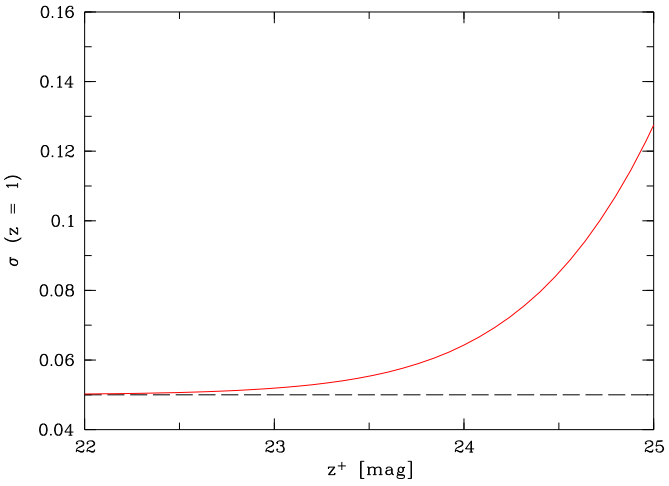

Figure 5.2: Standard deviation of the Gaussian function used as basis for the scatter of the model led photometric redshifts for a galaxy at $z = 1$ as a function of apparent z-band magnitude. The scatter gets significantly larger above $z = 21^{\mathrm{mag}}$.

photometric redshift on the object z-band magnitude can be found in the comparison sample, which consists of fairly bright objects. However, the quality of the photometric redshifts is known to decline for fainter objects, also evident in the increase of the width of their probability distributions as can be seen in Figure 5.4. Due to this, a dependence of the modeled scatter on the objects' apparent magnitude has to be introduced. The z-band is chosen as the basis here, with the magnitude dependence tweaked so that the increase in the redshift uncertainty matches the real data:

$$\sigma = (1+z) \cdot \left(0.025 \cdot \sqrt{1 + 10^{(0.9247 \cdot (\mathrm{mag}_z - 24.2))}}\right). \tag{5.2}$$

An example of the behavior of Function 5.2 for a galaxy at $z = 1$ is shown in Figure 5.2. Here, the dependence of the standard deviation of the Gaussian used to model the photo-z scatter is plotted against the apparent z-band magnitude of the object. For bright objects, the width of the scatter is $\sigma \approx (1+z) \cdot 0.025 = 0.05$, the value derived from the comparison between photometric and spectroscopic redshifts of the comparison sample. For fainter objects, the width of the scatter increases; the scatter for a $z = 24^{\mathrm{mag}}$ object is $\sigma = 0.065$, about 30% larger than for bright objects. Figure 5.3 shows the resulting modeled photometric redshifts plotted against the exact values from the simulation, similar to the plots shown in Section 4.7.1.

5.2.3.2 Photometric redshift probability distributions

Another vital input for the cluster search is the probability distribution for the photometric redshifts, given by a Gaussian centered on the photometric redshift. Its width is estimated by the multi-color classification based on the distance to the neighboring templates compared to the best-fitting template in the n-dimensional color space. The

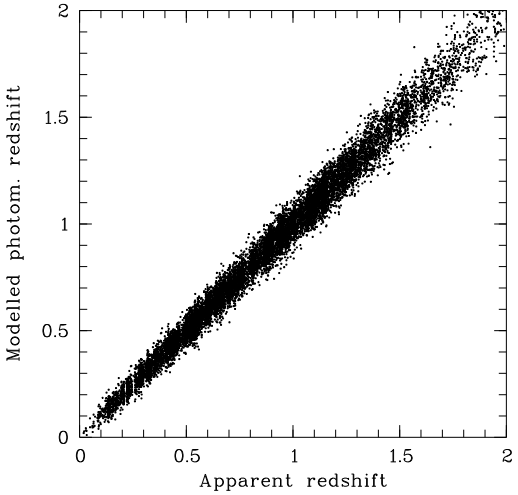

Figure 5.3: Modeled photometric redshifts plotted against the exact apparent redshifts for objects from the mock catalog in the redshift range $0 \leq z \leq 2$.

reliability of the fit of course also depends on the accuracy of the measurements and hence on the brightness of the object. For the mock catalog, the photometric redshift errors have to be modeled to match the real data as closely as possible. They need to resemble the behavior of the real data with respect to three aspects:

- The dependence on the object magnitude
- The dependence on the object redshift
- The overall distribution of their widths

The last condition is not independent of the first two. If both the dependence on magnitude and redshift are modeled perfectly, the overall distribution of the widths of the errors is automatically similar to the real data if the model catalog matches the observed one. Since no model can be perfect, the overall distribution of the widths can serve as a consistency check here.

Even though the multi-color classification can assign a second probability distribution peak to objects whose photometric redshift estimate is ambiguous, only the more significant peak is used in the HIROCS cluster search algorithm. Due to this, all objects in the mock catalog get single peak redshift probability distributions assigned.

To model the probability distributions, two steps are taken. First, a function similar to 5.2 is used to introduce a dependency on both the z-band magnitude as well as the simulated photometric redshift. Its parameters are tweaked to match the behavior of the real probability distributions:

$$\Delta z_{\mathrm{phot},1} = \left(0.012 \cdot \sqrt{0.01 + 10^{(0.1 \cdot mag_H - 13)}}\right) \cdot \left(1 + 2.3 \cdot z_{\mathrm{phot}}^{2.3}\right) + 0.003 \tag{5.3}$$

When applying the above formula, any object's probability distribution is fully determined by its brightness in the z-band and its redshift. In the real data, the situation is a lot more complex. One has to keep in mind that the modeled errors only depend

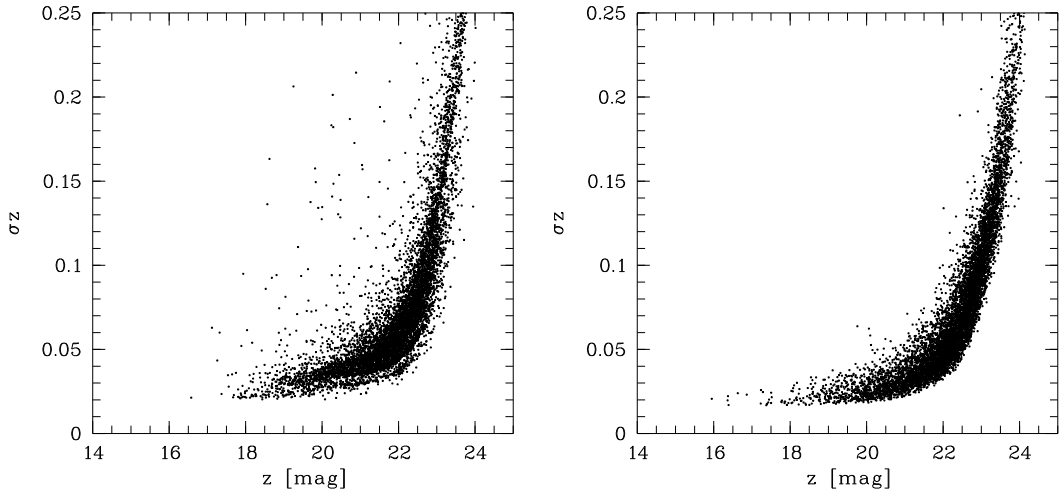

Figure 5.4: Estimated photometric redshift errors as a function of object brightness in the z-band. Left panel: Real data from the COSMOS field. The scatter at the high-magnitude border is due to non-uniform depth of the data. Right panel: Simulated data. Note that the real data used for comparison is the H-band selected catalog used to search for distant galaxy clusters which has photometric redshift errors that have been increased to more conservative values, see Section 6.2.

on the object's z-band magnitude and its modeled photometric redshift, whereas the error estimates in the real data also depend on other object properties. Since they are estimated based on the distance of the measurements to the library objects, they depend for example on the sampling of the parameter space by the library. Peculiar objects for which no exact match in the template library is found will be assigned larger error estimates, even if they are relatively bright, an effect that cannot be modeled here. Instead, a random component is introduced to Δz_{phot} in a second step, consisting of two components. The first one is a pure random value, while the second component also depends on the object's redshift, reflecting the effect that the photometric redshifts in the COSMOS data become more uncertain with increasing redshift due to the low number of near-infrared filters:

$$\Delta z_{\mathrm{phot}} = \Delta z_{\mathrm{phot},\,1} + \mathrm{random}\,1 + \mathrm{random}\,2\,(z_{\mathrm{phot}})\,, \tag{5.4}$$

with both $\mathrm{random}\,1 > 0$ and $\mathrm{random}\,2\,(z_{\mathrm{phot}}) > 0$.

The behavior of the simulated redshift probability distributions has to match the real data as closely as possible in the three criteria given above. The resulting scatter is plotted as a function of z-band object magnitude in comparison with real data in Figure 5.4. It should be noted that the real data used for comparison is the H-band selected catalog used to search for distant galaxy clusters which has photometric redshift errors that have been increased to more conservative values, see Section 6.2. The modeled estimates of the photometric redshift errors agree very well with the distribution of the real errors.

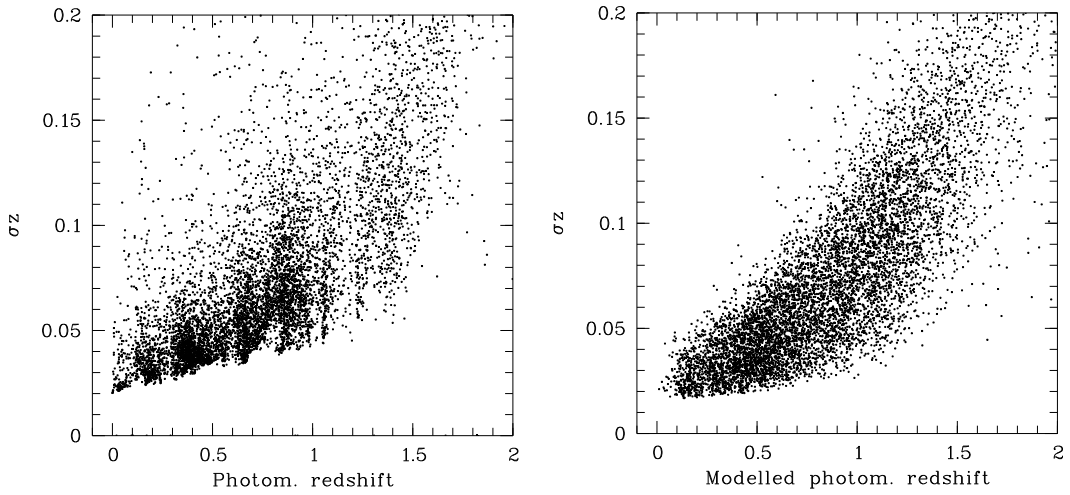

Figure 5.5: Comparison of the photometric redshift errors as a function of photometric redshift. Left panel: Real data from the H-band selected catalog in the COSMOS field. Right panel: Simulated data.

Figure 5.5 compares the dependence of the redshift probability distributions on the photometric redshifts for the real data (left panel) and the mock data (right panel). The overall agreement is quite good. The almost linear increase of the lower boundary between $z = 0$ and $z = 1.5$ for the real data could not be modeled. While in the plots it looks like the model overestimates the width of the probability distributions above $z \approx 0.7$, this is not the case. The average values for the real and the mock data agree well for the redshift range $0.5 \leq z \leq 1.5$ relevant for the cluster search. However, the distribution of values is somewhat different, with the real data being more extreme in the number of small-width objects as well as the maximum probability distribution widths.

As explained above, while not being independent from the two comparisons between the photometric error estimates of the real and the mock data given above, their overall distributions can serve as a consistency check. Figure 5.6 shows the histograms of both error estimates overplotted, with the black line for the real data and the red line giving the distribution for the mock data. The area of the mock catalog has been reduced to 0.6 square degrees here to match the object count in the real data. While having a good overall agreement, the width of the modeled photometric redshift probability distributions slightly overpredicts the real values, as can be seen in the deficit of mock objects with error estimates between 0.04 and 0.08 together with their excess between 0.08 and 0.18 when compared to the real data. Overall, the similarity between the properties of both datasets is pleasing.

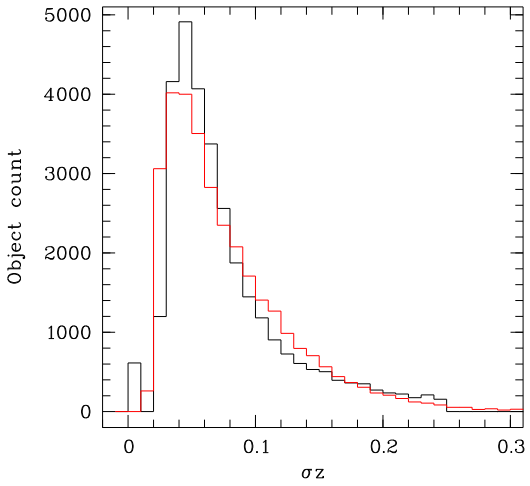

Figure 5.6: Histogram of the data plotted in Figure 5.4, with the area of the mock catalog reduced to 0.66 square degrees to match the object count in the real data. Black line: Objects from the H-band selected object catalog in the COSMOS field with more conservative error estimates assigned. Red line: Modeled errors in the mock catalog. The modeled photometric redshift error estimates slightly overpredict the real values, with a deficit between 0.04 and 0.08 and an excess between 0.08 and 0.18.

5.2.4 Cluster search on the mock catalog

5.2.4.1 Selection of an object catalog from the mock sky

In order to get a realistic assessment of the cluster finding methods capabilities and properties, the mock catalog used for comparison has to resemble the real data as closely as possible. In concordance with the catalog used to perform the distant cluster search, and representing the envisaged final depth of the HIROCS COSMOS data set in the H-band, a magnitude cut is applied to select a subset of the mock catalog. Ideally, this cut should be chosen to be $H = 21.4^{\mathrm{mag}}$, same as for the catalog used to search for distant clusters in the COSMOS field. Unfortunately, as discussed above, there is an unresolved issue with the observed H-band magnitudes in the mock catalog. The objects H-band magnitudes do not play a special role in the analysis of the mock data in later steps, so the near-infrared magnitude issue is of no importance if another criteria can be found to select an object catalog with similar properties. The simulated z-band magnitudes do not suffer from the effects of the dust absorption model, so a brightness cut in z is defined in a way that its effect is the same as the cut in H. This cut cannot be constant, instead it has to have a slope towards fainter objects at higher redshifts. A magnitude limit

$$\mathrm{maglimit}_z = \mathrm{mag}_z - z_{\mathrm{phot}} - 22.25 \tag{5.5}$$

is chosen for this, with its location being close to the completeness limit in the z-band for the H-band selected sample. Figure 5.7 shows a comparison of the real data plotted in black with the mock sky data plotted in red. The overall agreement in object brightness

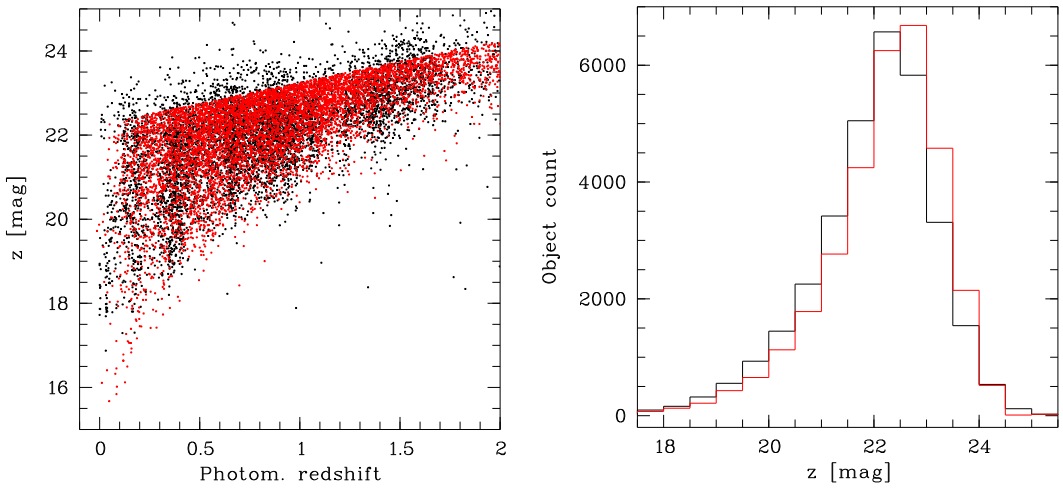

Figure 5.7: Left panel: Observed frame apparent z-band magnitudes plotted vs. photometric redshift for the H-selected COSMOS data (black) and the modeled data (red). Right panel: Histogram of the objects shown in the left panel. The agreement between simulations and real data is good, except for the outliers caused by very red objects, with the objects in the mock catalog being slightly fainter on average than the objects in the H-selected catalog.

and number density is good, but with some outliers at fainter magnitudes for the same redshift in the real data than in the mock catalog (left panel). These are mostly very red objects which are bright enough in the H-band to be included in the sample, but fall under the brightness cut in the z-band (see also Figure 5.12). This possibly affects the cluster detection in the mock catalog, especially at higher redshifts, since the very red objects predominantly populate the massive halos. The magnitude histogram of all objects in the right panel shows that the objects in the mock catalog are slightly fainter than the objects from the H-selected COSMOS catalog on average, but the difference is negligible for the analysis to follow. The mock catalog resulting from the z-band magnitude cut contains 3 063 660 objects on the area of about 51 square degrees in the redshift range $0 \leq z \leq 2$, or 60 071 objects per square degree. The difference in number density compared to the real H-band selected catalog in the COSMOS field, which contains 49 694 objects per square degree, is about 20%.

Figure 5.8 shows the distribution of the absolute rest-frame V-band magnitudes as a function of redshift for the mock catalog. One needs to note that in the simulation data, the rest-frame magnitudes are given without dust extinction.

If the mock catalog is a good model of the real catalog, the redshift distribution of objects has to be similar. Figure 5.9 shows a histogram of the simulated photometric redshifts in the mock catalog (red line) overplotted on the distribution of photometric redshifts in the H-selected COSMOS catalog (black line). The agreement between the two catalogs is very good. Some large scale structures in the real data showing as spikes at redshifts below unity. The dip in object density at $z \approx 1.1$ in the real data is most likely an artifact of the photometric redshift code. At higher redshifts, the agreement between the two catalogs is excellent.

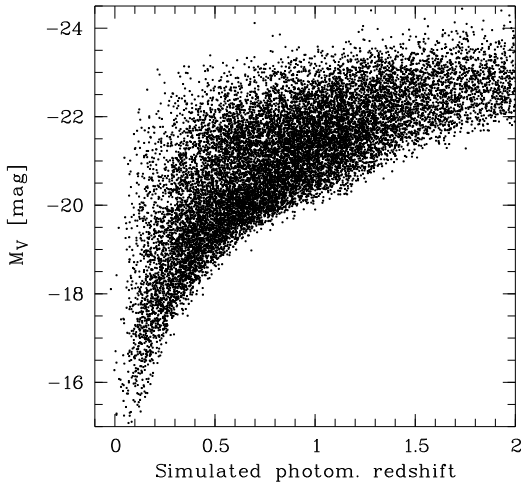

Figure 5.8: Distribution of the absolute rest-frame V-band magnitudes for objects from the z-band selected mock catalog as a function of modeled photometric redshift. Note that these magnitudes do not include dust absorption.

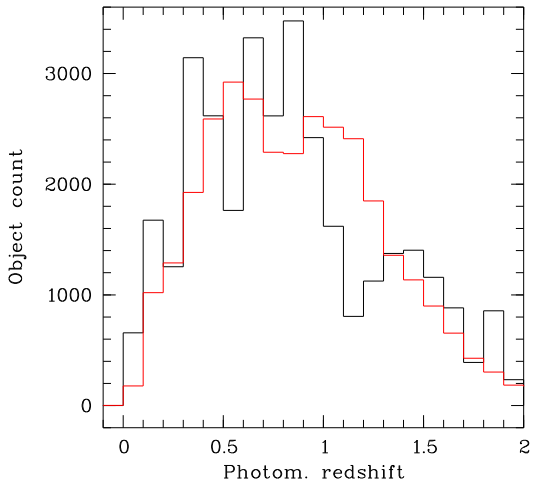

Figure 5.9: Photometric redshift histogram of the objects from the H-band selected COSMOS catalog (black line) and from the mock catalog (red line). The area of the mock catalog has been reduced to 0.66 square degrees to match the H-selected catalog.

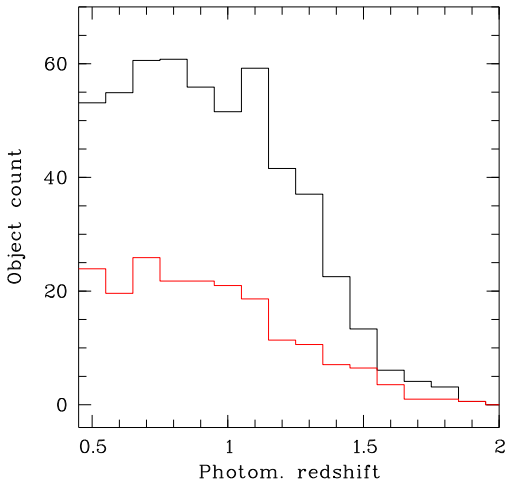

Figure 5.10: Differential redshift distribution of galaxy groups with halo masses $5 \cdot 10^{13}\,\mathrm{M}_\odot \leq M < 1 \cdot 10^{14}\,\mathrm{M}_\odot$ (black line) and galaxy clusters with halo masses $M \geq 1 \cdot 10^{14}\,\mathrm{M}_\odot$ (red line) in the mock catalog reference sample.

In summary, the properties of the mock catalog created using the steps described above are in good agreement with the real data from the H-band selected COSMOS catalog, thus providing the basis for a reliable assessment of the properties of the HIROCS cluster search method.

5.2.4.2 The reference sample

The spectroscopic mock sky is used to create a reference sample containing all clusters and groups with three or more halo members above the z-band magnitude cut in the redshift range $0.4 \leq z \leq 2.0$. To create this reference sample, the cluster selection algorithm is run on the spectroscopic mock sky selecting objects which are close together in projection on the sky, redshift and live in a parent dark matter halo of the same mass. A small redshift range has to be used here for cluster member selection because the cosmological redshifts of objects living in the same halo can be slightly different.

In total, the mock catalog reference sample contains 85299 halos with $N_{\mathrm{gal}} \geq 3$ and with virial masses ranging from $1.7 \cdot 10^{11}\,\mathrm{M}_\odot$ to $1.11 \cdot 10^{15}\,\mathrm{M}_\odot$, 5092 of which are structures with virial mass greater or equal than $5 \cdot 10^{13}\,\mathrm{M}_\odot$ in the target redshift range. There are 1548 clusters above $1 \cdot 10^{14}\,\mathrm{M}_\odot$ in the reference sample. Figure 5.10 shows the differential redshift distribution of galaxy groups with halo masses $5 \cdot 10^{13}\,\mathrm{M}_\odot \leq M < 1 \cdot 10^{14}\,\mathrm{M}_\odot$ (black line) and clusters with halo masses $M \geq 1 \cdot 10^{14}\,\mathrm{M}_\odot$ (red line). At $z \approx 0.4$, structures with four members above the magnitude cut span the mass range $2.7 \cdot 10^{12}\,\mathrm{M}_\odot$ to $1.3 \cdot 10^{13}\,\mathrm{M}_\odot$, with the median being $3.15 \cdot 10^{12}\,\mathrm{M}_\odot$.

A vital information for the cluster search is the relation between the halo mass of a cluster and the number of its members. With the HIROCS search method, the clusters are detected as galaxy overdensities with respect to the field. If a cluster contains too few members, it cannot be detected as an overdensity even if its halo mass is quite

large. Figure 5.11 shows the relations between the number of members brighter than the z-band magnitude cut applied and the virial mass of the dark matter halo. In all redshift bins, the number of galaxies depends linearly on the halo mass, but with a significant scatter of up to one order of magnitude in the number of members for halos of the same mass. Some outliers with a large halo mass but few member galaxies are present, resulting from the halos being truncated by the field borders. Above $z = 1$, there is a significant number of high-mass halos with less than 10 member galaxies, making them very hard to detect with the HIROCS cluster search method. Above $z = 1.2$, the majority of halos falls into this category, even most of the highest-mass ones. Figure 5.12, plotting L^* for passive cluster galaxies vs. z for various observed bands, illustrates that when targeting higher redshift clusters, selecting galaxies by a brightness cut in a near-infrared instead of an optical band can avoid this problem since the brightness decrease with redshift gets less steep with increasing wavelength. Unfortunately, a near-infrared brightness cut cannot be applied to the simulation data used here due to the issues with the near-infrared colors. With real data, HIROCS uses a limit of $H = 20.9^{\mathrm{mag}}$, sufficient to detected galaxies 1^{mag} fainter than L^* at $z = 1.5$.

The large scatter in the number of members for halos of the same mass makes number counts not the most reliable tracer of halo mass. As Figure 5.13 shows, the total luminosity of a cluster is a much better tracer of halo mass than the number of galaxies. The scatter is significantly smaller, except for the highest-redshift halos where the small number of members above the brightness cut increases the scatter substantially.

5.2.4.3 Cluster search on the photometric sample

An ideal cluster catalog, like the reference sample discussed above, would be complete (containing all halos above a certain threshold once and only once), pure (without false positive detections), and unbiased (selecting all halos without preferences for halos with certain properties). With real data, all cluster selection algorithms have shortcomings which need to be characterized. To assess the HIROCS cluster search and selection algorithm performance under realistic conditions, a cluster search in the redshift range $0.4 \leq z \leq 1.4$ is conducted on the photometric mock sample resembling the real H-band magnitude selected data in the COSMOS field. One has to note that the cluster search at higher redshifts, where structures have only very few members above the magnitude cut, is a worst-case scenario. Many very red objects, expected to be present predominantly in groups and clusters, are excluded from the sample by the brightness cut in the z-band which had to be adopted due to the defective near-infrared magnitudes of the simulated galaxies. Due to this and the overall number of massive structures at higher redshifts being very small and thus not statistically significant, the cluster search has to be restricted to $z \leq 1.4$ here.

For each cluster candidate, the following properties are stored:

- The central coordinates in projection on the sky
- The redshift (averaged over all members)
- The number of $> 3\sigma$ overdense members
- The maximum number of members from a single central halo

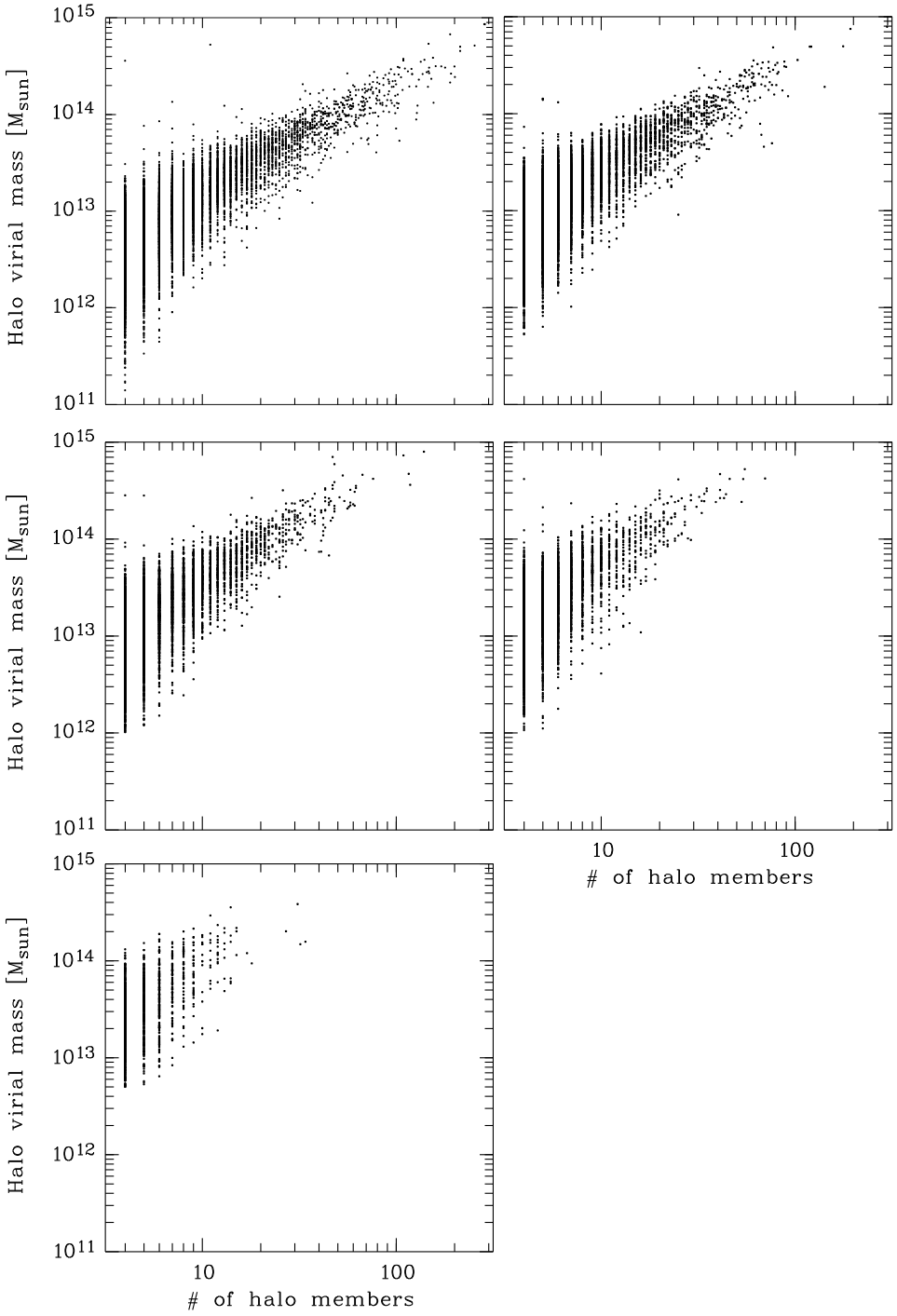

Figure 5.11: Relation between the dark matter halo mass and the number of halo member galaxies above the brightness cut in the redshift slices $0.4 \leq z < 0.6$ (top left panel), $0.6 \leq z < 0.8$ (top right panel), $0.8 \leq z < 1.0$ (middle left panel), $1.0 \leq z \leq 1.1$ (middle right panel), and $1.2 \leq z \leq 1.4$ (bottom panel). Both axes are plotted in logarithmic scale.

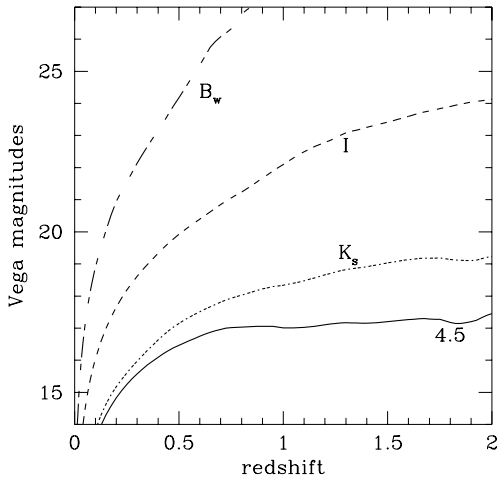

Figure 5.12: L^* for cluster galaxies vs. z, in the observed B_W (long dashed), I (short dashed), K_s (dotted), and IRAC 4.5μ (solid) bands, adapted from (Eisenhardt et al. 2008). The curves are based on a Bruzual and Charlot (2003) model where stars are formed in a 0.1 Gyr burst beginning at $z_f = 3$.

- The corresponding halo mass
- The number of central halos contributing members

In the following, the term "halo" is used to refer to the objects in the reference sample, while "clusters" is used for the groups of overdense objects found in the photometric catalog.

Cluster-halo matching and false positive rate The member galaxies link back the clusters found on the mock sky to the dark matter halos in the simulation. A cluster is considered to be a valid detection if it fulfills one of the following two conditions:

- $\geq 50\%$ of its members stem from the same central halo, which has at least four members above the brightness cut.

For the majority of these detections, the cluster position and redshift are in excellent agreement with those of the halo contributing the majority of the members. Some exceptions are caused by filaments of the most massive halos being found as individual clusters, causing an offsetted center with respect to the parent halo. Requiring at least four members above the brightness cut in the parent halo ensures that the number of chance matches is small.

Figure 5.14 shows the fraction of clusters from the total HIROCS catalog selected by the above criterion as a function of overdense cluster members. At all redshifts, less than 60% of the clusters have more than half of their members stemming from a single parent halo. There is a trend towards a lower fraction of true members per cluster with increasing redshift, caused by the increased scatter in photometric redshift which makes member selection less secure. At $z = 0.5$ and $z = 0.7$, 56% of clusters in the catalog

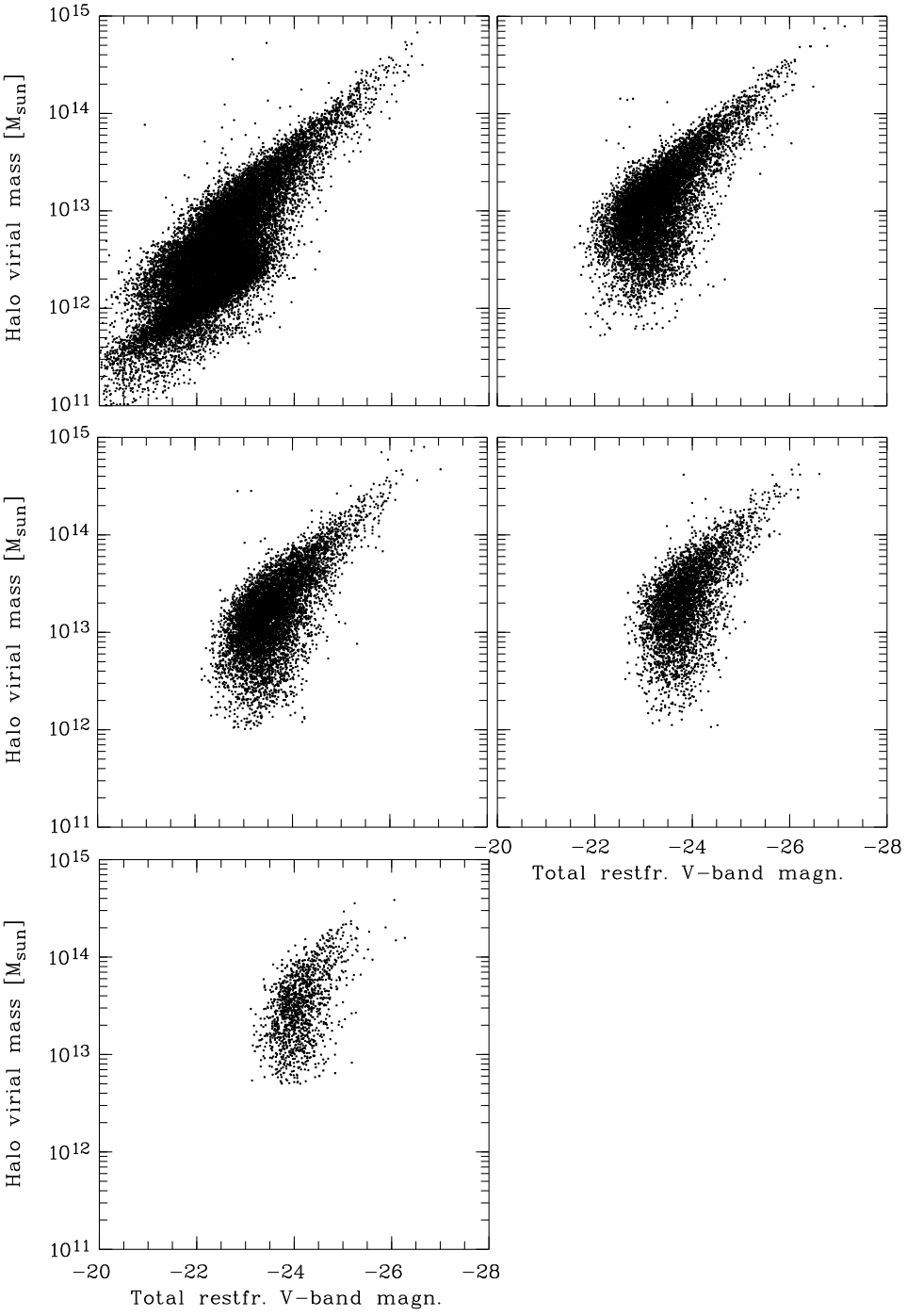

Figure 5.13: Relation between the dark matter halo mass and the total cluster luminosity in the redshift slices $0.4 \leq z < 0.6$ (top left panel), $0.6 \leq z < 0.8$ (top right panel), $0.8 \leq z < 1.0$ (middle left panel), $1.0 \leq z \leq 1.1$ (middle right panel), and $1.2 \leq z \leq 1.4$ (bottom panel).

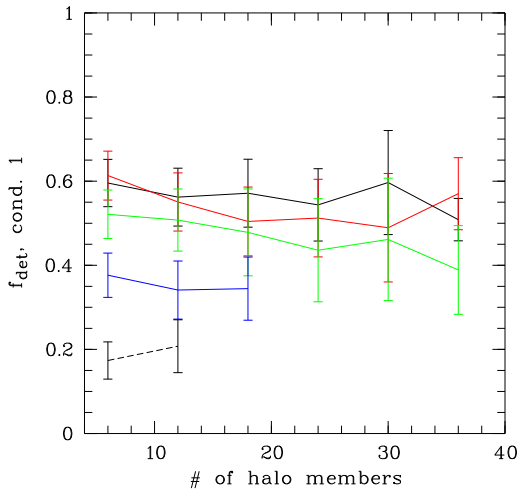

Figure 5.14: Fraction of cluster candidates with $\geq 50\%$ of their members stemming from a single central halo as a function of overdense cluster members for $z = 0.5$ (black line), $z = 0.7$ (red line), $z = 0.9$ (green line), $z = 1.1$ (blue line), and $z = 1.3$ (dashed black line). The last number of members data points also contain all structures above this member count. In the $z = 1.1$ and $z = 1.3$ redshift bins, the richness cutoff has to be placed low because the number of halos above the member counts shown is not statistically significant. The fraction missing to unity is the false positive detection rate. Errors are 1σ.

meets the criterion. This fraction decreases to 49% at $z = 0.9$, 36% at $z = 1.1$, and 18% at $z = 1.3$. Apart from some outliers caused by low number statistics, only a slight decrease with respect to the number of cluster members is visible, mostly caused by shape asymmetries in the richest clusters. On average, the number of halos contributing members to a cluster is $N_{\rm gal}/3$ (slightly less for rich halos than for poor ones, and with a slightly increasing fraction towards higher redshifts). One has to note that in this and all following analyses, statistics involving only halos with a high number of members are affected by the low overall number of these objects, especially at higher redshifts.

Closer inspection of the remaining clusters shows that the HIROCS detection algorithm can very reliably detect the positions and redshifts of massive halos, but the selection of the individual member galaxies is strongly contaminated because of the scatter in photometric redshifts. To take this into account, a second criteria to select true halo detections is adopted:

- Less than 50% of its members stem from the same halo, but positional and redshift detection are correct: A circle of 1 Mpc in projection on the sky around the central coordinates encompasses $\geq 75\%$ of the halo members and the detected cluster redshift is within ± 0.05 of the true redshift. The parent halo contains at least four members above the brightness cut.

Figure 5.15 shows the fraction of clusters selected by fulfilling at least on of the above criteria in the different redshift bins. There is an increase in the number of false positives

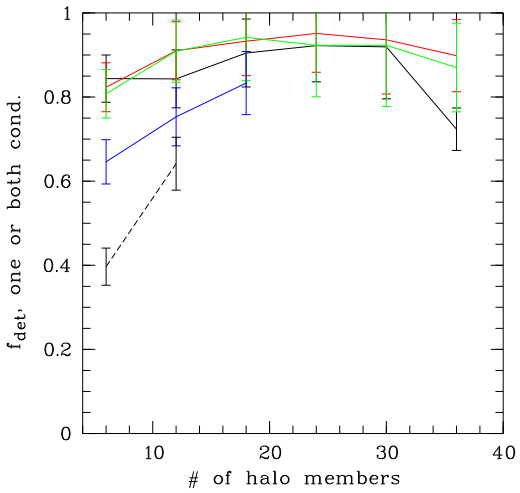

Figure 5.15: Fraction of halos fulfilling at least one of the two selection criteria as a function of overdense cluster members. For details see the description of Figure 5.14.

in the $z = 1.1$ and $z = 1.3$ bins visible; performance at the lower redshifts is comparable. There is also an increase in the fraction of false positive detections for richly populated, massive halos. The reason for this is that parts of the most massive halos are being detected as individual clusters. Some of them who do not have 50% or more of their members stemming from the central halo are not fulfilling the second criteria as well. Consisting only of a part of the central halo's members, their central position is offsetted with respect to the parent halo center. Due to this, the 1 Mpc circle around their center does not encompass 75% or more of the parent halo members. This is most likely to happen for the most extended, massive halos containing most members, thus creating the observed increase in the false positive detection rate.

On average, clusters selected with the first criterion tend to underestimate the true number of halo members, while clusters fulfilling only the second condition are biased high.

In the light of these results, a different approach to select cluster members was tried. The galaxies above the overdensity cut were used only to determine the center coordinates and redshift of the clusters. Next, all galaxies living within a circle of 1 Mpc radius around the cluster center and being at most 1σ of their photometric redshift probability distribution away from the cluster redshift were selected as cluster members. The contamination rates achievable with this procedure were not lower than by selecting the overdense galaxies as members, so this approach was discontinued.

As Figure 5.14 and Figure 5.15 show, the false positive rate does not depend strongly on the minimum number of overdense objects, as long as this is kept in a reasonable range. Thus, $N_{>3\sigma} = 6$ is chosen as the minimum number of members per candidate to include it in the list of clusters. The catalog created from the mock sky comprises 3018 clusters in the redshift range $0.4 \leq z \leq 1.4$. Above $z = 1.4$, the results are no longer statistically significant and most likely substantially affected by the need to apply the

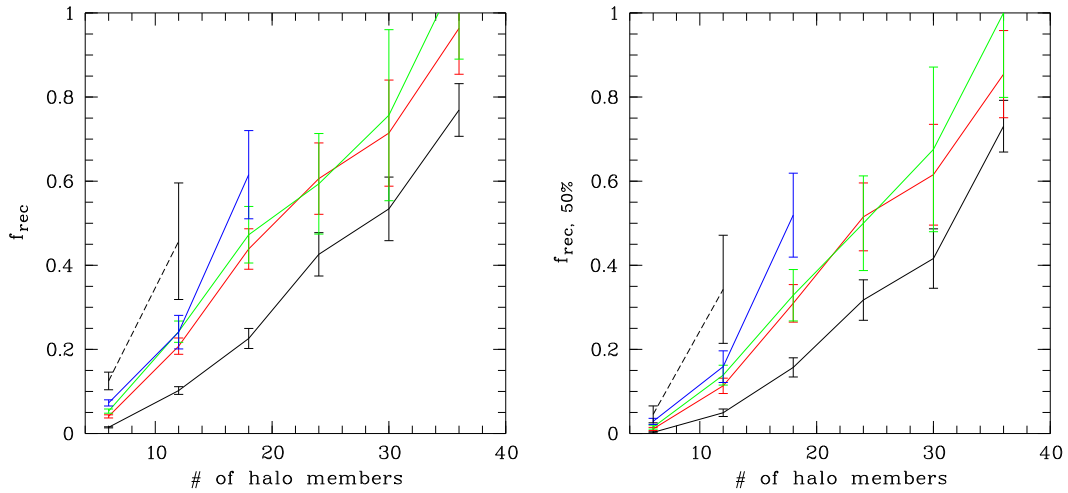

Figure 5.16: Fraction of real clusters recovered with respect to the total number of halos as a function of the number of halo members. Left panel: All detected clusters; right panel: clusters with $\geq 50\%$ of their members stemming from the same central halo. Black line: $z = 0.5$; red line: $z = 0.7$; green line: $z = 0.9$; blue line: $z = 1.1$; dotted black line: $z = 1.3$. The last number of members data points also contain all structures above this member count. In the $z = 1.1$ and $z = 1.3$ redshift bins, the richness cutoff has to be placed low because the number of halos above the member counts shown is not statistically significant. Errors are 1σ.

brightness cut in the z-band, so this redshift range is omitted in the analysis to follow even though it is accessible in principle.

Recovery rate One key characteristic of a cluster search method is its recovery rate, *i.e.* the fraction of true clusters recovered from the data as a function of cluster mass (and hence number of members) and redshift. The recovery rate is closely related to the false positive rate - when using the same search method, a very sensitive search with a high recovery rate (*e.g.* down to poorer systems) is also prone to yield a higher number of false positives than a less sensitive one.

It is desirable to characterize the recovery rate of a cluster selection algorithm with respect to the cluster / halo mass. However, the quantity underlying the HIROCS cluster selection method is the number of halo members, and as is shown in Figure 5.11, there is no one-to-one correspondence between the halo mass and the number of members. In fact, the scatter in the virial mass for halos with the same number of member galaxies can be up to an order of magnitude. So as a first step, the recovery rate based on the number of halo members is plotted in Figure 5.16 for the five redshift bins. Within each redshift bin, the difference in recovery rate between the two selection criteria is most obvious for halos with a smaller number of members. The majority of reliably detected rich halos fulfills both selection criteria, making selection functions based on either one perform almost equally well.

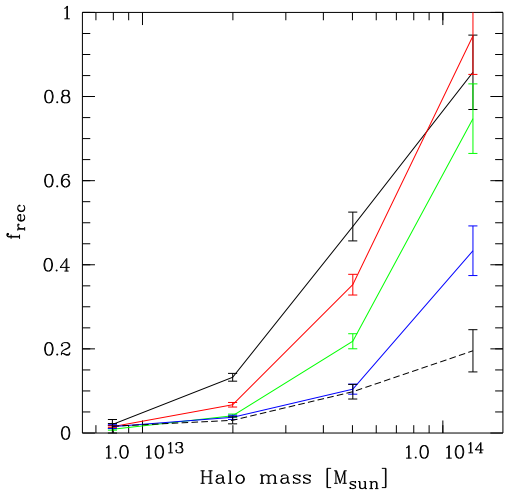

Figure 5.17: Fraction of recovered mock sky halos as a function of the halo virial mass in the redshift bins $z = 0.5$ (black line), $z = 0.7$ (red line), $z = 0.9$ (green line), $z = 1.1$ (blue line), and $z = 1.3$ (dotted black line). The last mass bin also contains all structures more massive. Errors are 1σ.

The redshift trend of the recovery rates seems to be counter-intuitive at first: At low redshift, a lower fraction of halos of the same richness is recovered than at high redshifts, even though because of the brightness cut halos should be more readily accessible here. But in return, the fact that at lower redshifts halos of the same mass contain more members above the brightness cut than at higher redshifts is the cause of this behavior because they represent a totally different mass range. For example at $z = 0.5$, halos with 10 members correspond to halo masses from $2.0 \cdot 10^{12}\,\mathrm{M}_\odot$ to $5.5 \cdot 10^{13}\,\mathrm{M}_\odot$, whereas at $z = 1.3$, a halo with 10 member galaxies above the brightness cut has a virial mass in the range $1.8 \cdot 10^{13}\,\mathrm{M}_\odot \leq M \leq 1.8 \cdot 10^{14}\,\mathrm{M}_\odot$. So while having the same number of members, a $z = 1.3$ halo consisting of 10 galaxies represents a much more massive and thus much rarer object than its counterpart at $z = 0.5$. The 3σ overdensity cut used to identify clusters is calculated with respect to the mean density in the field at the respective redshift, thus the massive high-redshift halos are more likely to be represented by a group of overdense objects than low-redshift halos of the same richness. On average, there are 7.3 halos per square degree at $0.4 \leq z \leq 0.6$ with ≥ 10 member galaxies, compared to just 0.3 at $1.2 \leq z \leq 1.4$. Figure 5.17 shows that the cluster recovery rate with respect to mass scales with redshift as expected: Less structures of the same mass are found at higher redshifts compared to lower ones, due to the reasons given above.

This also influences the mass selection function of the cluster search. Figure 5.17 presents the fraction of halos recovered by the cluster search as a function of halo mass in the different redshift bins. Here, the recovery rate scales with redshift as expected: The cluster selection is able to recover less massive halos at lower redshifts than at higher redshifts. In the $z = 0.5$ bin, a recovery rate of at least 20% is achieved for halos with virial mass $\geq 3.1 \cdot 10^{13}\,\mathrm{M}_\odot$, whereas in the $z = 1.3$ bin, the virial halo mass limit needs to be at least $2.0 \cdot 10^{14}\,\mathrm{M}_\odot$.

The real detection rate depending on the mass of the halo contributing the plurality of

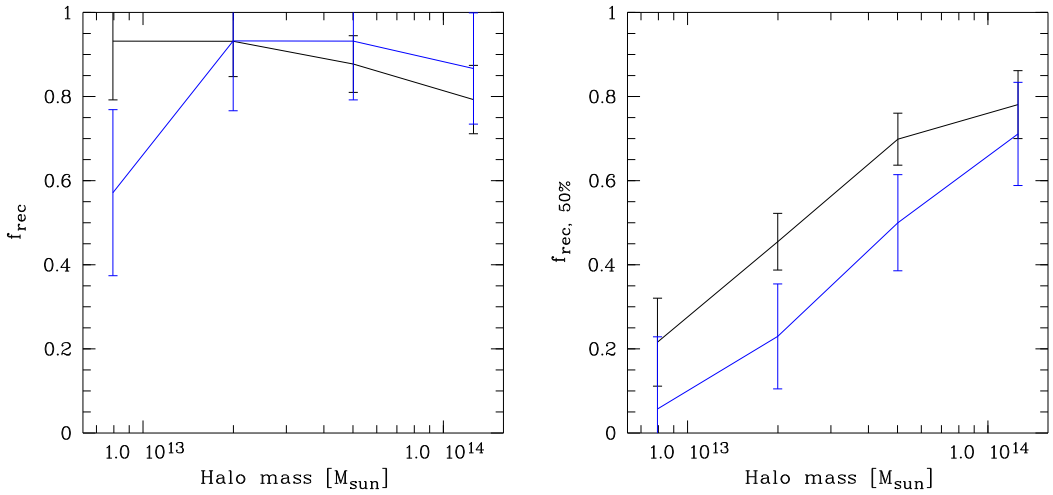

Figure 5.18: Fraction of real clusters with respect to all candidates found as a function of the virial mass of the central halo contributing the plurality of the cluster members. Left panel: All detected clusters; right panel: clusters with $\geq 50\%$ of their members stemming from the same central halo. Black line: $z = 0.5$; blue line: $z = 1.1$. The fraction missing to unity is the false positive detection rate. Errors are 1σ.

the galaxies is plotted in Figure 5.18. The left panel shows the fraction of real clusters fulfilling at least one of the two selection criteria, while only clusters with $\geq 50\%$ of their members stemming from a single central halo are plotted in the right panel. In this plot, the false positive rate is the fraction missing to unity. In these plots, the result is illustrated showing only the two redshift bins $z = 0.5$ and $z = 1.1$. The data in the other redshift bins apparently follow the trend indicated here. However the error bars are too big for the differences to be statistically significant, thus the plot is restricted to two redshift bins for clarity. The fraction of real detections is almost constant around 90% in the low-redshift bins, while at higher z the number of false positives increases for low-mass halos (which are low in the number of members at the same time) because the sample is contaminated by projections of several smaller halos appearing to be one cluster. The HIROCS cluster search is able to reliably select clusters with underlying halos above $M \approx 5 \cdot 10^{13}\,\mathrm{M}_{\odot}$ in all redshift bins. There is a monotonic increase in the fraction of clusters fulfilling the first selection criterion. The slope is comparable for all redshift bins, with the curve being shifted to higher masses for the higher redshift. In the highest mass bin, the fraction of all real detections is nearly the same in both diagrams, meaning that the majority of the most massive halos above $10^{14}\,\mathrm{M}_{\odot}$ that are recovered by the HIROCS cluster search have $\geq 50\%$ of their members stemming from the same central halo.

Halo masses from cluster properties It would be desirable to be able to infer the total mass of a cluster from the properties of its members found by the overdensity search. However, as Figure 5.11 shows, the number of galaxies / cluster mass relation suffers from substantial scatter already if the detections would be perfect, and the difficulties to

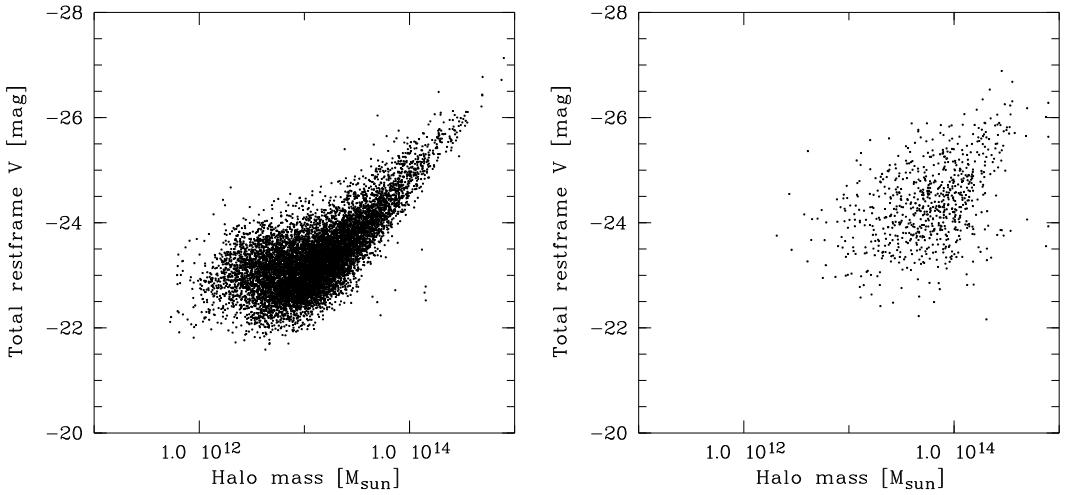

Figure 5.19: Relation between the dark matter halo mass and the total cluster luminosity in the redshift range $0.6 \leq z < 0.8$. Left panel: Halos from the reference sample (same plot as in Figure 5.13). Right panel: Clusters found by the overdensity algorithm.

determine the real cluster members using photometric redshifts can only make it worse. The relation between the total luminosity of the cluster members above the brightness cut and the cluster mass is tighter than the one for the member count, see Figure 5.13. Figure 5.19 gives a comparison between this relation for halos from the reference sample shown in the left panel, and the clusters recovered by the overdensity algorithm in the right panel. While the overall mass / total luminosity relationship is still present in the clusters, the scatter is too big to be able to make precise estimates of the cluster mass based on the total luminosity of the overdense members. Furthermore, the distribution of data points in the two panels do not overlap exactly because the correlation between the number of recovered overdense members on the total number of halo members is often biased low.

Variation of the overdensity limit The choice for the overdensity cutoff, which is used to select cluster candidate members, to be at 3σ above the mean field density is somewhat arbitrary. In first tests done with COMBO-17 data, this cutoff value has worked well, producing no spurious detections on randomly shuffled data and good recovery rates for artificial clusters inserted into the real data.

One of the main advantages of using mock sky data to assess cluster finder properties over inserting artificial clusters into real data is the presence of a correct cosmological context for the clusters. This allows for the determination of a more realistic false positive detection taking into account projection effects. Together with the recovery rate, the false positive detection rate is the most important quantity to trace when doing parameter studies of the effects of different overdensity cutoff limits. In a naive picture, these two main quantities that change under a variation of the overdensity limit can be expected to behave as follows:

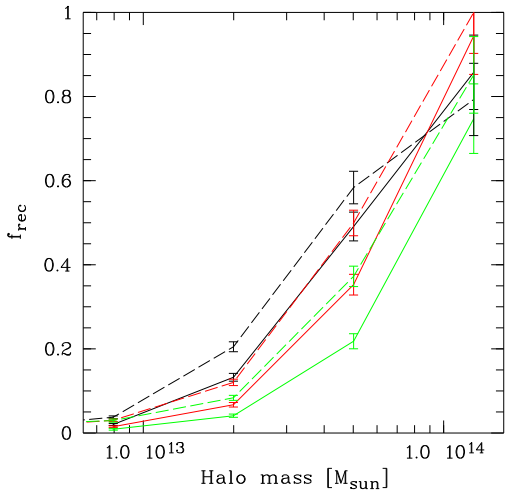

Figure 5.20: Fraction of recovered mock sky halos as a function of the halo virial mass in the redshift bins $z = 0.5$ (black line), $z = 0.7$ (red line), $z = 0.9$ (green line). Continuous lines: 3σ cluster search. Dashed lines: 2.5σ cluster search. Errors are 1σ.

- Lower overdensity limit - higher completeness, but also with higher contamination rate, especially for structures with few members. The lower the overdensity threshold is, the more can the cluster selection be expected to behave like a LSS finder due to smearing out of the clusters with the inclusion of connecting filamentary structures.
- Higher overdensity limit - lower recovery rate, with a lower false positive detection rate.

The optimal value for the overdensity for the real data cut cannot be inferred from the mock sky, because simulations and real data still do not match well enough. The quality of the mock sky is good enough though to demonstrate the behavior of the cluster sample recovered when alternating the overdensity cut. Here, two values are regarded exemplary: An overdensity cut lowered to 2.5σ, and a higher cut chosen to be 4σ.

An overdensity cut of 2.5σ results in a cluster catalog containing 4778 clusters in the target redshift range $0.4 \leq z \leq 1.4$, of which 3653 or 76% are real detections. 1963 or 41% have $\geq 50\%$ of their members stemming from one single halo. These numbers have to be compared to those from the 3σ search recovering 3018 clusters in the target redshift range, of which 2473 or 85% are real detections, and the fraction of clusters with $\geq 50\%$ of their members stemming from a single halo is 51% here. As expected, the lower overdensity cut has a higher recovery rate than the search conducted with the 3σ limit. The recovery rates based on halo mass for the cluster searches with overdensity limits of 3σ and 2.5σ are compared in Figure 5.20, showing the better performance of the 2.5σ cluster search in this respect. One needs to keep in mind here that this comes at the price of a higher contamination of the sample.

With an overdensity cut of 4σ, the cluster catalog comprises 1312 clusters at $0.4 \leq z \leq$

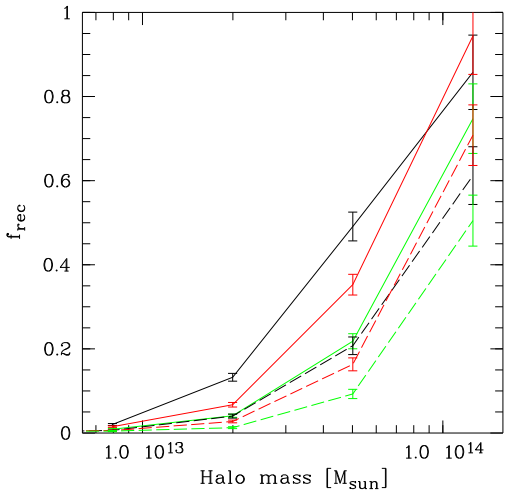

Figure 5.21: Fraction of recovered mock sky halos as a function of the halo virial mass in the redshift bins $z = 0.5$ (black line), $z = 0.7$ (red line), $z = 0.9$ (green line). Continuous lines: 3σ cluster search. Dashed lines: 4σ cluster search. Errors are 1σ.

1.4. Of these, 1155 or 88% are real detections, with 765 or 58% having $\geq 50\%$ of their members contributed by a single halo. These fractions are only marginally better than for the 3σ search. From the total number of real detections in the catalog it is apparent that the recovery rate of the 4σ search has to be significantly worse than the one from the 3σ search. Figure 5.21, giving the fractions of recovered halos with respect to their masses, illustrates this: The dashed lines, representing the 4σ cluster search, are well below the continuous lines from the 3σ search in all redshift bins. The shape of the selection function is very similar in both cases, but the 4σ search can only select more massive systems than the search using the 3σ limit.

This brief parameter study shows the effects of varying the overdensity cutoff limit used for cluster member selection. A tradeoff has to be made here between achieving a high recovery rate and having a low contamination of the sample at the same time. At present, the simulated mock sky cannot be used to calibrate the overdensity cutoff for the real cluster searches due to several shortcomings in the model like the imprecise near-infrared colors used for the magnitude cut in the real data, but it can be used to illustrate that a value of 3σ for the overdensity limit is within the range of reasonable values yielding good results, and most likely not far away from the optimum.

Case studies: Three examples A case study of three systems gives examples of the different properties of the cluster candidates found and illustrates the complications introduced by blending and projection effects for cluster detection methods employing photometric redshifts. All clusters are among the richest detections and trace the most massive halos with virial masses above $2.5 \cdot 10^{14}\,M_{\odot}$. They appear as reasonable clusters in the projection on the sky and in the redshift histograms.

The first cluster, consisting of 210 galaxies contributed by 47 different halos, is detected

at $z = 0.416$. 153 of its members (73%) stem from a halo of mass $3.07 \cdot 10^{14}\,\mathrm{M}_\odot$ at $z = 0.419$, which has 85% of its 179 member galaxies included in the cluster, making this cluster a very clean detection. Since the halo underlying this cluster detection is very extended, only 66% of its members are included in a 1 Mpc circle around the detected position, thus this halo does not meet the second criterion for a cluster detection and is verified as a true cluster by its member fraction only. The projection of its member galaxies on the sky, shown in the upper left panel of Figure 5.22, has the nearly round, centrally concentrated shape expected for a relaxed system. The histogram of the photo-zs (lower left panel) has a clear peak due to the contribution of the members of the underlying halo. The plot of the precise redshifts in the right panel shows the contamination objects to be evenly distributed in the fore- and background of the cluster.

The second cluster at $z = 0.77$ has 169 members stemming from 64 different halos. Almost all galaxies, 50 out of 55, of its most massive contributing halo with $M = 2.84 \cdot 10^{14}\,\mathrm{M}_\odot$ at $z = 0.76$ are included. In this case, the contamination is high, with the fraction of members living in the same halo being only 30%, but with 95% of the members of the underlying halo being included in a 1 Mpc circle around the estimated cluster center, and the redshift and positional detection of the most massive halo being very accurate, this cluster is considered a real detection fulfilling the second condition. Figure 5.23 shows similar plots as Figure 5.22 does for the first cluster. In the projection on the sky, the cluster looks slightly extended, but reasonable with its core slightly east of the barycenter of its members. The histogram of photometric redshifts does not peak as clearly as the one from the first cluster studied in detail, but this could also be due to photometric redshift uncertainty alone, as can be seen by the photo-z distribution of the members of the most massive contributing halo (dashed black line). The distribution of redshift shows that the contaminating galaxies partially are members of other fairly massive halos very close in redshift to the most massive one, and others from various halos being affected by the photo-z scatter. A contaminated cluster detection like this poses some problems for the selection of targets for follow-up spectroscopy.

The third cluster is detected at $z = 0.508$. Of its 557 members, 241 or 43% come from a very massive $M = 8.63 \cdot 10^{14}\,\mathrm{M}_\odot$ halo at $z = 0.510$, that contributes 85% of its members to the cluster. In addition, there are 173 other halos contributing members. This candidate meets neither of the two criteria and is thus not considered a real detection. The reasons becomes apparent when looking at it in detail in Figure 5.24. In the projection on the sky, it is apparent that the cluster candidate consist of several subclumps, with the center of the most massive contributing halo forming the most prominent one. Other clumps are located along its filaments. Because of the filaments, the central coordinates of the detection, located at the barycenter of the members, are offsetted with respect to the central concentration of the members of the most massive halo. As a result, less than 75% of its members are included in a 1 Mpc circle around the center of the detection. Thus, this candidate also fails to meet the second criterion for a real detection. The photo-z histogram shows a peak at the redshift of the most massive halo, partially due to its members, partially due to contaminating galaxies peaking at the same redshift. The distribution of the member redshifts shows that apart from the most massive contributing halo, there are two other quite prominent halos slightly in the background also contributing members. However, with 19 members the second most

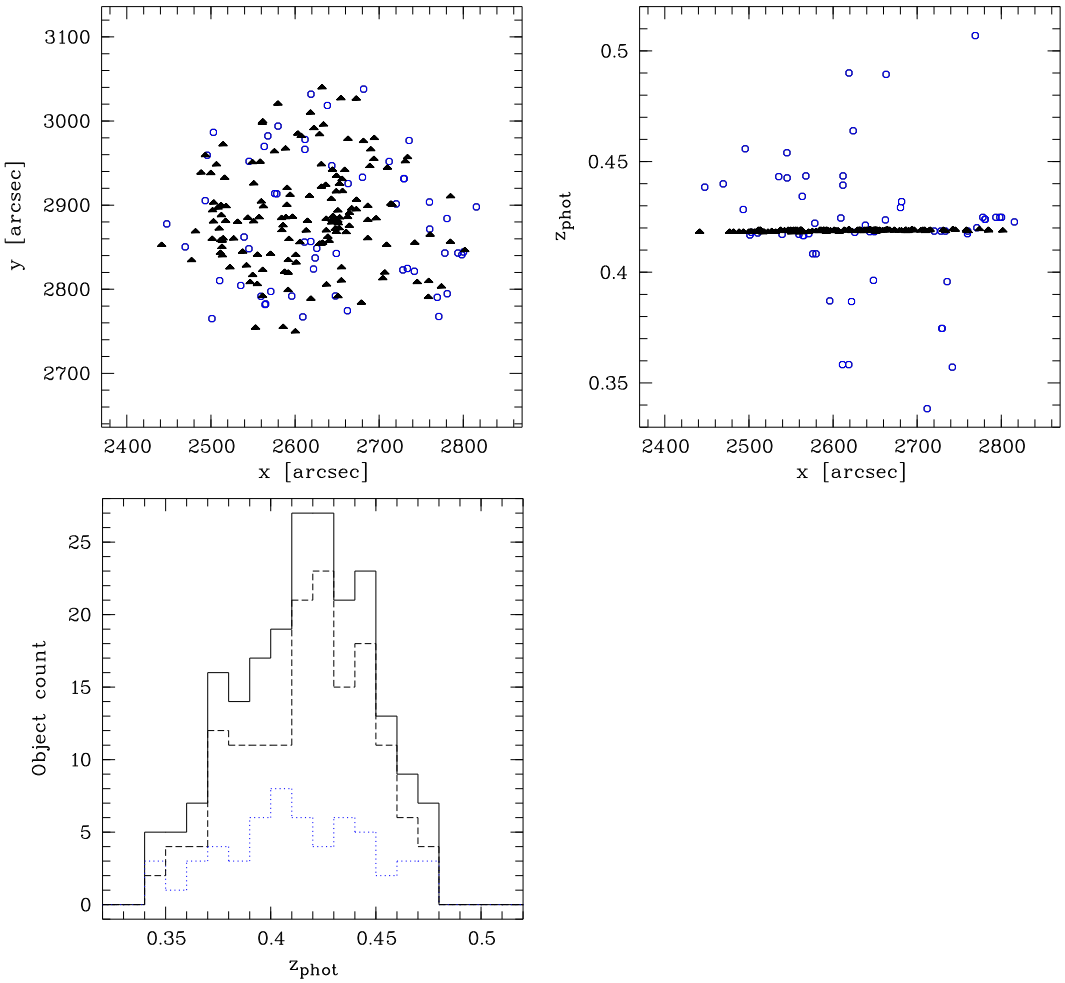

Figure 5.22: Example of a reliable cluster detection at $z = 0.42$ fulfilling the first condition. The projection of member galaxies on the sky is given in the top left panel. A histogram of the photometric redshifts of the clusters members is plotted in the lower left panel (black line: all cluster members; dashed line: members of the underlying halo; dotted line: contamination galaxies), and the right panel shows the distribution of the true redshifts of its member galaxies. Filled triangles symbolize members of the halo contributing the plurality of the members, open circles represent contaminating galaxies.

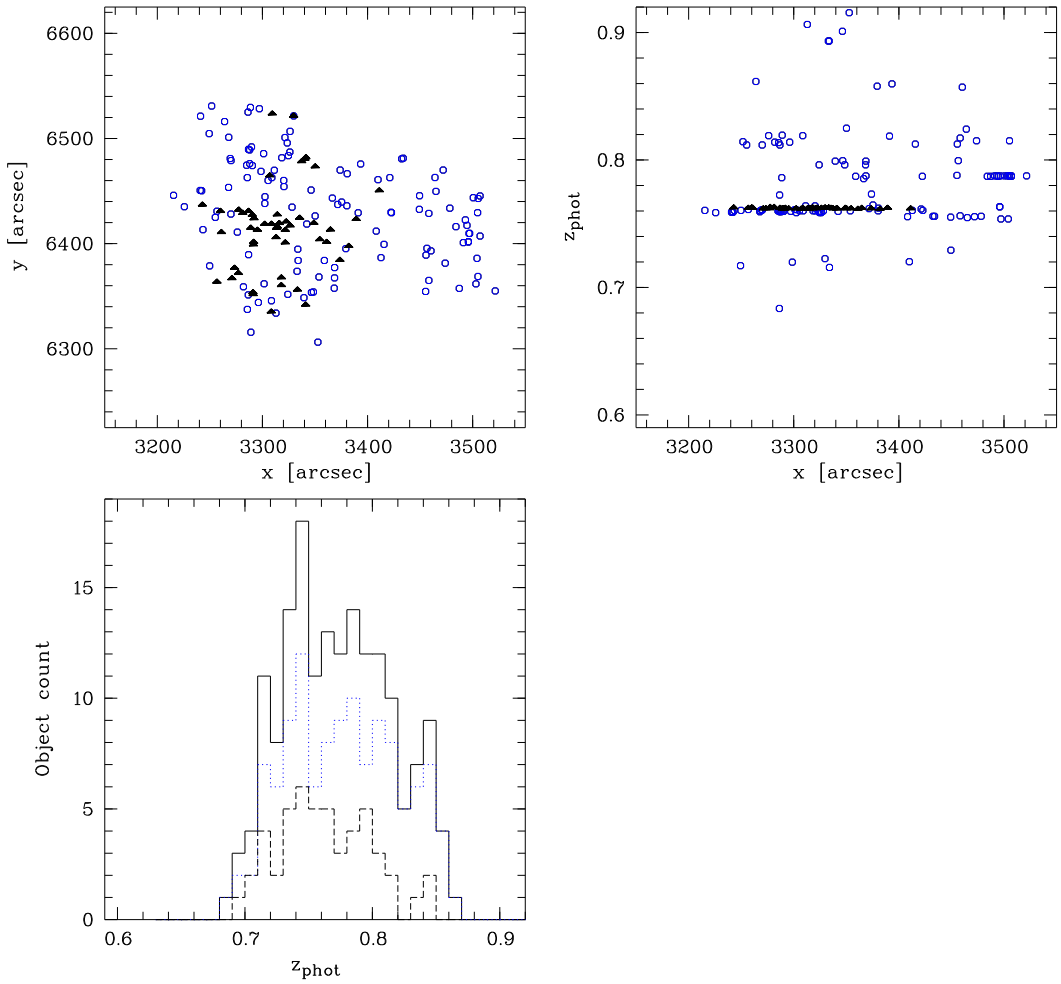

Figure 5.23: A cluster detection at $z = 0.77$, fulfilling the second criterion. For a description of the panels see Figure 5.22.

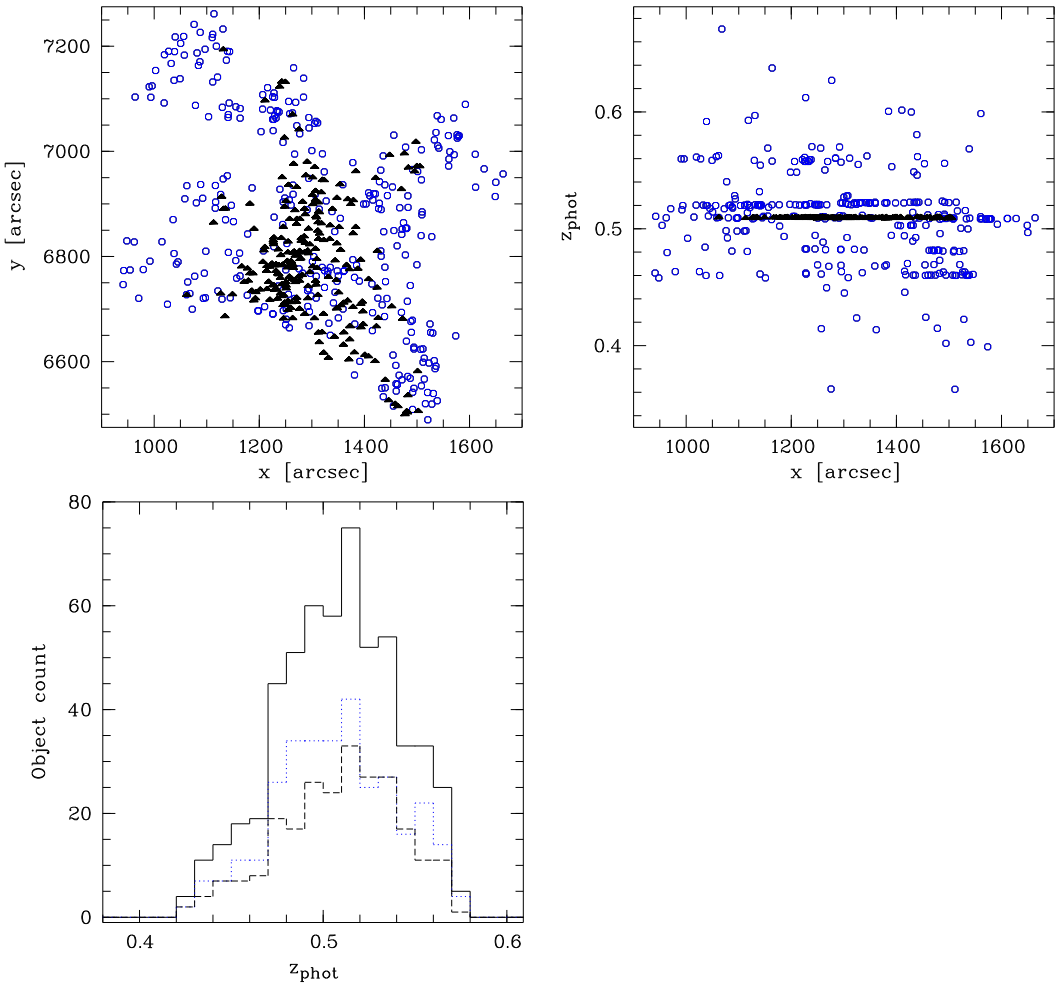

Figure 5.24: A cluster candidate at $z = 0.51$ which meets neither of the two criteria for a real detection. For a description of the panels see Figure 5.22.

contributing halo only plays a minor role for the total member count. This cluster nicely illustrates the issues with the automated detection of very massive halos on the mock sky. Even though this candidate meets neither of the two selection criteria, in reality it would be easier to spectroscopically confirm its most massive contributing halo than in case of the second cluster studied in detail because of its shape in projection on the sky. Here, the observer would naturally pick the main clump as target area for follow-up observations, and in this region the most massive halo contributes the plurality of member galaxies. This example resembles the LSS from the COSMOS field shown in Figure 5.1.

5.2.4.4 Summary and discussion

To select cluster candidates from the galaxies with local densities above the HIROCS cluster search limit, an algorithm based on a friends-of-friends member selection in

projection on the sky as well as redshift has been implemented.

Using a mock sky based on a semi-analytic extension of the Millennium Simulation imprinting galaxies on the dark matter halos, the characteristics of the HIROCS cluster search method are assessed under realistic conditions. The mock sky's properties are modeled to match those of the object catalog used to conduct the HIROCS search for high-redshift galaxy clusters in the COSMOS field as closely as possible.

It is being shown that the position and redshift of halos can be recovered reliably, but the contamination rate of the cluster member galaxies is significant, especially for the lower-mass systems. On average, $N_{gal}/3$ different halos contribute galaxies to a cluster candidate. The false detection rate increases with redshift, from 17% at $z = 0.5$ to 53% at $z = 1.3$. The latter value is probably biased high due to the magnitude cut used to select the object catalog being in the z-band instead of the H-band as in the real data, resulting in red objects predominantly found in galaxy clusters being excluded.

A parameter study illustrates the effect of varying the overdensity limit on the cluster catalog. A tradeoff has to be made between a high recovery rate and a pure sample. It is shown that an overdensity limit of 3σ is a good compromise for the mock sky. The current state of the simulations does not allow to transfer an optimal overdensity limit determined on the models to the real data. Thus, this parameter study is only used to illustrate the effects of varying the overdensity limit.

As an anticipation of the results from the cluster search on the COSMOS field and the comparison with the X-ray based cluster search there, the recovery rate based on the virial mass can be compared with the masses estimated from the X-ray temperatures of the clusters recovered on the COSMOS field. There, clusters with X-ray masses (M_{500}) of $1-4\cdot10^{13}\,\mathrm{M}_{\odot}$ are common in the redshift range $0.7 \leq z \leq 0.9$. On the mock sky, the recovery rate for halos with these masses is low, see Figure 5.17. One has to note here that the masses inferred from the X-ray emission are lower than the total virial masses of the halos due to halo substructures with the conversion factor not known precisely (S. Khochfar, priv. comm.), so no direct translation of the results is possible at the moment.

With the current state of the simulations, the mock sky cannot be used to fine-tune parameters of the cluster search algorithm like the overdensity limit to yield optimal results with the real data. The simulated data is good enough though to allow more general statements about the performance and characteristics of cluster search methods. In a similar attempt, Cohn et al. (2007) have used the Millennium Simulation based galaxy catalog to perform galaxy cluster searches using a red sequence cluster finder on the simulated r- and z-bands employing color and magnitude cuts to select members of massive halos and increase their contrast with respect to the field. They find very high recovery rates (see Figure 4 of their paper) and almost no bias of the cluster sample. The contamination rates are very low, increasing from $\sim 10\%$ at $z = 0.41$ to $\sim 20\%$ at $z = 0.99$. With their filter choice, no conclusions about the performance of red sequence cluster finders at higher redshifts can be drawn. The Cohn et al. study is a best-case estimate since the knowledge about the halo properties of the simulation can be used to precisely select the red sequence galaxies, and no luminosity uncertainties resembling real data are applied to the simulated magnitudes. Furthermore, the number of blue

galaxies in massive halos is underpredicted by the simulation, leading to an artificial increase in the contrast for cluster searches targeting red galaxies.

The significant contamination of the overdense cluster members with fore- and background galaxies is likely to pose problems for follow-up observations, especially for spectroscopy where only a limited number of objects can be targeted. With the photometric redshift uncertainties currently achievable, cluster member selection based on photometric redshifts alone is likely to include many non-cluster galaxies. Since HIROCS aims at conducting a cluster search without prior assumptions about cluster galaxy properties such as colors, the yield of follow-up spectroscopy can be expected to be 50% on average at best for lower-redshift clusters. For high-redshift candidates, this rate may drop to 25% (50% sample contamination in conjunction with 50% contamination of members). Results from follow-up spectroscopy of cluster candidates from the IRAC shallow survey (Eisenhardt et al. 2008) which also selects cluster members based on photometric redshifts alone raise the hope that in reality the contamination rates may be somewhat lower. One needs to keep in mind here though that the object selection in the IRAC shallow survey is done on the 4.5μ images, thus introducing a preference for the very red objects.

6 Cluster search and sample

As described in Section 2.2.2, HIROCS searches for galaxy clusters as overdensities of objects in 3D space (projected RA and DEC plus redshift). Since the reliability of the photometric redshift estimation is vital for this search method, the cluster search in the COSMOS field is done using two different catalogs, depending on the target redshift range, due to the inhomogeneous depth of the H-band data:

- Photometric redshifts without deep near-infrared bands are reliable up to $z \approx 1$, so for the cluster search in the redshift range $0.5 \leq z \leq 1$, the catalog of the whole COSMOS field can be used.
- Above $z \approx 1.1$, deep near-infrared data is required for precise photometric redshifts. Thus, the search for galaxy clusters above $z = 1$ is restricted to the area with sufficiently deep H-band data, since the public COSMOS Ks data which cover the whole field are too shallow for precise redshift estimations.

There exist large samples and quite a number of detailed studies of galaxy clusters at $z < 0.5$, so these are not included in this study.

With the photometric data alone it is not possible to reliably separate virialized clusters from unrelaxed galaxy overdensities. For simplicity, in the following sections the term "galaxy cluster" is used for all candidates found by the HIROCS method. Here, this term also includes overdensities which are possibly not virialized yet and will turn into virialized structures only later in their evolution, and structures usually referred to as galaxy groups.

6.1 Galaxy clusters at $0.5 \leq z < 1$

6.1.1 Catalog selection

The whole 2.1 square degrees of the COSMOS field can be used to search for galaxy clusters in the intermediate redshift range $0.5 \leq z < 1$. Some of the 653 785 objects in the HIROCS master catalog for the COSMOS field are very faint and thus do not have reliable photometric redshifts, as discussed in Section 4.7.2. Selection according to photometric redshift errors does not provide a homogeneous sample, instead a brightness cut has to be used. With the uniform depth of the COSMOS optical data, a clean selection across the field is provided by including only objects with a SNR of five or greater in z^+ in the catalog used to do the intermediate redshift galaxy cluster search. In total, this catalog contains 346 627 objects. All of them have photometric redshifts assigned, and only a handful have photometric redshift errors of $\Delta z_{\mathrm{phot}} > 0.2$. 210 442 of these objects have photometric redshifts below $z = 1.1$.

Comparison with spectroscopic redshifts (see Section 4.7.2) shows that the sample contamination from catastrophic outliers is negligible in this redshift interval. If at all, catastrophic misclassification of objects from this redshift range would lead to them being assigned higher photometric redshifts, thus scattering these objects out of the sample. In principle, this could lead to a reduced significance of the structures targeted by the cluster search. However, with 8% overall outliers, the effect is not very prominent. Its influence is being further reduced by the fact that also the number density of field galaxies at the respective redshift, which the local densities are normalized to, is lowered by this.

6.1.2 Cluster search and catalog

The cluster search is done using a redshift scatter of $\sigma_z = 0.025$, as determined from comparison with the spectroscopic sample (see Figure 4.17; the result of $\sigma_z = 0.023$ there was rounded up). The selection algorithm described in Section 5.1 is used to extract structures with 6 or more members with a local overdensity of at least 3σ.

The resulting candidate catalog comprises 188 structures with redshifts ranging from $z = 0.42$ to $z = 1.05$. A small number of structures outside the target redshift range is included because some of their overdense members are scattered into the target redshift interval, serving as starting points for the selection function (see discussion of the cluster selection procedure). The majority of these are very massive clusters with many member galaxies.

The 2σ redshift interval around the mean structure redshift used to select member objects can result in some members not being included in the structure if their redshifts are outside this range. For some of the most massive structures, there are enough of these outlier objects to be selected as structures themselves, forming fore- and background groups around the massive structures. These spurious groups are cleaned from the candidate catalog by adding their members to the central structure, and all clusters outside the $0.5 \leq z < 1$ redshift range are removed to have a consistently selected sample. After cleaning, 152 clusters are left in the candidate list. This number should be compared to the 420 photometric overdensities of bulge-dominated galaxies found by Finoguenov et al. (2007). There, they are used to identify X-ray emitting clusters by cross-correlating their positions with those of diffuse X-ray emission. In total, they find 72 X-ray clusters at $0 \leq z \leq 1.3$ with this method in the COSMOS field. For a color-blind search like HIROCS, one would expect an even higher number of overdensities to be found. Even when taking into account that the Finoguenov et al. sample of photometric overdensities spans the larger redshift range $0 \leq z \leq 1.3$, the number of HIROCS clusters is still smaller in comparison. This is another indication that the 3σ overdensity cut of the HIROCS cluster search is rather conservative.

As a consistency check, a cluster search is performed on only the passive galaxies. The HIROCS algorithm finds 42 overdensities with $0.5 \leq z \leq 1.0$. The clusters found with this search are a subset of the 152 clusters found in the color blind search; no additional clusters are found. Again, the discrepancy in the total number of clusters is very large compared to the number of galaxy overdensities from the Finoguenov et al. paper. Here, one should recall the results presented in Chapter 5, indicating that the HIROCS

overdensity cut is a good compromise, yielding a high fraction of recovered clusters while still keeping the false positive fraction relatively low.

Figure 6.1 shows the members of all $0.5 \leq z \leq 1.0$ candidates plotted in the projection on the sky, and Figure 6.3 splits the sample into four redshift bins for clarity. The most prominent overdensity in the field is the LSS at $z = 0.73$ north-west of the field center studied in detail by Guzzo et al. (2007), having close to 1000 members above the 3σ overdensity limit. The more massive HIROCS candidates agree very well with the peaks of the LSS found by Scoville et al. (2007b). When comparing their structures to Figure 6.3 one needs to note that in Figure 6.3, all structure members are plotted at the central structure redshift, not including the distribution of individual member redshifts which cause rich structures in (Scoville et al. 2007b) to be smeared out over several redshift bins. The poorer HIROCS candidates do not show up as distinct peaks in the Scoville et al. LSS plots due to being smoothed out by the adaptive smoothing algorithm used to trace the LSS. However, the peak in the differential redshift distribution of the HIROCS clusters at $z \approx 0.9$ and their positions are in agreement with the extended but rather featureless LSS contours in the $z = 0.85$ and $z = 0.95$ panels of Figure 3 in the Scoville et al. paper.

The redshift distribution of the HIROCS candidates is in overall agreement with the distribution of bulge-dominated galaxies from (Finoguenov et al. 2007), shown in Figure 6 there. When comparing both histograms, one needs to note that the cluster redshifts of the two samples differ systematically by about 0.05, with those of the HIROCS sample being lower. Due to this, features in the Finoguenov et al. cluster redshift distribution are found at lower redshifts in Figure 6.8 (see also Section 7.4 for an example).

6.1.3 Comparison with X-ray based cluster selection

6.1.3.1 Structure matching

Comparing the HIROCS cluster sample to the sample selected in X-rays by Finoguenov et al. (2007) allows an assessment of the biases introduced by the two different cluster search methods. Finoguenov et al. identify 72 X-ray emitting galaxy clusters in the redshift range $0.08 \leq z \leq 1.25$ in the COSMOS field. 29 of these fall into the redshift interval $0.5 \leq z \leq 1.0$ probed by the HIROCS cluster search on the full COSMOS field, and 3 additional X-ray clusters can be identified with "redshift outliers" of the HIROCS sample at $z < 0.5$, which are not included in the final cluster list. The X-ray cluster IDs and redshifts from (Finoguenov et al. 2007) are overplotted on the HIROCS clusters in Figure 6.1 and 6.2, respectively. Unfortunately, almost all X-ray clusters above $z = 1$ are located in areas of the COSMOS field which are not part of the distant clusters search due to shallow or nonexistent H-band data, thus the comparison between the HIROCS sample and the one from (Finoguenov et al. 2007) is limited to $z \leq 1$. One $z > 1$ X-ray cluster, no. 102, is included here due to an estimated redshift of $z = 0.89$ in the HIROCS sample.

Of the 29 X-ray clusters in the redshift range probed by the HIROCS cluster search, 21 can be unambiguously identified with HIROCS clusters. The agreement in cluster redshift between (Finoguenov et al. 2007) and HIROCS is usually very good. The

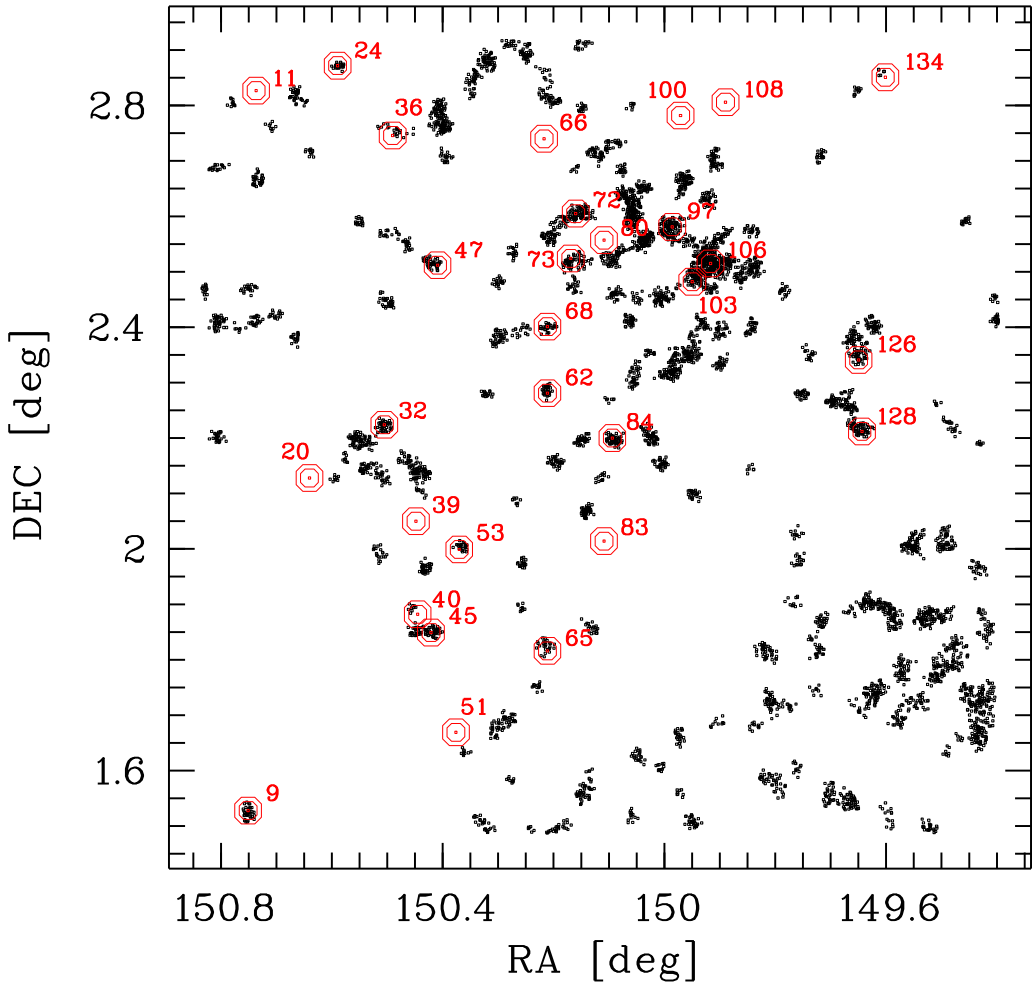

Figure 6.1: Members of the 152 HIROCS clusters in the redshift range $0.5 \leq z \leq 1.0$ in the projection on the sky plotted as black dots. Overplotted in red are the positions of X-ray emitting clusters from (Finoguenov et al. 2007) between $z = 0.5$ and $z = 1.0$. The X-ray cluster IDs refer to Table 1 of their paper. The size of the red symbols does not give the area of the diffuse X-ray emission, the shape should just guide the eye better.

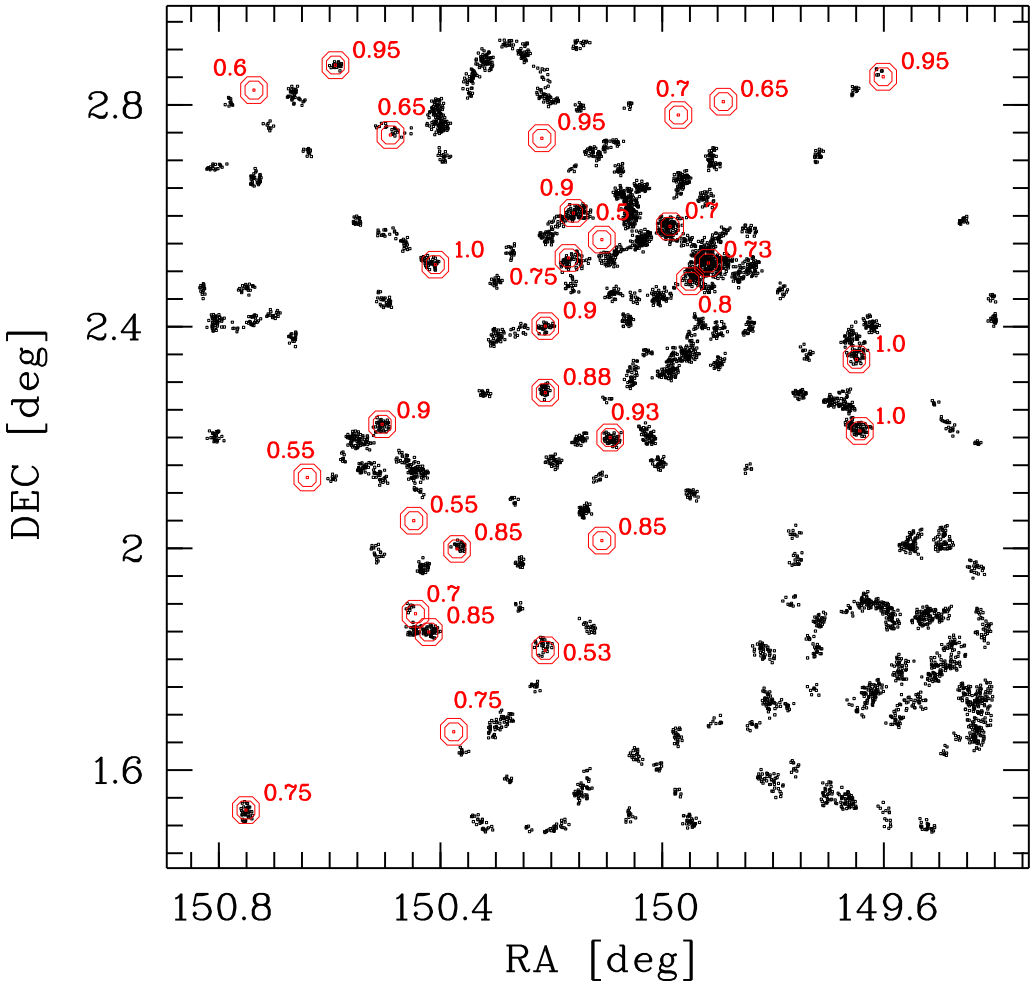

Figure 6.2: The same plot as in Figure 6.1, but with the cluster redshifts from Finoguenov et al. (2007) given.

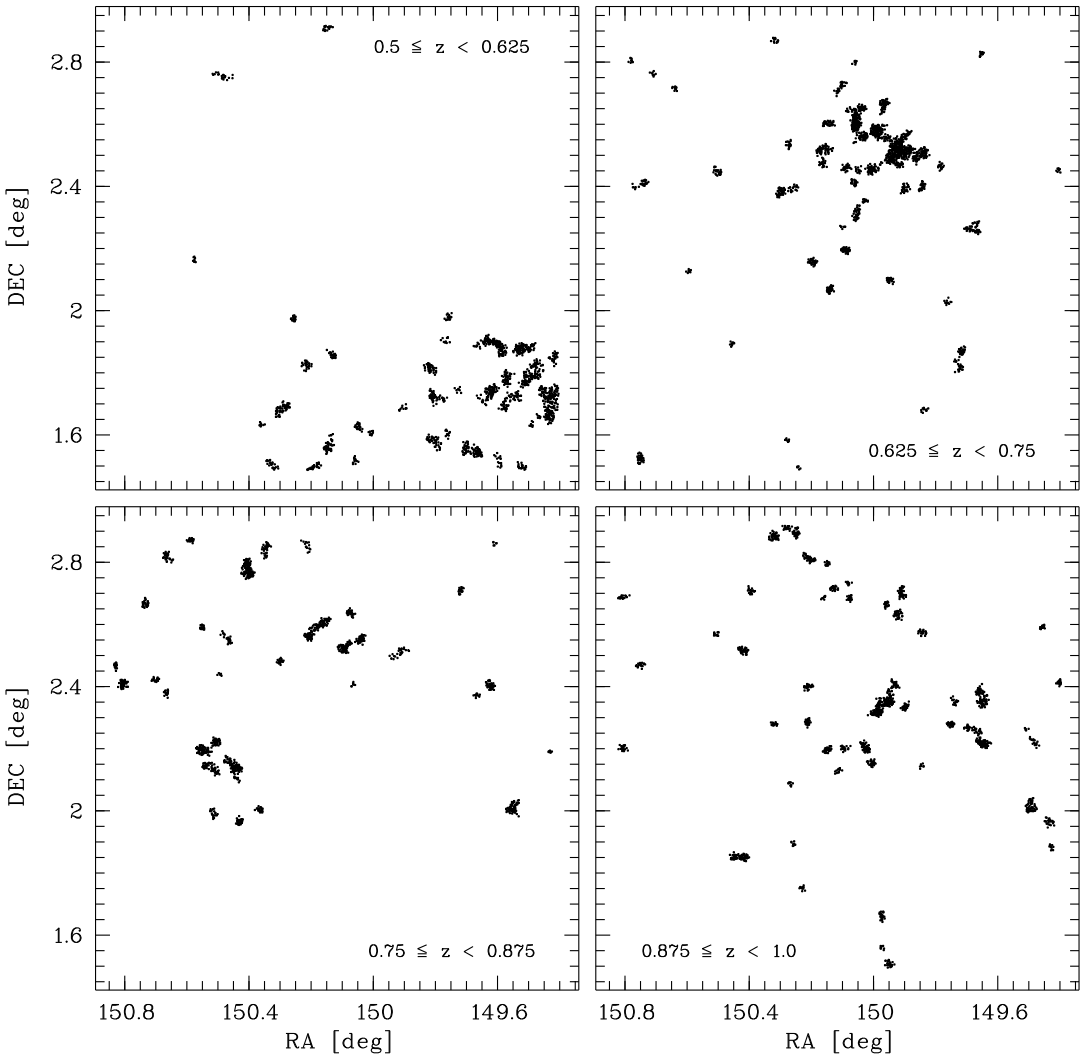

Figure 6.3: The $0.5 \leq z < 1$ HIROCS cluster sample split up into four redshift bins. Top left: $0.5 \leq z < 0.625$; top right: $0.625 \leq z < 0.75$; bottom left: $0.75 \leq z < 0.875$; bottom right: $0.875 \leq z < 1$.

remaining 8 X-ray clusters without counterparts in the HIROCS cluster list are discussed individually in the following, with the IDs referring to Table 1 of the Finoguenov et al. paper:

- X-ray cluster 11, $z = 0.60$: There are no members above the 3σ overdensity threshold of the HIROCS search. The cluster is visible in the HIROCS data when the overdensity cut is lowered to 2σ.
- X-ray cluster 20, $z = 0.55$: No detection in the HIROCS data.
- X-ray cluster 39, $z = 0.55$: Like X-ray cluster 11, visible with an overdensity limit of 2σ.
- X-ray cluster 51, $z = 0.75$: Just 4 galaxies above an overdensity of 2σ are present at the cluster's location. No significant detection even at lower overdensities. No real detection in HIROCS.
- X-ray cluster 66, $z = 0.95$: No detection in the HIROCS catalog.
- X-ray cluster 83, $z = 0.85$: Detection at $z = 1.0$ when the overdensity limit is lowered to 2σ.
- X-ray cluster 100, $z = 0.70$: No detection.
- X-ray cluster 108, $z = 0.65$: No detection.

X-ray clusters 51, 100 and 108 are in the same redshift range as the very prominent LSS at $z = 0.73$. Since the overdensity used for the HIROCS cluster search is calculated with respect to the mean density across the whole field in the respective redshift interval, it could in principle be the case that this LSS with its high density values raises the mean density in a range around its redshift so much that less prominent clusters are not detected above the 3σ overdensity limit. In this case, their members should be visible when lowering the overdensity limit. However, for the three X-ray clusters listed above, no significant galaxy grouping is visible at the X-ray cluster's position even with very low overdensity cutoffs.

3 out of the 8 X-ray clusters without counterparts in the HIROCS cluster list are detected when lowering the overdensity limit to 2σ. This is somewhat unexpected, since X-ray emitting clusters should be virialized and in a mature state of their evolution, thus being rather centrally concentrated, leading to high overdensity values in their central regions. In addition to the low number of HIROCS clusters compared to the 420 early-type galaxy overdensities found by Finoguenov et al. (2007), this again argues for the 3σ overdensity cutoff used to identify clusters in HIROCS being on the conservative side also when dealing with real data. Even though an overdensity of galaxies at the location of diffuse X-ray emission verifies the presence of a galaxy cluster, the three clusters that have no galaxies above an overdensity of 3σ are not included in the further analysis since they are not automatically selected by the HIROCS cluster search. A cross-check of the X-ray candidates with objects from the COSMOS DR1 catalog also shows no significant overdensity of bulge-dominated galaxies for X-ray clusters 20 and 108.

For some clusters the peak in the galaxy overdensity is somewhat offsetted with respect to the center of the X-ray emission, which seems to match better with the location of

the brightest passive cluster galaxies, an interesting effect also noted by (Finoguenov et al. 2007)

A large part of the southeast quadrant contains neither X-ray emitting clusters nor HIROCS-selected galaxy overdensities at $0.5 \leq z < 1$. Additionally, no LSS is detected there by (Scoville et al. 2007b), so this part of the field seems to target a void of the cosmic web. The X-ray data in large areas of the southwest part of the field, where many HIROCS cluster candidates and matching LSS detections in the Scoville et al. (2007b) paper are detected in this redshift range, are rather shallow. If this is the only reason for the absence of X-ray detected clusters there is unclear at present.

The question now arises if there are any galaxy overdensities consisting predominantly of blue members which would thus be omitted by the cluster search conducted by Finoguenov et al. (2007) who match the diffuse X-ray emission with overdensities of bulge-dominated galaxies. Indeed, Figure 2 of their paper shows some X-ray contours from diffuse emission without a matching overdensity of passive galaxies, thus being candidates for this phenomenon. However, a cross-correlation with the HIROCS cluster catalog yields only four objects in addition to the (Finoguenov et al. 2007) sample where the X-ray emission could perhaps coincide with a HIROCS cluster. This is slightly above the likelihood of chance projections, but by no means a high fraction. Due to the lack of detailed information about the nature of the X-ray sources, no further studies about the nature of these objects can be conducted at this point. Results from a follow-up study of a group of four $z \approx 0.85$ candidates, with only one of them detected in the X-rays, are presented in Chapter 8. There are indications that one of them, while being X-ray dark, is in fact a large galaxy group or a galaxy cluster.

6.2 The $z \geq 1$ cluster sample

This section describes a more refined analysis of the search for distant clusters than the preliminary one published in (Zatloukal et al. 2007). While the underlying cluster catalog is the same, cluster selection has been automatized, and additional IRAC data are available. In this analysis, the redshift range of the cluster search has been extended down to $z = 1$. The overall results are still valid, but details change, like the selection of individual clusters.

As explained above, the search for distant clusters of galaxies has to be restricted to areas with sufficient H-band depth. The cutoff was chosen to be $H = 21.4^{\mathrm{mag}}$ (3σ). But just imposing this magnitude cut on the complete HIROCS master catalog does not yield usable results. If one does not restrict the catalog to those H-band pointings with sufficient depth, it will contain objects in areas which are more crowded (the deep pointings), resulting in higher local densities, and objects in sparsely populated areas (the shallow pointings) which have artificially lower local density because of catalog incompleteness. The objects in the shallow pointings distort the local density measurements for the rest of the catalog, because local density is computed with respect to the mean density in the whole field at the given redshift. To avoid this effect, objects have to be selected both by magnitude and by pointing, taking into account only those pointings with sufficient depth.

At the time the analysis for this thesis is done, there are 11 OMEGA2000 pointings in the COSMOS field which reach the desired depth, covering a total area of 0.66 square degrees (see Figure 3.4). Objects are selected in these pointings according to the magnitude cut in H. The resulting catalog contains 32 798 objects down to $H = 21.4^{\mathrm{mag}}$. Because of the extremely deep optical data, all objects have photometric redshifts assigned, and the catalog is essentially complete down to $H = 21.4^{\mathrm{mag}}$.

In this catalog, the photometric redshifts for the old and passive galaxies with $0.35 \leq z \leq 1.2$ are not corrected as described in Section 4.7.2. This does not affect the results since the correction is negligible inside the redshift range used to search for galaxy clusters with this catalog.

In general, high-redshift objects are expected to be relatively faint, and the photometric error estimates given by the multi-color classification are to be regarded as a lower limit. To account for this, the error estimates of the photometric redshifts, which directly influence the local density calculation, are manually corrected to more conservative values than the ones given by the multi-color classification code. A combination of the internally expected error and a lower limit that increases with magnitude and scales with $(1 + z)$ is used:

$$\sigma_{z,\mathrm{eff}} = (1+z) \cdot \sqrt{0.04 \cdot \left(1 + 10^{0.8 \cdot \left(\mathrm{mag}_{z^+} - 22.5\right)}\right) + \left(\frac{\sigma_z}{1+z}\right)^2} \tag{6.1}$$

This formula is motivated by the fact that in the low brightness regime, photon noise translates linearly into photometric redshift uncertainty, whereas at the bright end there is a minimum error. A similar approach is chosen in (Wolf et al. 2004) and described in more detail there. A $z^+ = 24^{\mathrm{mag}}$ galaxy at $z = 1.4$ thus has a median redshift error of $\sigma_{z,\mathrm{eff}} \sim 0.30$. This increase in the photometric redshift error "smears" out the structures in redshift space, lowering their contrast with respect to the field. Since clusters are traced as overdense structures compared to the field, the lowest traceable structure mass would be increased since on average only richer structures contain a sufficient number of objects with overdensities $\geq 3\sigma$, while the false detection rate is expected to be lower.

6.2.1 Comparison with preliminary results

For the sample published in (Zatloukal et al. 2007), all structures with six or more overdense members within a $2'$ aperture were manually picked as cluster candidates from a two-dimensional plot of overdense galaxies in redshift slices to avoid projection effects. The resulting sample comprises 15 cluster candidates with redshifts $1.22 \leq z \leq 1.55$ (see Table 2 of this paper; an ^{L} is added to their IDs to avoid confusion with the sample selected for this thesis). Three of these were rejected from the final list; two (1^L and 11^L) because of low significance after correction for field contamination, and one (7^L) for likely being a projection effect.

For this thesis, the analysis from (Zatloukal et al. 2007) is repeated employing the automated cluster selection algorithm and extending the redshift range above $z = 1$ to supplement the $0.5 \leq z < 1$ cluster search on the whole COSMOS field. The same object table as in (Zatloukal et al. 2007) is used for this analysis.

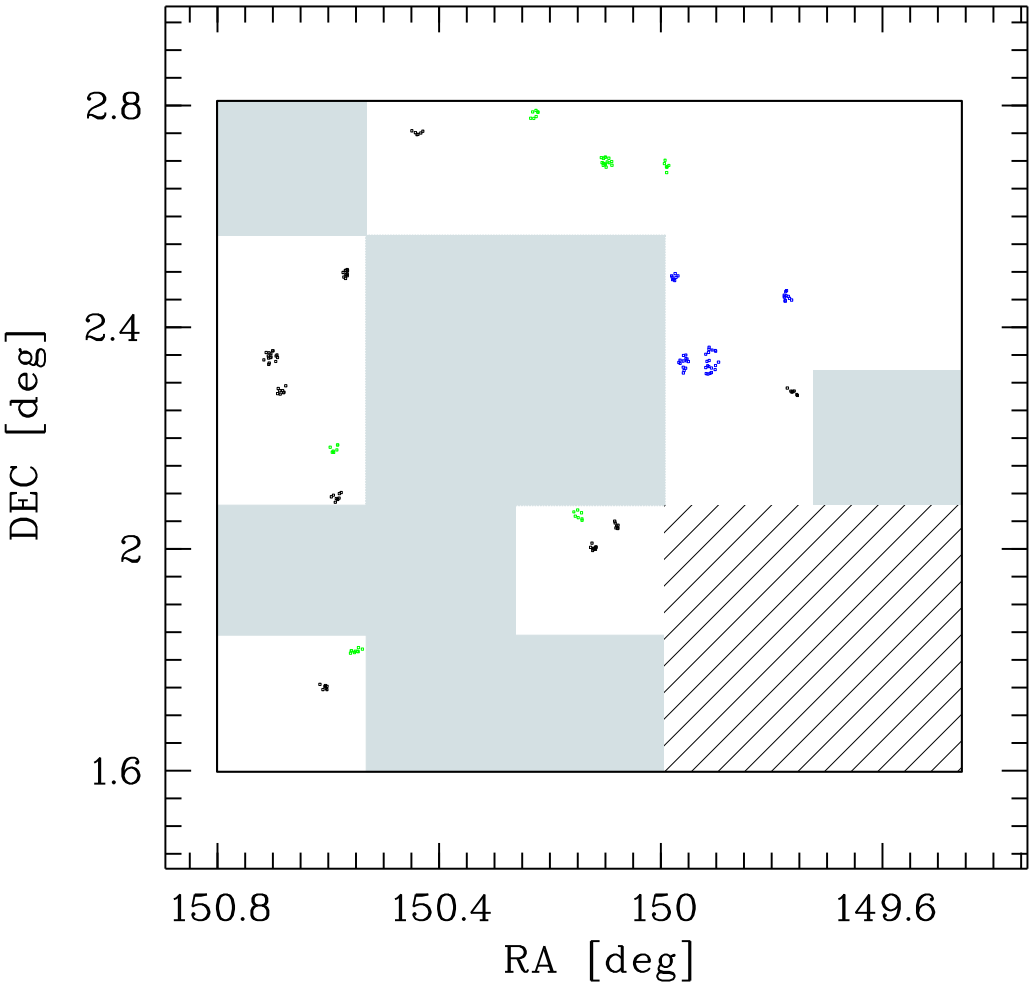

Figure 6.4: HIROCS clusters at $z \geq 1$ in the 0.66 square degrees of the COSMOS field with deep H-band coverage. The H-band will not cover the full COSMOS field, as indicated by the black box. Regions with shallow H-band data are shaded; the southwest quadrant is completely lacking H-band coverage. Black: $1 \leq z < 1.2$; green: $1.2 \leq z < 1.4$; red: $1.4 \leq z < 1.6$.

Table 6.1: Candidates from (Zatloukal et al. 2007) no longer included in the sample extracted by the automatic cluster selection. Column #1 gives the candidate ID referring to Table 2 of the paper, column #2 states a brief explanation why this candidate is not picked by the automatic selection.

ID	Reason
4^L	Two undersignificant subclumps
9^L	Too spread in redshift
10^L	Six galaxies, too far apart in projection
11^L	Too spread in projection

To account for the lower accuracy of the photo-zs for the faint galaxies in the target redshift range, the input value for the expected redshift scatter is increased to $\sigma = 0.035$ in the automated cluster selection. In the 0.66 square degrees covered by the catalog, 20 clusters in the redshift range $1.0 \leq z \leq 1.59$ are selected. Figure 6.4 shows them in projection on the sky. 12 of these are identical to candidates from (Zatloukal et al. 2007); no additional candidates are selected in the redshift range $z \geq 1.2$ since the automated cluster selection is more conservative than the manual one. In fact, only 11 out of the 15 original candidates are being automatically selected since one of them, 14^L, is split up into two components at slightly different redshifts. Table 6.1 lists the candidates no longer included in the sample extracted by the automatic selection with the reasons indicated. The members of one candidate, 9^L, are so spread in redshift that only less than six members are included in the redshift range the automatic selection uses. The three other candidates are too spread in projection on the sky so that they consist of subclumps with none of them being significant enough to be selected as a cluster.

6.3 Object group search on the 4.5μ selected catalog

As shown in Section 4.7.1.2, the IRAC data cannot be used to enhance the quality of the photometric redshifts with the current version of the galaxy template library. However, the object catalog selected on the IRAC 4.5μ-band can be used to perform a consistency check of the cluster search based on photometric redshifts. The idea is to select a subcatalog of objects whose colors are consistent with passively evolving galaxies in the redshift range $1.2 \lesssim z \lesssim 1.7$ and a brightness consistent with cluster galaxies at these redshifts, and search for overdense groups among them. The 4.5μ-selected catalog is an ideal basis for this, since the passively evolving galaxies are especially prominent in the near-infrared IRAC bands.

6.3.1 Color range selection

Even though the galaxy templates' colors involving IRAC bands are not matching the data well enough to permit reliable photometric redshifts, they can still be used to

determine the rather rough color cut used for the object selection envisaged here. Figure 6.5 shows the $(z^+ - 3.6\mu)$ colors of the library templates plotted against redshift at $0 \leq z \leq 2$. Here, $(z^+ - 3.6\mu)$ was chosen instead $(z^+ - 4.5\mu)$ to be comparable to other surveys using red-sequence cluster finding techniques together with IRAC data, like SpARCS (Wilson et al. 2006). The left panel shows only the dust-free objects, with the passive galaxies plotted in red. As to be expected, the passive galaxies in the $z \geq 1$ cluster candidates almost exclusively have dust-free SEDs, so the restriction to the color range of dust-free galaxies for this selection method is valid.

Among the dust-free galaxies, there is no significant degeneracy between the colors of the passively evolving objects and those of galaxies actively forming stars in the target redshift range $1.2 \lesssim z \lesssim 1.7$. However, this picture changes when dust extinction is included, as the right panel of Figure 6.5 shows. There is a color degeneracy between old and dust-obscured objects, no longer permitting a reliable redshift selection based only on object color. Due to this and the rather imprecise selection of objects in a relatively broad color range, no additional candidates will be added to the sample based on this selection. Since HIROCS is intended to be a cluster survey using as little priors about cluster properties as possible, this has never been the goal of this IRAC near-infrared color selection in the first place. However, the sample generated here can serve as a consistency check for the existing candidates in the sense that an overdense group of objects bright in the near-infrared and with colors consistent with passively evolving galaxies, at the position of a cluster candidate found as an overdensity of galaxies in 3D space based on photometric redshifts, is a strong support for its true cluster nature.

Since the galaxy template library is known to not be precise for the colors involving IRAC bands, the color interval used is chosen to be a little larger than inferred from the template library. In addition, a cut in object magnitude is imposed on the catalog: $16^{\mathrm{mag}} \leq 3.6\mu \leq 18.5^{\mathrm{mag}}$, motivated by the brightness of luminous cluster galaxies. This leaves only those objects whose brightness and color are consistent with passively evolving cluster galaxies in the target redshift range.

6.3.2 Search for groups of objects based on the $(z^+ - 3.6\mu)$ color selection

Since the redshift range of the objects is pre-selected using the color cut, the cluster search is this time done in two dimensions only. To be able to still use the local density calculation program, all objects are assigned the same artificial redshifts and redshift errors.

The resulting table contains almost no objects with overdensities of 3σ or higher, so the overdensity cut is lowered to 2σ. Figure 6.6 shows the objects from this selection, marked in red, overplotted on the cluster candidates from the 3D cluster search, plotted in black.

Of the 11 HIROCS clusters in the target redshift range, 6 are in common with groupings of 4.5μ-selected objects that have colors consistent with passive evolution. If the overdensity cut is lowered slightly more, there is also a group of these objects at the position of the HIROCS cluster at $z = 1.55$. There are additional groups of these objects in areas where the H-band data is shallow and no reliable 3D cluster search can

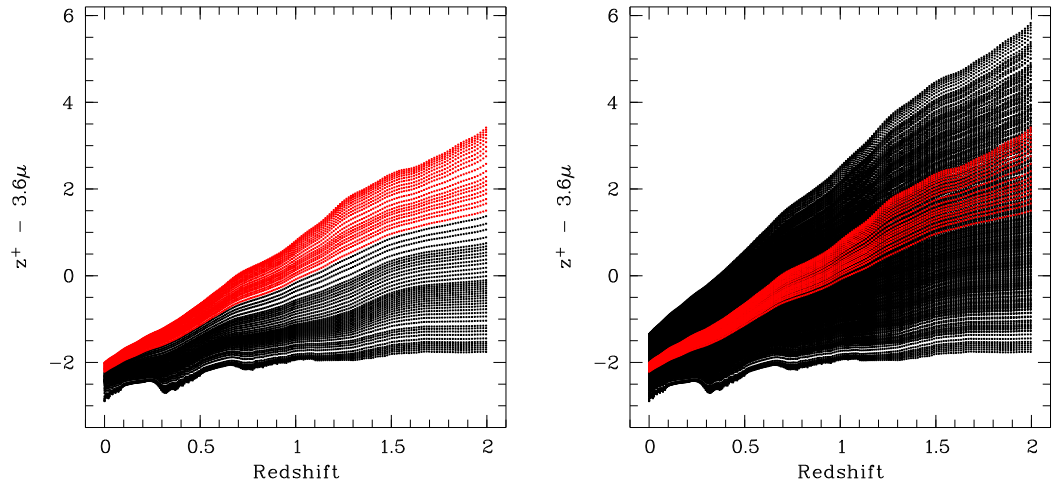

Figure 6.5: $(z^+ - 3.6\mu)$ colors of the library templates at $0 \leq z \leq 2$. Left panel: Dust-free passive templates are plotted in red, dust-free starburst and star forming templates plotted in black. Right panel: Plotted in black are all templates from the galaxy library, with dust extinction ranging from dust-free to very dusty. Overplotted in red are the dust-free passive templates.

be performed in this redshift range to date. These will be analyzed in the future, when deeper near-infrared imaging is available for the COSMOS field.

For those HIROCS clusters that are verified as groups of overdense objects in the $(z^+ - 3.6\mu)$ color-selected search, not all objects are in common with the IRAC-based passive galaxy selection. Figure 6.7 shows a cluster at $z = 1.20$ as an example. On the H-band image in the left panel, the cluster members found by the 3D search are encircled. Additional objects at the cluster position are found using the search based on the $(z^+ - 3.6\mu)$ color selection. They are marked with squares on the 4.5μ image in the right panel, with their photometric redshift estimates given. For almost all of them, the photometric redshifts are not secure, with probability distribution widths of about ± 0.25, mostly because they are too faint in the H-band. Nonetheless, they are bright in the 4.5μ-band and are thus included in the color-selected sample. On the other hand, many $z \geq 1$ cluster members are star forming and thus have blue SEDs, excluding them from the search on the color-selected sample which has been selected to contain only passively evolving galaxies in the first place.

In summary, the color-selected search for overdensities of passively evolving galaxies provides additional evidence for 7 of the 11 HIROCS clusters in the redshift range $1.2 \leq z \leq 1.6$. No overdensity of IRAC-selected objects is found for four of the clusters.

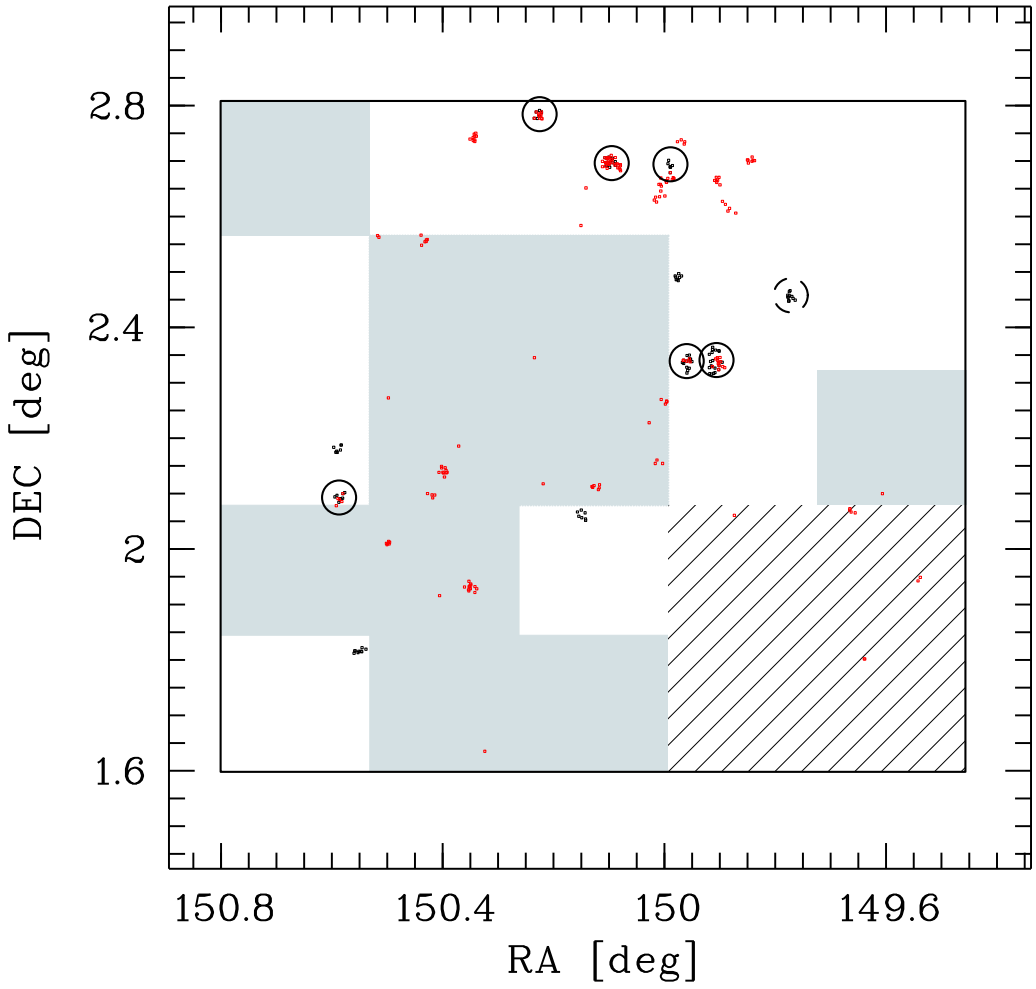

Figure 6.6: Color-selected objects from the 4.5μ-selected catalog having overdensities of 2σ or more in the 2D cluster search (red) overplotted on the HIROCS clusters from the 3D search using photometric redshifts (black). Structures which are in common are encircled. See Figure 6.4 for description.

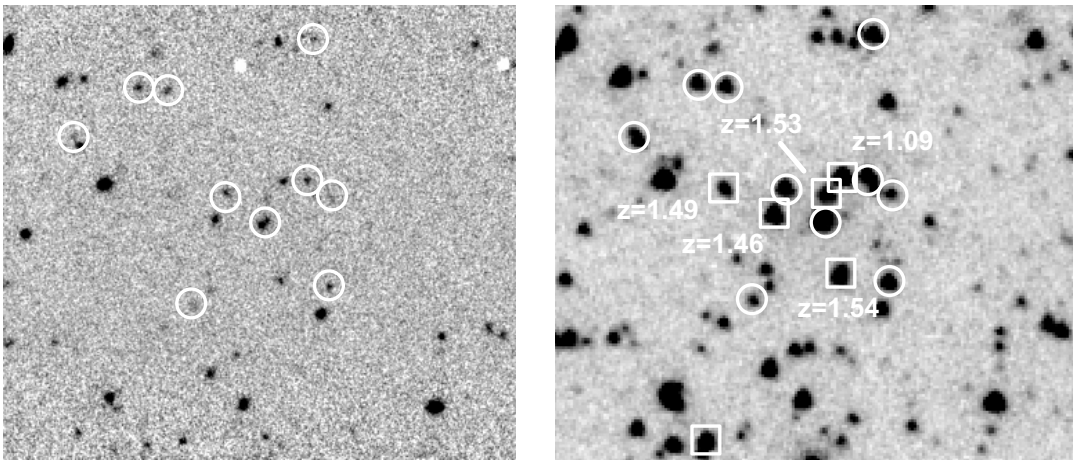

Figure 6.7: Example of additional cluster members being selected by the search based on the $(z^+ - 3.6\mu)$ color selection. Left panel: H-band image of a HIROCS cluster at $z = 1.20$ with the overdense cluster members encircled. Right panel: The same cluster on the 4.5μ-band image. In addition to the encircled members found by the 3D overdensity search, there is a number of red galaxies selected in addition by the search using the $(z^+ - 3.6\mu)$ color selection, marked with squares.

6.4 Results and discussions

6.4.1 Properties of the cluster sample

In total, the HIROCS cluster sample for the COSMOS field comprises 172 galaxy clusters in the redshift range $0.5 \leq z \leq 1.59$, 20 of which stem from the $z > 1$ cluster search on the 0.66 square degrees with deep H-band coverage. The complete list of candidates is given in Appendix A. A differential redshift distribution of the HIROCS candidates is shown in Figure 6.8 (solid line), whereas the red line shows the distribution of HIROCS clusters that can be matched with X-ray detected clusters from (Finoguenov et al. 2007). The histogram of X-ray identifications has a prominent spike because many of the HIROCS clusters matching X-ray detections fall into this redshift bin. The height of the spike looks exaggerated due to low number statistics. The numbers of the $z > 1$ cluster sample are scaled here to match the larger area covered by the $0.5 \leq z < 1$ cluster search. It should be noted that the redshift range of the X-ray matched sample is truncated due to the lack of H-band coverage (see Section 6.1.3.1). Above $z = 1.59$, no clusters are found in the part of the COSMOS field probed here. The depth of the data should be sufficient to trace clusters out to higher redshifts, however this is complicated by the increasingly broad probability distributions of the photometric redshift. Rich clusters with many bright members are easier to detect at high redshifts, but a 0.66 square degree field is unlikely to contain a distant rich cluster due to their low number density, see *e.g.* Figure 5.10. The 7.25 square degree IRAC shallow cluster survey sample (Brodwin et al. 2008) also contains only few clusters above $z = 1.5$, even though they also should in principle be detectable with this dataset. Their search method also

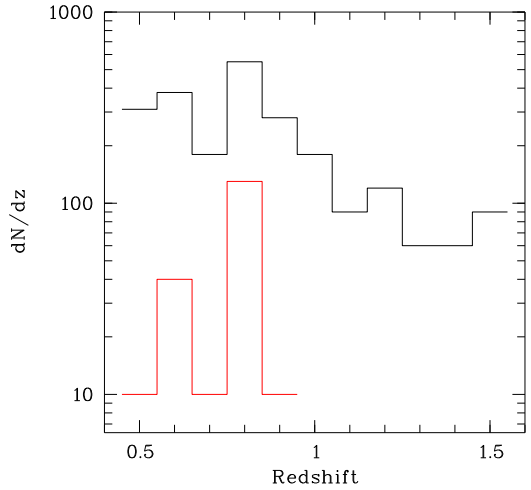

Figure 6.8: Differential redshift distribution (dN/dz) of the complete HIROCS cluster sample for the COSMOS field (black line) and the HIROCS cluster that can be matched with X-ray emission (red line). The values for $z \geq 1$ clusters have been scaled to match the larger area covered by the $0.5 \leq z < 1$ cluster search.

employs photometric redshifts to trace galaxy clusters as galaxy overdensities, but with object selection on the 4.5μ band, which favors passive galaxies. It is not clear at the moment whether this effect is due to photometric redshift inaccuracies, or a real decline in the abundance of galaxy clusters at these redshifts. In case of the IRAC shallow survey, it could also be caused by a selection effect due to a Butcher-Oemler like increase of the blue galaxy fraction in distant clusters.

The nature of the HIROCS candidates without X-ray emission is unclear at the moment. The cluster search on the mock sky shows that the HIROCS algorithm reliably recovers the position and redshifts of massive halos (see Section 5.2.4.3), with false positive rates for massive systems of only $\sim 20\%$ up to $z \approx 1$. This argues that most of the HIROCS candidates in this redshift range trace massive dark matter halos. Above $z = 1$, the rate of false detection rises, but the simulations can not be used for a precise estimate of the exact value due to the near-infrared brightness issues. Of the 11 $z \geq 1.2$ candidates, 7 can be matched with overdensities of galaxies selected in the 4.5μ-band having $(z^+ - 3.6\mu)$ colors consistent with passively evolving galaxies in the same redshift range, providing strong support for their true cluster nature. Without spectroscopic follow-up observations, it is not clear whether the remaining ones are projection effects or less evolved systems whose member counts are dominated by blue, star forming galaxies. More insight, albeit without a definite answer regarding their cluster nature, will be gained after evaluation of the additional near-infrared Y-band data taken with OMEGA2000 during follow-up observations of the most distant HIROCS cluster sample in spring 2008. Unfortunately, they could not be included into this thesis in the remaining time. In addition, public availability of COSMOS medium-band and additional near-infrared data expected in the end of 2008 will boost the accuracy of the photometric redshifts significantly.

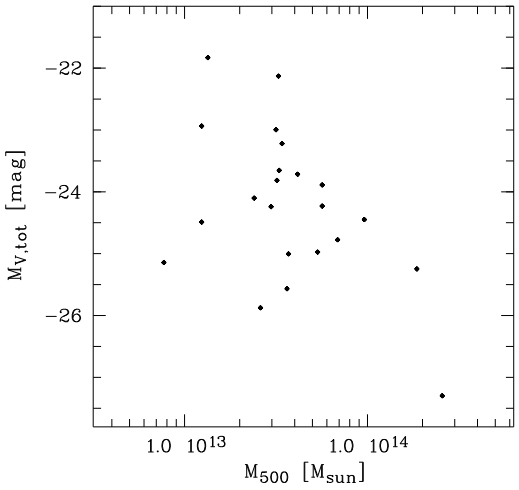

Figure 6.9: Rest-frame $M_{V,\mathrm{tot}}$ plotted against M_{500} inferred from X-ray emission.

6.4.1.1 Mass range probed

Since galaxy number counts and total cluster luminosity are not good tracers of the total cluster mass, the only estimate of the masses probed by the HIROCS sample come from comparisons with masses $M_{500,\mathrm{X}}$ of clusters in common with the X-ray selected sample. These show that the sample includes objects down to $M_{500,\mathrm{X}} = 2.6 \pm 0.38 \cdot 10^{13}\,\mathrm{M}_{\odot}$ at $z = 0.89$ (X-ray cluster #102) and $M_{500,\mathrm{X}} = 1.24 \pm 0.22 \cdot 10^{13}\,\mathrm{M}_{\odot}$ at $z = 0.68$ (X-ray cluster #73).

As already noted in Section 5.2.4.3, the correlation between the cluster total rest-frame luminosity and its mass is not very tight, as is also shown in Figure 6.9 for the HIROCS COSMOS clusters. Due to this, the cluster masses cannot be determined more precisely at this point without *e.g.* velocity dispersion measurements from spectroscopic observations.

According to the newest parameters for the concordance cosmology from WMAP (Dunkley et al. 2008), one cannot expect to find many rich high-redshift clusters in a field of 0.66 square degrees. The estimated number density for $\sigma_8 = 0.8$ are 11 $M \geq 5 \cdot 10^{13}\,\mathrm{M}_{\odot}$ systems per 0.66 square degree (M. Bartelmann, priv. comm.), matching the observations (albeit without completeness or contamination corrections).

6.4.1.2 Radial profiles

Figure 6.10 shows the radial distribution of the overdense cluster galaxies for the 12 clusters with a sufficient number of members, determined from the position of the most overdense object in the cluster. All clusters exhibit profiles with a sharp central peak. In those cases where the structure consists of several subclumps of comparable richness, the radial profile appears non-monotonic. Since the structures with the highest local densities are usually also the most massive ones, they are most likely to feature multiple subclumps / subclusters. Some of the most massive ones, like the LSS at

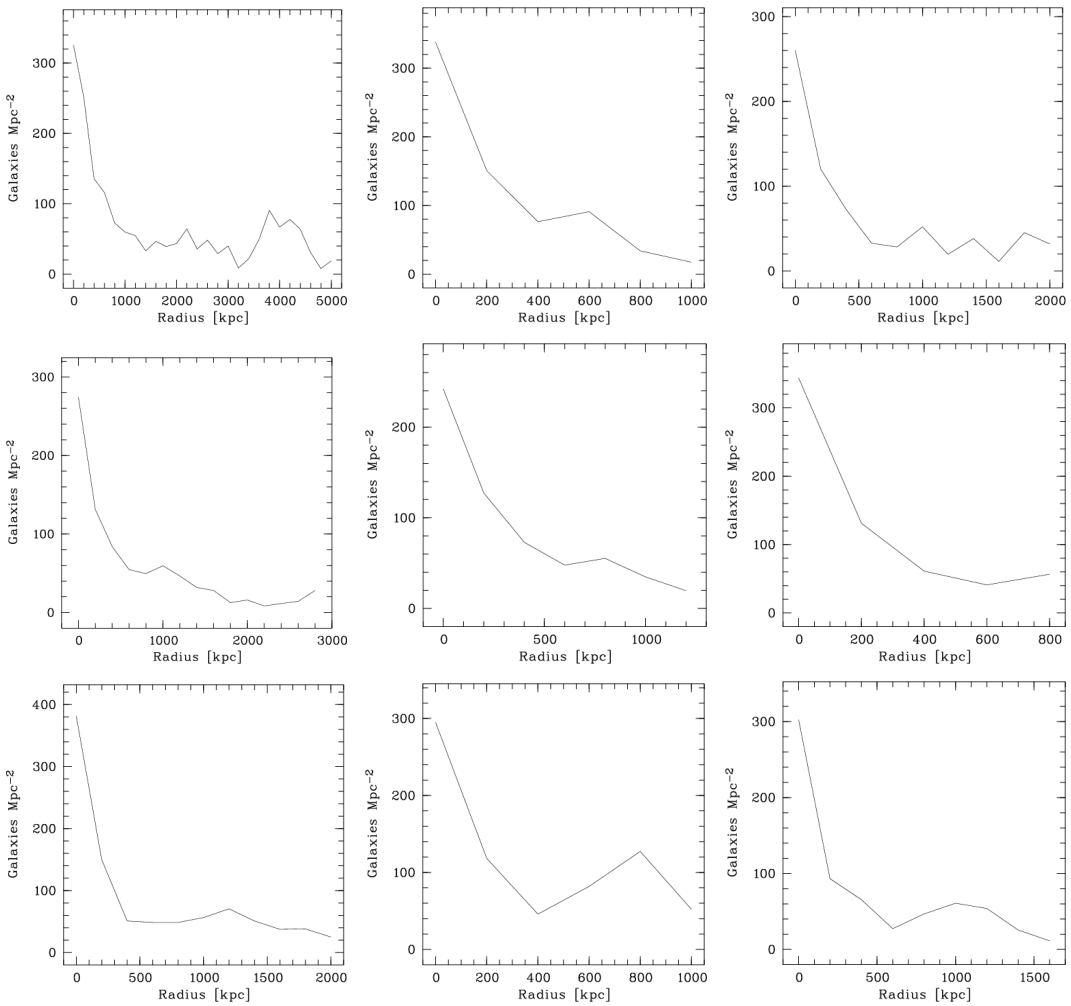

Figure 6.10: Radial distribution of the overdense cluster members determined from the position of the most overdense object in the cluster. The structures are ordered with respect to decreasing maximal local density. No background correction is applied. These profiles show only the most overdense members, so their radii are smaller than those quoted in the literature for typical galaxy clusters.

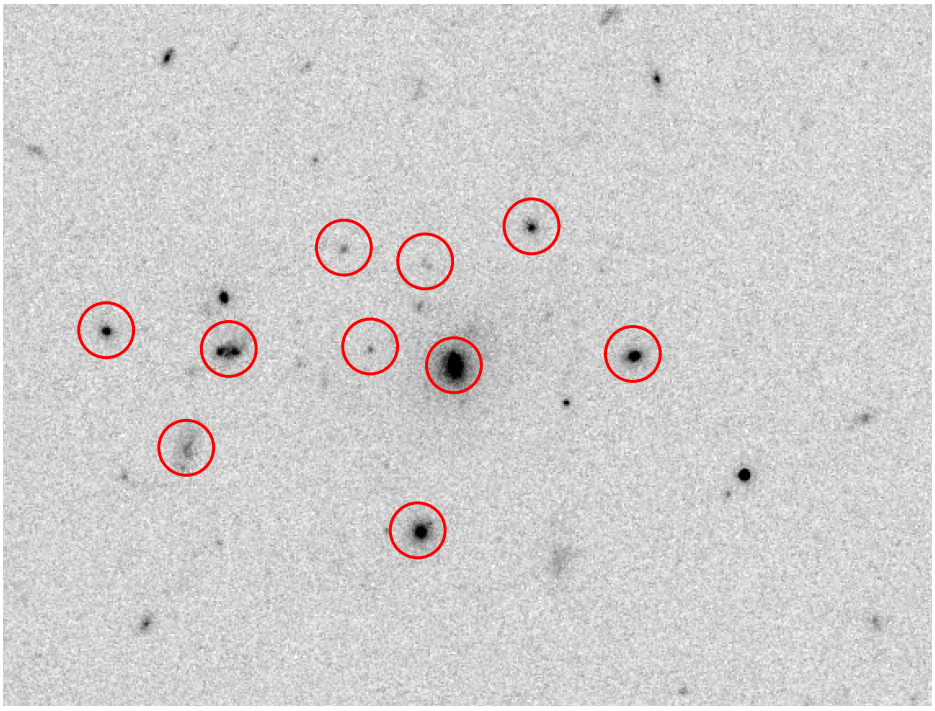

Figure 6.11: ACS image of the core region of a $z = 0.88$ X-ray galaxy cluster with the BCG being a radio source. Cluster members are encircled.

$z = 0.73$ shown in the top left panel of Figure 6.10, even consist of several X-ray bright subclusters. They may eventually collapse to form a single, centrally concentrated, extremely massive cluster. As noted in Section 5.1, the cluster selection program picks them as a single object if they are close together in redshift and projection on the sky, *i.e.* belong to the same LSS.

Some profiles flatten in the outer parts of the cluster. Together with the distribution of the overdense members on the sky, this suggests the presence of an infall region, in the case of the second-from-left panel in the top row of Figure 6.10 even an infalling subclump. Again, spectroscopy would be needed to precisely probe the kinematic substructure of the clusters.

6.4.1.3 Radio sources

A number of the HIROCS clusters contains radio sources. Ten X-ray emitting clusters, or about half of the HIROCS X-ray sample, show a radio source in their core region, close to the center of the X-ray emission. In many cases, the source is one of the central early-type galaxies of the cluster, in some even the brightest cluster galaxy (BCG). Figure 6.11 shows an ACS image of a $z = 0.88$ X-ray cluster core region having a BCG with radio emission. These clusters are possible test cases to study the influence of feedback effects on the ICM as well as the surrounding galaxies.

6.4.2 Color evolution

For comparing the properties of the X-ray detected clusters and those solely found as galaxy overdensities and to study the evolution of these properties with redshift, the cluster sample has to be split up into redshift bins. A tradeoff has to be made between having many redshift bins for detailed analysis of the cluster evolution and the number of clusters per bin which has to be large enough to permit reliable statistics. For this study, the limiting factor is the number of X-ray clusters, and as a result the minimal number of two redshift bins is chosen: A low-redshift bin with $0.5 \leq z < 0.75$, an intermediate-redshift bin with $0.75 \leq z \leq 1$, and a high-redshift bin with $z > 1$. Due to the limited H-band area coverage discussed above, the $z > 1$ sample does not contain X-ray clusters.

A basic characteristic of galaxy clusters and a vital input for a number of galaxy cluster search algorithms is the higher fraction of passive galaxies in clusters compared to the field. This is an incarnation of the morphology-density relation (Dressler 1980) when extending it to very low density environments (*i.e.* the field). In a color-magnitude diagram, these passive galaxies form a well-defined "red sequence" at lower redshifts, in contrast to the "blue cloud" of star forming galaxies. Figure 6.12 shows the fraction of passive members, selected using a rest-frame color-magnitude relation, plotted over the total rest frame V-band magnitude of the $\geq 3\sigma$ overdense cluster members. Galaxies are regarded as being passive if they lie above a line defined by

$$(280 - V) = -0.125 \cdot V - 0.875 \tag{6.2}$$

in the color-magnitude diagram. This division separates the red sequence from the blue cloud and is indicated in Figure 6.13 as a red line. No evolution correction was applied for the separation equation between the low- and intermediate-redshift bin. Figure 6.12 shows the total luminosity of all passive cluster members subtracted from the total cluster luminosity. In both redshift bins, the X-ray clusters populate the upper left area of the plot[1]. This means that they are among the most luminous clusters in the respective redshift bin, and a substantial part of their total luminosity stems from passive galaxies (they are "red", in contrast to "blue" structures whose major luminosity contributors are star forming galaxies). This is especially apparent in the intermediate-redshift bin. In the low-redshift bin, the clusters are more diverse. While the X-ray clusters still are more luminous with a high fraction of their luminosity contributed by passive galaxies, there are also very luminous structures without X-ray emission, both blue and red.

For both cluster populations, an evolution effect is visible when comparing the two diagrams. In the low redshift bin, the clusters, both X-ray dark and bright ones, are redder on average than in the intermediate redshift bin. However, the population is still diverse with about the same overall fraction of clusters being blue.

In an attempt to quantify these effects, X-ray emitting and non-X-ray emitting "average" clusters are compiled in both redshift bins by combining the members of each type in

[1] The case for this in the $0.5 \leq z < 0.75$ redshift bin is not that strong due to the low number of X-ray bright clusters. The effect becomes more prominent when including also the "redshift outlier" clusters that were not included in the final sample, who also fall into the upper left corner of the plot.

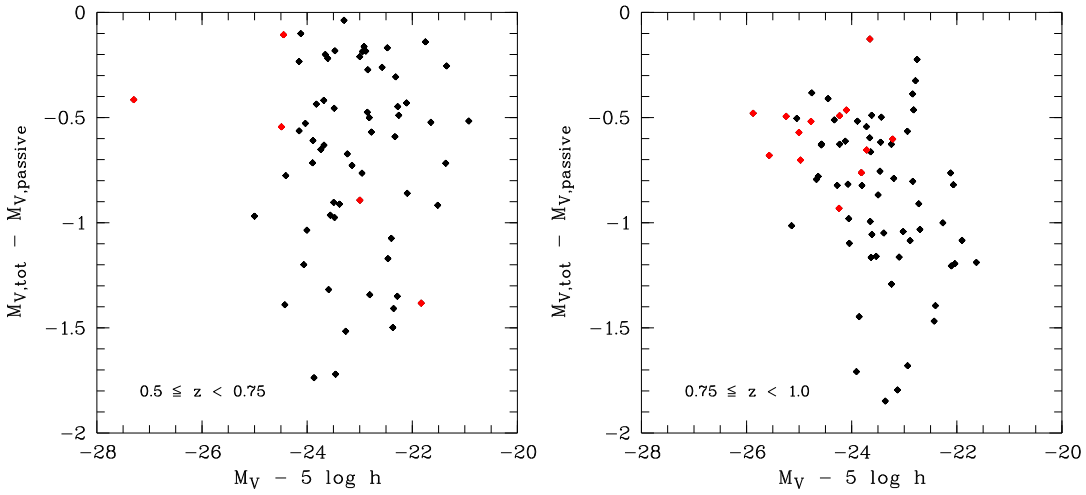

Figure 6.12: Difference in total luminosity of all cluster galaxies and the passive members plotted over the total rest frame V-band luminosity of the $\geq 3\sigma$ overdense cluster members. Left panel: Low-redshift bin. Right panel: Intermediate-redshift bin. Black dots show the clusters from the HIROCS sample without X-ray emission, red dots symbolize the HIROCS clusters which can be identified with an X-ray cluster from the Finoguenov et al. sample.

their rest frame, where their color and magnitude properties are unaltered by apparent luminosity and redshift-dependent K-correction. It is important to recall here again the results from the characterization of the cluster search algorithm using the simulated data. It was shown there that cluster membership of the overdense objects is not secure, with a substantial fraction of them not living in the same dark matter halo as the cluster. The cluster member population is redder and contains more passive galaxies than the field at the same redshift. Because of this, the projection effects adding non-members to the overdense galaxies will more likely increase the number of blue galaxies than add a passive one. As a result, the fraction of the red population in the HIROCS clusters is biased low.

Figures 6.13 and 6.14 show rest frame color-magnitude diagrams of these average clusters in both redshift bins, and Table 6.2 gives the total number of their members as well as the number of passive and star forming galaxies as selected from a rest frame color-magnitude relation. From the low- to the intermediate redshift bin, the fraction of passive galaxies is decreasing in both cluster populations, in agreement with a Butcher-Oemler type effect. In both redshift bins, the fraction of passive members is significantly higher in the X-ray emitting clusters than in the X-ray dark clusters. Their fraction of passive members is also decreasing more strongly, from 36% to 25%, when going out in redshift from $0.5 \leq z < 0.75$ to $0.75 \leq z < 1$. In the X-ray dark clusters, this fraction is going down from 20% in the low-redshift bin to 18% in the intermediate-redshift bin. In both redshift bins, the fraction of passive members in the X-ray emitting clusters is about 1.5 times higher than in the X-ray dark clusters.

The X-ray emitting clusters tend to be the most luminous ones in their redshift inter-

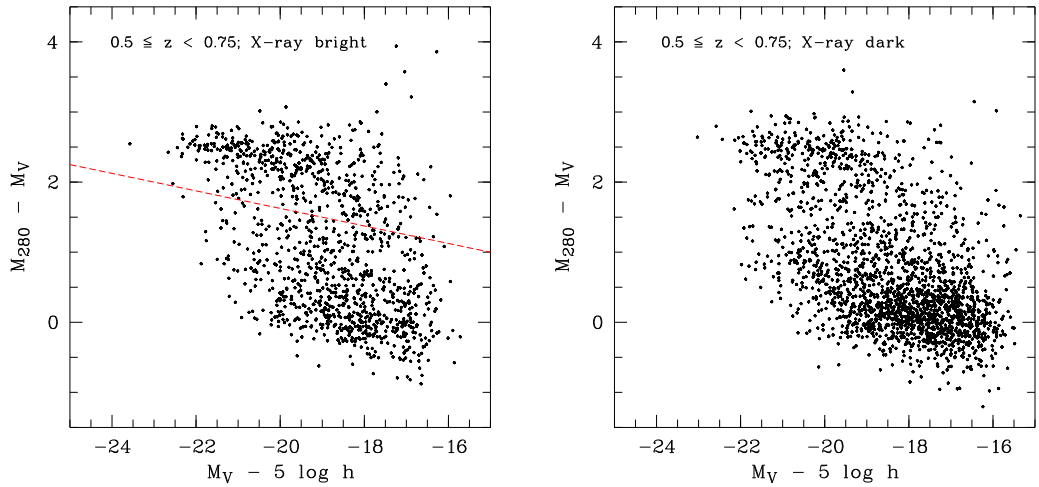

Figure 6.13: Rest frame color-magnitude diagrams of the members of the average clusters in the low-redshift bin. The x-axis gives the rest frame V-band brightness of the galaxies, the y-axis gives the $280 - V$ color. The 280 filter is an artificial box filter centered at 280 nm. Left panel: X-ray bright cluster. Right panel: Cluster with no X-ray emission detected.

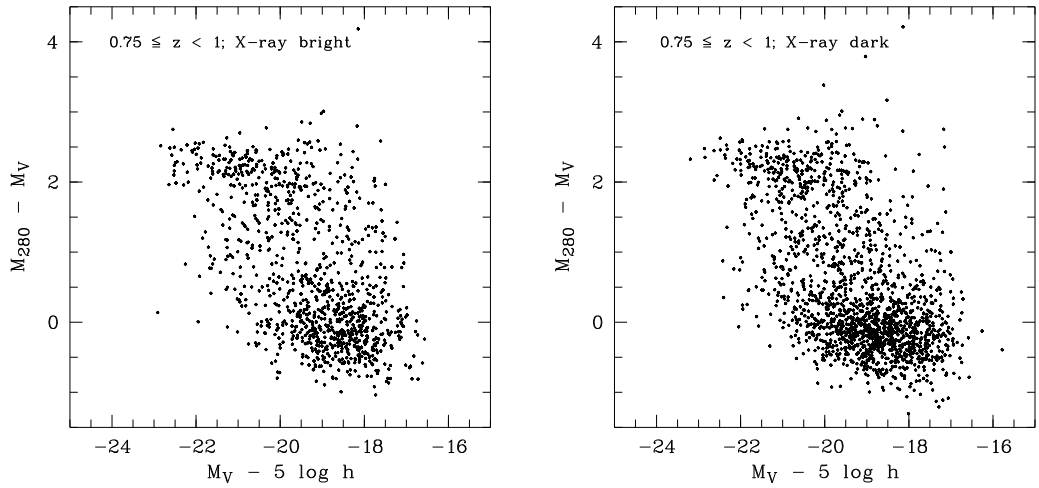

Figure 6.14: Rest frame color-magnitude diagram of the average clusters in the intermediate-redshift bin. The axes are the same as in Figure 6.13. Left panel: X-ray bright cluster. Right panel: Cluster with no X-ray emission detected.

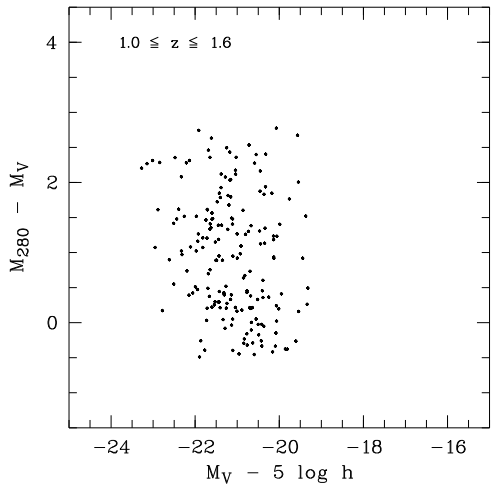

Figure 6.15: Rest frame color-magnitude diagram of all $z > 1$ clusters.

Table 6.2: Member counts of the average clusters in the low, intermediate, and high redshift bins: Total number of members $N_{\rm tot}$, number of passive members $N_{\rm passive}$, and the fraction of passive members $f_{\rm passive}$, with uncertainties indicated. Field galaxies are selected as those objects with overdensities below the peak of the local density Gaussian in this redshift range.

$0.5 \leq z < 0.75$			
	X-ray em.	X-ray dark	field
$N_{\rm tot}$	1217	2228	25608
$N_{\rm passive}$	439	442	3339
$f_{\rm passive}$	0.36 ± 0.02	0.20 ± 0.01	0.13 ± 0.002

$0.75 \leq z < 1$			
	X-ray em.	X-ray dark	field
$N_{\rm tot}$	1089	1814	24189
$N_{\rm passive}$	272	324	2695
$f_{\rm passive}$	0.25 ± 0.02	0.18 ± 0.01	0.11 ± 0.002

$1.0 \leq z \leq 1.6$		
	clusters	field
$N_{\rm tot}$	187	461
$N_{\rm passive}$	38	47
$f_{\rm passive}$	0.20 ± 0.04	0.10 ± 0.02

val, so it needs to be probed whether the fraction of passive galaxies is a function of cluster luminosity. However, the fractions of passive galaxies in the X-ray dark "average" clusters do not change if they are compiled only of members from clusters with total rest frame magnitude $m_V < -23$, the regime of the X-ray emitting clusters (see Figure 6.12). This holds for any magnitude cut applied, thus the HIROCS data do not show a dependency of the fraction of passive galaxies on the total luminosity and hence the mass of the galaxy clusters.

The statistical properties of the passive galaxies in the clusters need to be compared to those of field galaxies as a reference. In the HIROCS sample, the field galaxies are being selected as objects having an overdensity below the peak of the overdensity Gaussian in this redshift range (see Figure 2.6). It should be noted that this is a rather conservative cut to select field galaxies. Their member counts are also given in Table 6.2. In both redshift bins, the fraction of passive galaxies in the field is significantly lower than in both X-ray emitting as well as X-ray dark clusters. The X-ray emitting clusters have more than double the percentage of passive galaxies than the field, the X-ray dark clusters still about 1.5 times of the field percentage. The fraction of passive field galaxies also decreases slightly between the two redshift bins.

Several possible effects could explain this difference. While for galaxy overdensities with X-ray emission the cluster nature is secure, some of the other candidates are projection effects. The comparison with mock sky data in Chapter 5 suggest a false positive fraction of $20\% - 30\%$ in this redshift range. The effect that cluster membership is not secure for a large fraction of overdense objects, also noted there, does not play a role here, since this affects both X-ray dark as well as X-ray bright clusters. On the other hand, as Mulchaey et al. (1996) noted, local spiral-rich groups do not have X-ray emission associated with their intergalactic medium. This would be consistent with the finding that the fraction of blue members is higher in X-ray dark candidates.

Figure 6.15 shows a rest frame color-magnitude diagram of all $z > 1$ cluster members, similar to the "average" cluster color-magnitude diagrams in Figures 6.13 and 6.14.

The fraction of passive galaxies in the $z > 1$ clusters is between those for the X-ray active and X-ray dark $0.75 \leq z < 1$ clusters, but one needs to note that this percentage is relatively uncertain due to a rather small total number of member galaxies. Likewise, the fraction of passive field galaxies is also higher than at lower redshifts, but this number is also suffering from larger statistical uncertainties than its lower-redshift counterparts. Part of the increased fraction of passive members could also be due to the catalog selection. A magnitude cut in an infrared band favors red galaxies, thus decreasing the fraction of blue, star-forming galaxies in the sample.

In contrast to the clusters at lower redshifts, the $z > 1$ clusters also contain very luminous blue galaxies. Whereas in lower-redshift clusters red, passive galaxies dominate the total luminosity, a population of equally bright, but blue galaxies is found in the clusters above $z = 1$. The average total luminosity of the cluster members in the blue cloud increases with increasing redshift, an effect visible even when comparing the HIROCS clusters at $1 \leq z < 1.2$ with those at $z \geq 1.2$. Between these two redshift bins, the average luminosity of a blue cloud galaxy increases from $V = (-20.7 \pm 0.81)^{\mathrm{mag}}$ at $z \approx 1.1$ to $V = (-21.4 \pm 0.63)^{\mathrm{mag}}$ at $z \approx 1.4$. This is not due to the selection effect

that more distant objects need to be brighter intrinsically to fall above the magnitude cut. Instead, the clusters at $z \geq 1.2$ contain a significantly higher number of bright blue cloud galaxies than those at $1 \leq z < 1.2$.

To trace back the evolutionary effects visible in the color-magnitude diagrams to a Butcher-Oemler effect, an extremely careful analysis is required. A number of caveats complicates this task. First, the low-redshift clusters have to have the properties and be in the mass range of descendants of the higher-redshift clusters, since the fraction of blue galaxies depends on cluster properties like halo mass. In addition, the fraction of blue galaxies is also not only a function of the evolutionary state of the cluster, but also of distance from the cluster core (morphology-density relation). Thus, the way in which the aperture for the number counts is defined has to be consistent for the whole sample, and has to depend on the properties of the individual clusters. Furthermore, the definition of a "blue" galaxy has to be consistent for all the sample. Some authors choose constant color cuts, which cause even the oldest galaxies to move into the blue sample at some redshift due to remaining star formation. Other authors include evolutionary effects, and define a galaxy to be "blue" if it has a color bluer than some offset with respect to the reddest cluster galaxies at the respective redshift. This way, the oldest galaxies are defined as "red" in all the redshift bins. Careful subtraction of fore- and background objects is also required, taking into account cosmic variance. Usually, the cluster's outer radii are defined as the distance where object density reaches the value of the field, and the aperture to count objects in is chosen as R_{30}, the radius containing 30% of all cluster objects. In the COSMOS field, many clusters are part of LSS having several overdensities at the same redshift (see Chapter 7 for an example), complicating the definition of the cluster boundaries and R_{30}. In summary, the extensive analysis needed to confirm the evolutionary effects visible in the combined clusters in the different redshift bins to be a Butcher-Oemler effect is beyond the scope of this thesis.

An effect very noticeable when comparing Figures 6.15, 6.14, and 6.13 is that in the highest redshift bin, the clear separation between the red sequence and the blue cloud in the rest-frame color-magnitude diagram begins to disappear. This effect has also been found *e.g.* by Cirasuolo et al. (2007). Furthermore, the cluster and field populations become more similar with respect to the fraction of red galaxies as redshift increases, especially at $z > 1$ (see Table 6.2). The quantization of this effect suffers a bit from the uncertainties of the rest frame colors at high redshifts, because of the redshift estimations becoming more uncertain in addition to the intrinsically less accurate color measurements for fainter objects closer to the detection limit. The observed weakening of the color segregation is in agreement with the studies of Cucciati et al. (2006) and Cooper et al. (2007) (see Section 1.1.2), but one has to note that Cooper et al. see a bimodality in the color-magnitude diagram out to $z \sim 1.5$ both for low- as well as high-density environments (albeit with only a weak correlation of the red fraction with overdensity). When regarding the HIROCS results for distant clusters, one needs to keep in mind here that according to the characterization of the cluster selection using mock sky data, the fraction of contamination for the members of individual clusters is quite high (see Section 5.2.4.3). While this introduces a bias for blue galaxies at $z \lesssim 1.2$ because the red fraction correlates with overdensity at these redshifts, the sample should not be biased too much at higher redshifts according to the findings of Cucciati et al. (2006) and Cooper et al. (2007)

Mean colors of distant cluster galaxies and the epoch of star formation The formation age for the bulk of the stellar mass of cluster galaxies and their mass assembly history is still being debated, see Section 1.1.2 of the introduction. As discussed there, galaxy clusters are ideal laboratories to witness the evolution of early-type galaxies because of the accelerated pace in which environmental effects take place there. Assuming a simple star formation history for a galaxy with a single, short duration star formation episode (single burst, exponential decay), the change in color per time unit increases as the age the galaxy is observed at gets closer to the burst, thus providing better leverage to distinguish between different formation times. With increasing redshift, the mean colors of cluster galaxies thus become increasingly sensitive tracers for the formation age of this galaxy type's stellar content (also shown by the increasing color separation of the different evolutionary tracks in Figure 6.16).

The HIROCS distant cluster sample's colors can be used in an attempt to put constraints on the formation age of the bulk of the stellar content of cluster early-type galaxies. Of the available colors and in the target redshift range, $(z^+ - H)$ is tracing the effects of varying age of the stellar content best. Since Figure 6.15 shows that the distant HIROCS clusters are populated by a significant fraction of very blue galaxies, a color cut has to be introduced to prevent them from biasing the cluster's mean color, thus shifting the resulting formation age to later epochs. A rest frame color cut 1^{mag} bluer than the red sequence cut given in equation 6.2 is adopted here. This excludes the bluest galaxies while at the same time leaving room for some color evolution in the red sequence itself towards higher redshifts. Figure 6.16 shows the mean $(z^+ - H)$ colors of the distant clusters' red galaxies plotted against redshift. The evolutionary tracks of galaxies originating from a single burst of star formation at the redshift indicated, followed by exponential decay ($\tau = 1$ billion years), are overplotted in red. The cluster red galaxies' average colors show a clear preference for high formation redshifts. All clusters favor at least $z_f > 3$, and some even $z_f \geq 5$.

Two things need to be noted here. Firstly, the HIROCS sample has not been selected with any kind of preference for old stellar populations (except for the brightness cut in H used to select the catalog). The photometric redshift color library uses non-evolving templates, thus also introducing no preference. Secondly, according to number fraction, the clusters' population contains more blue than red galaxies, and some clusters actually contain only few red galaxies. So while the stellar content of the red galaxies (the cluster's original population?) in distant clusters is uniformly very old with $z_f \geq 3$, possibly even $z_f \approx 5$, they also contain a significant young stellar population in galaxies which are actively forming new stars. Thus, the $1.2 \leq z \leq 1.6$ clusters and their stellar content as a whole are not "old" and still in very active development, as to be expected due to hierarchical structure formation (see Introduction). This has important consequences for high-redshift cluster finding, as these predominantly blue clusters would very likely be missed by search methods like red sequence finders that require overdensities of red galaxies as cluster tracers.

On needs to keep in mind that according to the results of the cluster search on mock data described in Chapter 5, the HIROCS sample very likely contains some false positives due to projection effects. On the other hand, the findings for the formation redshift of red galaxies based on the HIROCS sample are in overall agreement with those from

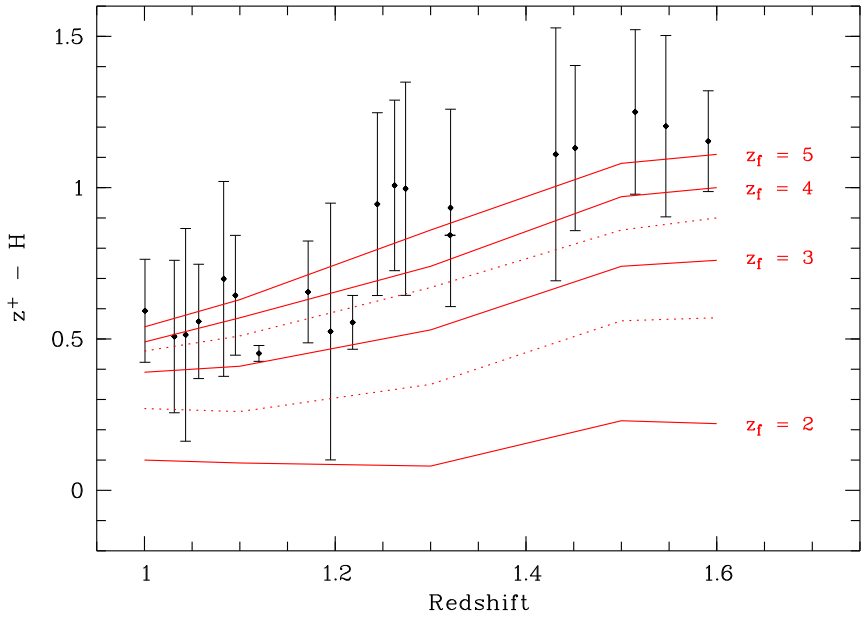

Figure 6.16: Mean $(z^+ - H)$ colors of the distant HIROCS cluster early-type galaxies plotted vs. redshift. Overplotted in red are the evolutionary tracks of galaxies from the template library used for the multi-color classification (see Section 2.2.1.2), originating from a single burst of star formation at the redshift indicated, followed by exponential decay ($\tau = 1$ billion years). Dotted lines correspond to $z_f = 2.5$ and $z_f = 3.5$; errors are 1σ.

other authors, *e.g.* the recent results of Eisenhardt et al. (2008). They find the colors of their clusters at $z \leq 1$ well matched by a "red spike" model (0.1 Gyr starburst) with $z_f = 3$. At $z > 1$, a wider range of z_f is required to fit the data, but the preference for high formation redshifts still remains. One needs to note here that on the one hand, their average cluster colors include all members regardless of their colors, potentially introducing a bias. On the other hand, the cluster sample is 4.5μ-band selected, thus favoring passive galaxies. In addition, the duration of the starburst in their model is significantly shorter than in the one used here, resulting in bluer colors at higher redshifts. Despite all these differences, both findings agree that the stellar content of early type galaxies in clusters formed very early at $z_f \geq 3$, as does *e.g.* the study by Faßbender (2007) who finds the formation redshift of cluster early types to be $z_f = 4.2 \pm 1.1$.

6.4.3 Morphologies of distant cluster members

Figure 6.17 shows representative best-fitting SEDs for cluster members[2] in three redshift slices: $z \sim 1.0$, $z \sim 1.2$, and $z \sim 1.5$. Figure 6.18 gives the corresponding ACS F814W images for the same objects. Here, one needs to note that in the rest frame of the cluster members, the F814W-band roughly corresponds to the u-band, thus being sensitive to signs of star formation. The three columns represent the three major SED types found in the distant clusters: The left column shows the very red, passively evolving galaxies expected to be found in galaxy clusters. These objects fall onto the red sequence in the color-magnitude diagram shown in Figure 6.15. Not surprisingly, the ACS images show bulge dominated morphologies for these objects. The central and the right columns show objects which are neutral and blue beyond the 4000 Å-break, respectively. These objects are members of the blue cloud, with their SEDs indicating recent or ongoing star formation. On the ACS images, many of these galaxies exhibit a clumpy structure, indicative active star formation.

Figure 6.19 gives more examples of blue cluster members with a clumpy substructure. Two major types can be distinguished among them: Galaxies with totally disturbed morphologies (left column) and galaxies with a regular overall shape, but having regions with very active star formation (right column). The galaxies in the central column are in between these two extremes. There is also quite a large fraction of interacting members in the clusters. Here, the term "interacting" is used cautiously, meaning that these members are not only close in the projection on the sky, but also show features typical of galaxy-galaxy interaction, like tidal arms and bridges. Examples of interacting galaxies are shown in Figure 6.20. The majority of the interacting members is neutral or blue beyond the 4000 Å-break (see Figure 6.17), though not as blue as the galaxies with a clumpy substructure. There are also some interactions among passive galaxies in the clusters, where both partners show no signs of star formation. In addition to the galaxies showing clear signs of interaction, there are also galaxies having close neighbors, but without visible direct connection. Since defining comprehensive criteria

[2] As has been shown in Chapter 5, cluster membership cannot be guaranteed for all overdense objects. Thus, the results presented here have to be taken with a grain of salt and need to be verified by spectroscopic follow-up observations.

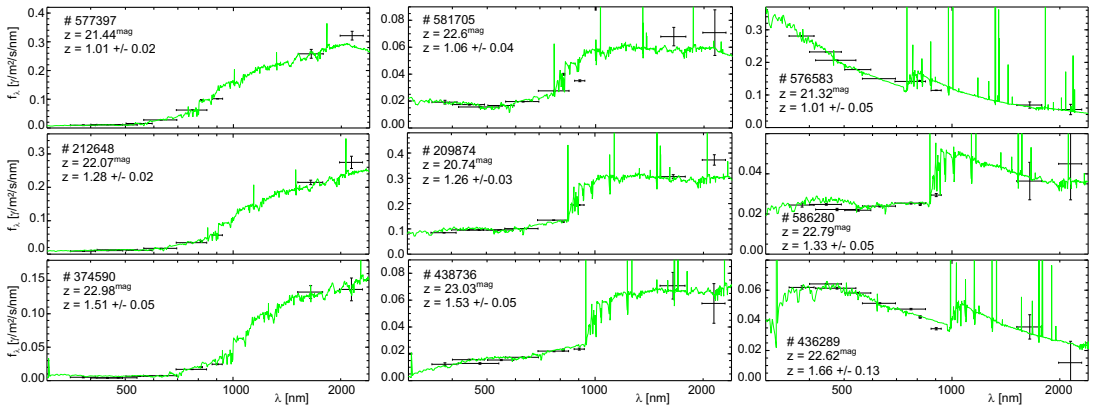

Figure 6.17: Representative best-fitting SEDs for cluster members at $z \sim 1.0$, $z \sim 1.2$, and $z \sim 1.5$. The measurements are indicated with the filter width and the errors. The columns represent the three major SED types found: red, neutral and blue (left to right) beyond the 4000 Å-break.

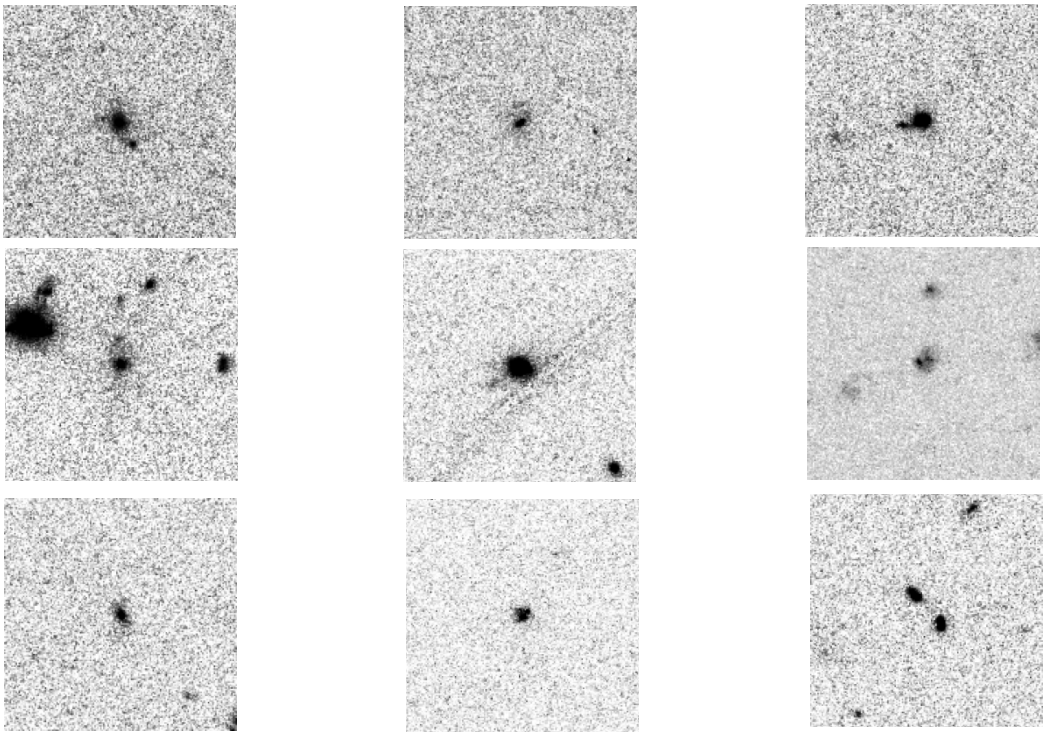

Figure 6.18: Corresponding ACS F814W images of the galaxies from Figure 6.17. The red galaxies show smooth, bulge-dominated morphologies, while many of the blue galaxies exhibit a clumpy structure, pointing to recent or ongoing star formation. Many of the galaxies show close neighbors and signs of interaction.

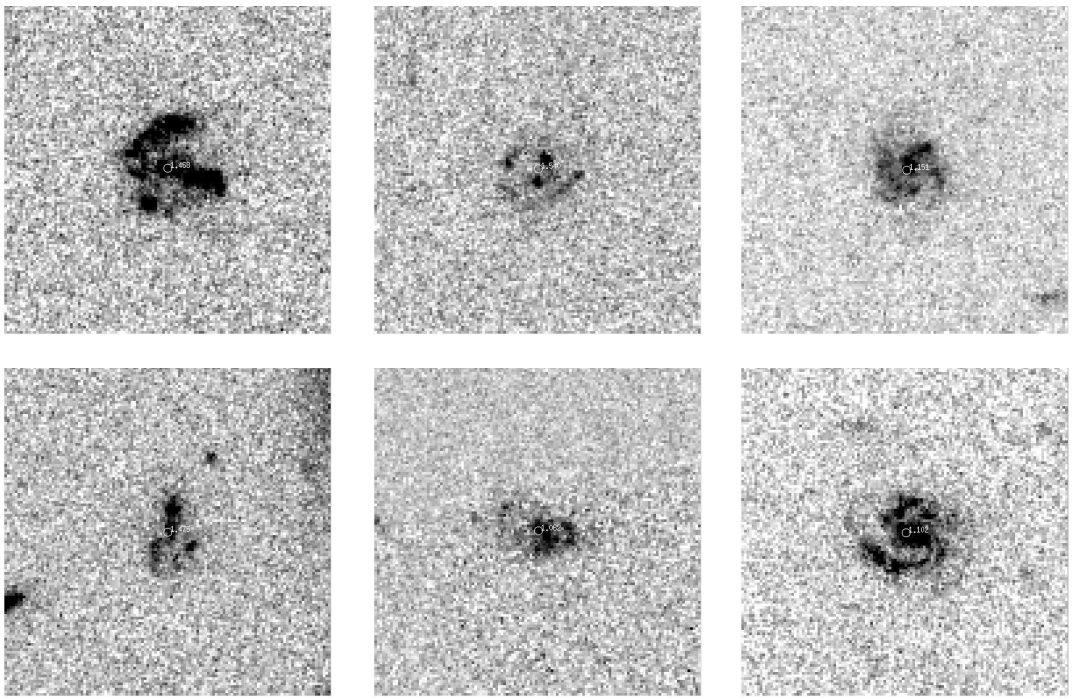

Figure 6.19: Example ACS F814W images of $z \geq 1$ cluster members with a clumpy structure. Two major types can be distinguished here: Galaxies with totally disturbed morphologies (left column), and galaxies with a regular overall shape, but having clumpy regions, most likely indicating very active star formation (right column). The galaxies in the central column are in between these two extremes. The top left panel shows an example of a bright blue-cloud galaxy not found in clusters at lower redshifts. Images are $6.25''$ on a side.

Table 6.3: Number counts of galaxies with a clumpy morphology (left column), interacting galaxy pairs (central column) and galaxies with close neighbors (right column) among the cluster members. The interacting galaxy pairs are so close that they cannot be separated with ground-based imaging, thus each pair is a single object in the master catalog. In contrast to this, close neighbors are counted individually. Galaxies are not contained in more than one group, so even though for example interacting galaxies may exhibit a clumpy substructure, they are only counted as interacting. "Interacting" and "Close neighbors" override "Clumpy morphology" here.

	Clumpy morphology	Interacting pairs	Close neighbors
No. of members	36	30	31

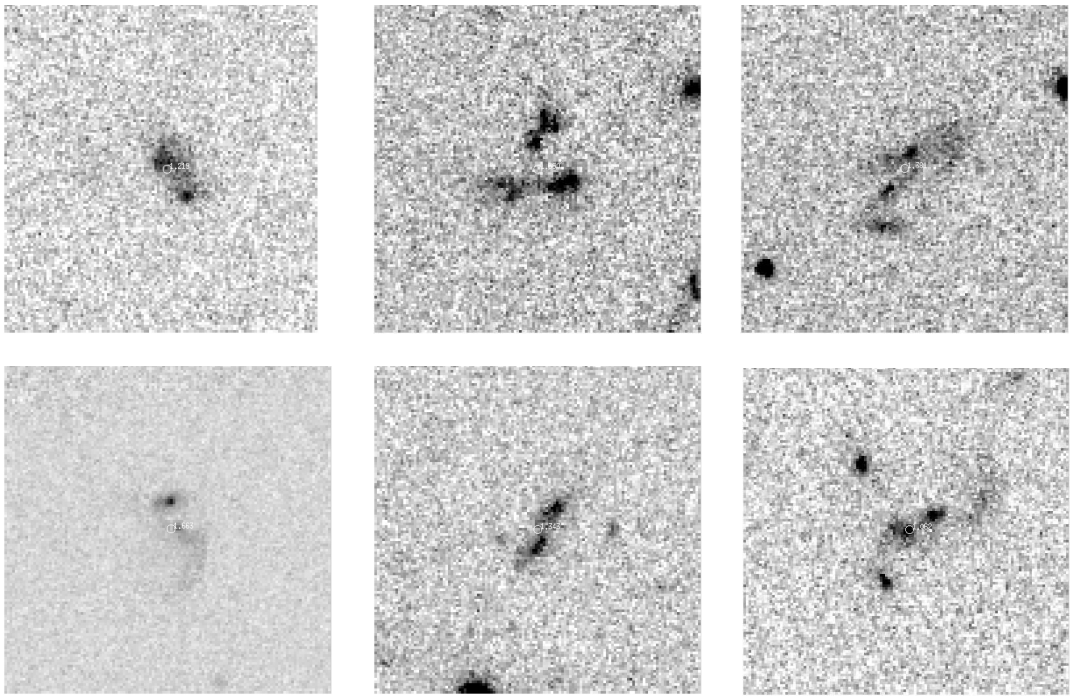

Figure 6.20: Example F814W images of interacting $z \geq 1$ cluster members. The majority of the interacting galaxies are blue cloud members with neutral of blue SED shapes beyond the 4000 Å-break. The bottom right panel shows an example of an interaction of two red-sequence objects. Images are 6.25″ on a side.

to select mergers among these is by no means straightforward, they are treated as a separate group here and are not included in the list of interacting galaxies.

Table 6.3 gives the number counts of the cluster members with clumpy morphology, interacting pairs and galaxies with close neighbors. Out of the 187 $z \geq 1$ cluster members, 97 or about 50% fall into one of these three categories. One needs to note that galaxies are not counted twice here, so even though for example interacting galaxies may exhibit a clumpy substructure, they are only counted as interacting. The properties "Interacting" and "Close neighbors" override "Clumpy morphology" here.

The number of galaxies with a clumpy morphology, pointing to areas of active star formation, does not come as a surprise regarding the fraction of blue cloud members in the clusters. The population of extremely bright blue cloud members which is not found in clusters at lower redshifts is formed by galaxies undergoing very violent episodes of star formation. An example object is shown in the top left panel of Figure 6.19. This galaxy, featuring extremely dominant star forming areas has a total rest frame V-band luminosity of $M_V = -22.78^{\mathrm{mag}}$, comparable to the brightest passive galaxies found in clusters at these redshifts.

The interacting galaxy pairs are so close in the projection on the sky that they cannot be separated with ground-based imaging and are treated as single objects in the catalog, and thus also during photometry. They could thus possibly be classified as peculiar objects, because the measured SEDs actually are a superposition of two, possibly different SEDs, resulting in an overall shape which cannot be fitted well by any library template. For all mergers shown here, the template fit is reliable, so they have to either be mergers of galaxies of similar type, or the total brightness has to be dominated by one of the two galaxies, resulting in an SED not too different from the templates. The interacting pairs account for about 16% of the cluster members (counting each pair as one object, as it is done in the master catalog). This fraction is higher than in present day galaxy clusters, where the large relative velocities caused by the total cluster mass prevent galaxies from merging. Out of the 30 interacting galaxy pairs, 5 have SEDs neutral beyond the 4000 Å-break, and 4 are mergers among two red-sequence objects.

In contrast to the interacting galaxies, each galaxy having a close neighbor is counted individually. They are far enough apart in projection on the sky that they can be separated also by ground-based imaging. In case the companion galaxy is fainter than $H = 21.4^{\mathrm{mag}}$ and thus not included in the catalog, only one galaxy from a close pair is counted as a cluster member. In contrast to the interacting galaxies, for which a clear connection is present, the sample of close pairs may be contaminated by projection effects.

There might be a bias towards interacting galaxies in the member selection introduced by the fact that their brightness is increased compared to regular galaxies, lifting them above the brightness cut or enabling a more reliable photometric redshift estimate. They are only picked as cluster members though if their SED resembles a template library object, because the photo-z of strange objects is too uncertain to be assigned an overdensity by the cluster search algorithm. This in return reduces the number of mergers among the cluster members, since only mergers involving objects of similar types are selected. How these two effects influence the fraction of mergers found in the

clusters compared to their true merger fraction cannot be quantified with the data in hand.

In addition to the formation redshift results presented in Section 6.4.2, and in combination with results about the mass growth of early type galaxies since $z \sim 1.5$ (*e.g.*, Faßbender 2007, but see also de Propris et al. 1999), these findings provide additional support for the hierarchical growth scenario. 16% of the members of distant HIROCS clusters are in direct interaction, and this value has been obtained with a very conservative definition of "interaction". Another 16% of the members are having close neighbors in projection on the sky. While not being the majority of the interactions, at least four clear dry mergers can be identified among the cluster members, consistent in their F814W (rest-frame u-band) morphologies as well as the appearance of the combined SED of the interacting pairs showing no signs of star formation. This provides direct evidence for mass growth of red sequence objects through merging, as predicted in the hierarchical structure growth scenario.

7 Follow-up MOSCA spectroscopy of four $z = 0.85$ candidates

7.1 Motivation

The primary aim of the HIROCS survey is to establish a sample of medium to high-redshift galaxy clusters with as little bias as possible. To achieve this, the detection method is not based on any assumptions about cluster properties, like X-ray emission, the existence of a prominent red sequence of galaxies, or other color properties. Clusters are simply traced by the fact that they are associations of galaxies close together in physical space. One of the main science drivers behind this is to be able to use this unbiased sample to shed light on various aspects of cluster evolution. Fundamental issues in this field which are still lacking observational answers include the question about the formation time of the first X-ray clusters, how the ICM forms, when it thermalizes, and its evolution with respect to redshift. Selection biases between X-ray- and optically selected cluster samples could occur if the evolution of the ICM is not in parallel with the evolution of the cluster galaxies at all redshifts. This is closely connected to the questions why some optically selected clusters are X-ray underluminous (Hicks et al. 2005) or even X-ray dark (Gilbank et al. 2004), and how the other characteristics of these clusters are different from those of the X-ray bright, "standard" clusters. The COSMOS field, with comprehensive data in many different bands including X-ray available, is ideally suited for studying the evolution of the relevant cluster properties.

During the HIROCS $0.5 \leq z < 1$ cluster search on this field, a group of four $z \approx 0.85$ cluster candidates close together in projection on the sky was found. The candidates are assigned the numbers 1-4 in this chapter. Their HIROCS COSMOS cluster catalog IDs are given in Table 7.1. Cluster 1 and 4 are extracted as a single structure by the cluster selection algorithm due to their similar redshift and a "bridge" object between them (see Section 5.1 for a discussion of this). Figure 7.1 shows the overdense cluster members in the projection on the sky.

In the LSS maps of the COSMOS field from (Scoville et al. 2007b), this group of cluster candidates also shows up as prominent peaks. Figure 7.2, taken from this paper, gives the LSS map of the respective redshift. Scoville et al. (2007b) extract the four candidates together as a single LSS at $z \approx 0.9$. There is a minor systematic offset between the HIROCS photo-zs and the ones from COSMOS in this redshift range (see Section 4.7.2), hence the different redshift estimations.

In Figure 7.3, the overdense members of the four cluster candidates are overplotted on the $0.1 - 2.4\,\mathrm{keV}$ X-ray image. One of the clusters, candidate #1, is very bright in X-rays (Finoguenov et al. 2007). With a total luminosity of $L_{0.1-2.4\,\mathrm{keV}}$, it is even

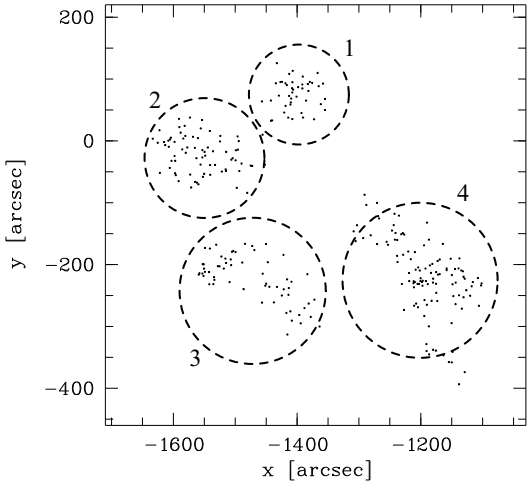

Figure 7.1: Map of the overdense cluster candidate members as projected on the sky.

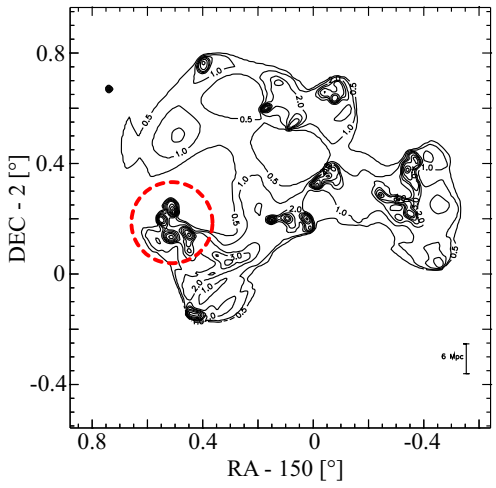

Figure 7.2: Large scale structure map of the COSMOS field at $z = 0.95 \pm 0.05$, taken from (Scoville et al. 2007b). The four cluster candidates are encircled. The difference in redshift to the HIROCS estimates results from systematic offsets between the HIROCS and the COSMOS photo-zs.

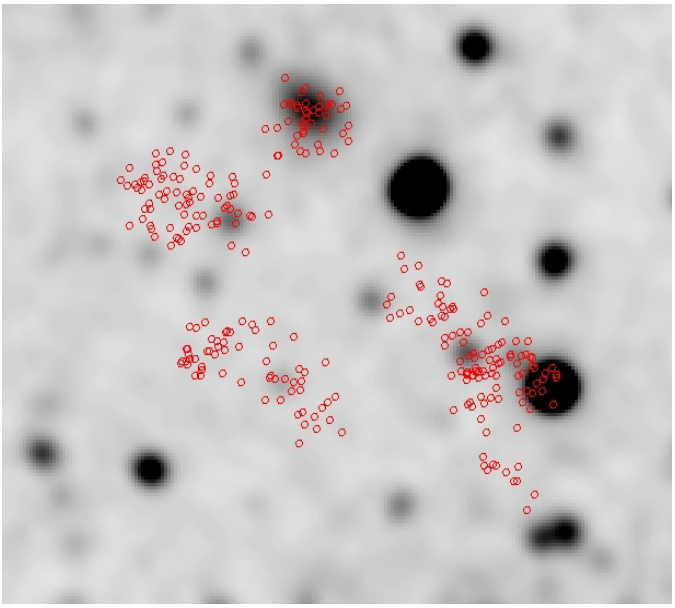

Figure 7.3: $0.5 - 2\,\mathrm{keV}$ X-ray image of the area around the cluster candidates in the COSMOS field. Overplotted as black circles are the overdense cluster candidate members. At the position of candidate #1, extended X-ray emission is clearly visible. No extended X-ray emission is detected for the other three candidates. The X-ray source underlying candidate #4 is pointlike.

the most luminous X-ray cluster in the whole COSMOS field, and its diffuse emission is easily spotted on the COSMOS X-ray sum images. For the other three candidates, no X-ray emission is detected, thus they are not in the list of COSMOS X-ray clusters (Table 1 of Finoguenov et al. 2007). The bright X-ray source underlying candidate #4 in Figure 7.3 is pointlike and thus not due to the cluster's ICM.

Candidates #1 and #2 exhibit a clear red sequence in the CM diagram. Figure 7.4 shows a restframe color-magnitude diagram of the members of candidate #2, with the red sequence clearly visible. The dashed line should just guide the eye and is not taken from galaxy evolution models.

Candidate #2 also sports a very bright passive galaxy close to, but not exactly in the center of its overdense members, similar in shape to a cD galaxy. The most interesting feature of this galaxy is the presence of prominent strong lensing arcs, readily visible in the HST ACS image of the galaxy shown in the left panel of Figure 7.5. The photometric redshift of the galaxy is $z = 0.87$, and the redshift estimation for the brightest lensing arc southwest of the galaxy is $z \approx 1.7$, albeit with a significant error. The redshifts for the galaxy and the brightest arc permit a mass estimate based on fitting a model to the observed data (M. Bartelmann, priv. comm.). The fit is shown in the right panel of Figure 7.5. Here, the position of the lens is fixed on the center of the galaxy, and the best fit is achieved if the source imaged in the two fainter arcs is at $z = 1.2$. The resulting masses, given in Figure 7.6 as a function of lens and source redshifts, indicate an object mass on the scale of a galaxy group with $M \approx 5 \cdot 10^{13}\,\mathrm{M}_\odot$. This mass is within

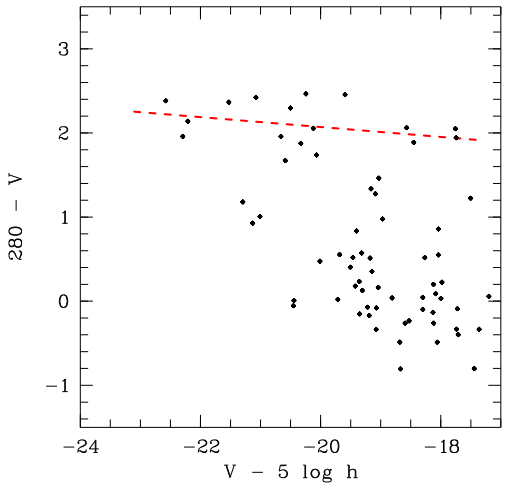

Figure 7.4: Restframe color-magnitude diagram of the overdense members of cluster candidate #2. The red sequence is clearly visible and marked by the red dashed line. Its slope does not result from simulations, but should just guide the eye.

the range of X-ray detected systems at the same redshift and well above the lower limit of $1.8 \cdot 10^{13}\,\mathrm{M}_\odot$ for a $z = 0.95$ cluster as listed in Table 1 of (Finoguenov et al. 2007).

Candidate #1 does not have a single central cD galaxy. Instead, there is a group of very bright elliptical galaxies very close together in its center. On ground-based images they are blurred by the seeing and almost appear as one object, but the ACS data show them to be clearly separated. In addition, a thin structure, possibly a giant arc, is visible (see Figure 7.7). Its photometric redshift is estimated as $z = 1.45 \pm 0.03$.

The two other candidates' red sequences are less prominent, and they exhibit neither X-ray emission nor strong lensing arcs. Candidates #1, #2 and #4 all show concentrations of members at or close to their centers in projection. This also results in high overdensity values for their central member galaxies, whereas the respective values for candidate #3 are lower. The radial profiles of candidates #1 and #2 show the falloff typical for galaxy clusters, whereas the galaxies of candidate #3 seem to be more loosely associated, possibly separated into two subclumps as indicated by its non-monotonic radial profile. The radial profile of candidate #4 exhibits a monotonic falloff, but is more extended than the sharp profiles of candidates #1 and #2. Table 7.1 gives a summary of the properties of the four candidates.

With their very diverse properties, these four cluster candidates form a group worthwhile to study in order to gain insights into why they are so different. Having a spatial extent of just $9'$, they are so close together in projection that they conveniently fit into the FoV of the MOSCA spectrograph of the Calar Alto 3.5 m telescope (Kuhlmann 1997), thus significantly reducing observing time. A proposal (F08-3.5-017) for a detailed spectroscopic study of $R < 24^{\mathrm{mag}}$ members to secure cluster membership for the galaxies, determine the dynamical properties of the candidates, get a handle on star formation rates and eliminate possible projection effects was granted 2.5 nights with MOSCA in January 2008.

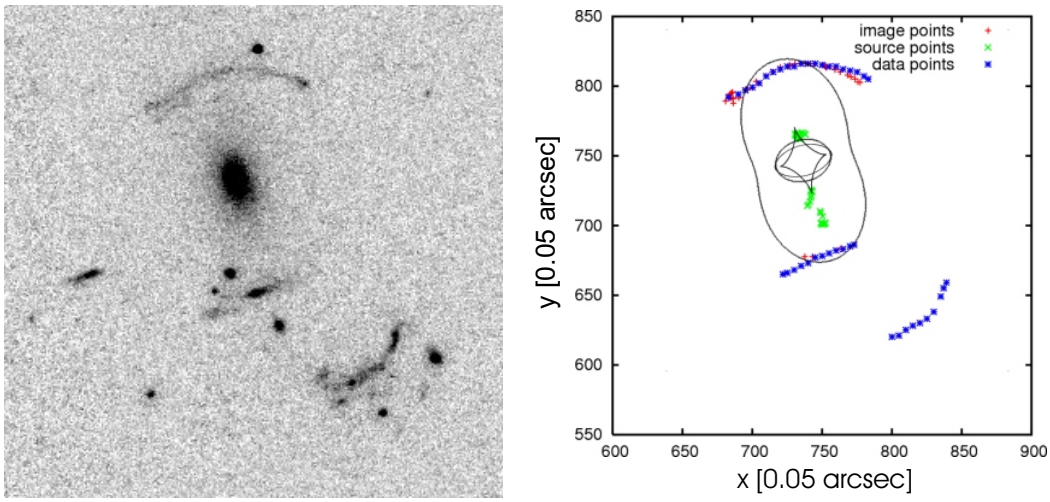

Figure 7.5: Left panel: cD galaxy in candidate #2 showing strong-lensing arcs. The lower right arc has a photometric redshift of $z = 1.7$. This galaxy is close to, but not in, the center of candidate #2. Right panel: Comparison between lensing model (red crosses) and data points (blue stars) for a source redshift of $z = 1.2$ for the upper arc shows good agreement (M. Bartelmann, priv. comm.). A NFW radial mass profile for the lens is assumed.

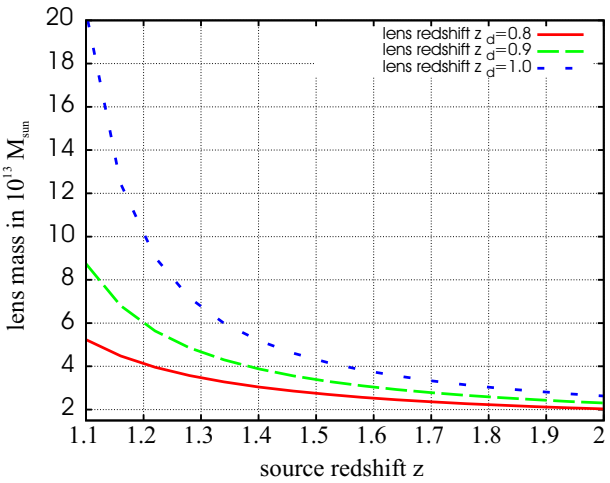

Figure 7.6: Lens mass estimate based on the model from Figure 7.5 for different lens and source redshifts. The photometric redshift of the lens is $z = 0.87$ (somewhat below the long-dashed line), the best fit for the source is $z = 1.2$. This results in a lens mass of $M \approx 5 \cdot 10^{13}\,\mathrm{M}_\odot$ (M. Bartelmann, priv. comm).

Table 7.1: Summary of the properties of the four $z \approx 0.85$ cluster candidates. Row "ID" gives their ID in the HIROCS cluster catalog for the COSMOS field in Appendix A, "$\overline{z}_{\mathrm{phot}}$" gives the estimated cluster redshift, the second row lists the number of overdense members. "X-ray emission" indicates if the cluster is included in the Finoguenov et al. list of X-ray clusters for the COSMOS field. Candidate #2 shows strong lensing arcs; for a structure in candidate #1 it is unclear at the moment whether or not it is an arc. The radial profile of a candidate can be monotonically decreasing ("mon.") or non-monotonic ("non-mon.").

candidate #	1	2	3	4
HIROCS cluster catalog ID	92	92	101	84
$\overline{z}_{\mathrm{phot}}$	0.84	0.84	0.81	0.87
# overdense	43	73	61	112
X-ray emission	$\surd$	-	-	-
Strong lensing	?	$\surd$	-	-
Radial profile	mon.	mon.	non-mon.	mon.

7.2 Observations

7.2.1 Layout of observations

The goal of the observations was to get as complete a coverage of $R < 24^{\mathrm{mag}}$ overdense candidate members as possible, with special emphasis on a good sampling of the dense central regions of the candidates. With four MOSCA masks, object selection is complete down to $R = 23.75^{\mathrm{mag}}$, with some of the $R = 24^{\mathrm{mag}}$ objects included in addition. For the very faint objects, slits were placed on more than one mask to ensure a sufficient SNR. Of the 108 overdense objects with $R < 24^{\mathrm{mag}}$, 82 can be assigned a slit on at least one of the four masks[1]. 10 objects are included twice, and two objects get a slit on three of the masks.

According to the HIROCS team's own exposure time calculator, which had been verified with CADIS and HIROCS observations, $6 \times 2.5\,\mathrm{ksec}$ of integration time per mask are sufficient to obtain a SNR of 7.5 for continuum objects with $R = 23.5$ when using the red_500 grism of MOSCA, assuming a seeing of $1.3''$. This SNR permits precise redshift determination even for objects without emission lines like the central cluster galaxies. Including overhead, a total of 20 h or 2.5 nights was envisaged for the observations.

7.2.2 Data acquired

Due to bad weather and technical problems, only half the scheduled observations could be carried out. For each of the masks 1 and 2, $6 \times 2.5\,\mathrm{ksec}$ exposures were taken on

[1] The number of objects here refer to the preliminary object table that was used to design the masks at the time of observation preparation. For Tables 7.1 and 7.2, the final cluster catalog which is slightly different is used.

January $14^{\text{th}} - 18^{\text{th}}$ 2008. Observing conditions were not exceptionally favorable, with some extinction and seeing in the range of $1.3'' - 1.6''$. The overall quality of the data is not very good; the SNR is much lower than expected. This might also be due to a slight rotation of the mask with respect to the objects on the sky when observations have been taken.

7.3 Data reduction

The data from this run were reduced by Kris Blindert. Care had to be taken of the pincushion distortion which MOSCA spectra are strongly affected by. If not properly corrected, this leads to poor subtraction of night sky emission lines, which are very strong at the wavelengths of interest to this study. The available distortion model (Kuhlmann 1997) is insufficient for this purpose; therefore, arc lamp exposures using a special purpose mask of small regularly spaced holes were obtained. The data from these exposures forms a regular pattern which accurately traces the pincushion distortion of MOSCA. This pattern was fitted to a smooth polynomial function, and used to transform all science exposures. Further data reduction was done via standard IRAF routines, using a hybrid of the multi-slit and longslit packages.

7.4 Results and Discussion

Due to the low SNR of the data, redshift determination for galaxies without emission lines was only possible for a few bright objects. In total, spectroscopic redshifts could be determined for 19 candidate members. Table 7.2 lists them grouped by candidates with their positions, photometric redshift estimates and spectroscopic redshifts given. The actual spectra are shown in Figures B.1-B.5. Unfortunately, the number of redshifts acquired is not sufficient to reach the envisaged goals of this study for any of the candidates.

In candidate #1, known to be a galaxy cluster because of its X-ray emission, two spectroscopic redshifts can be determined to $z_{\text{spec}} = 0.846$ and $z_{\text{spec}} = 0.838$, in excellent agreement with the estimated cluster redshift of $z = 0.84$ based on photo-zs (see Table 7.1; the redshift estimate given by Finoguenov et al. (2007) is $z = 0.90$); Figure B.1 shows their spectra. These two galaxies are separated by $\sim 1300\,\text{km/s}$ in velocity space. This is within the limits for the velocity dispersion of a massive galaxy cluster, especially since the X-ray mass estimate for this object is $M_{500} = 1.86 \cdot 10^{14}\,\text{M}_{\odot}$, the third highest one in the whole COSMOS field regardless of redshift. Figure 7.7 shows the central region of this cluster, with the galaxy at $z_{\text{spec}} = 0.838$ encircled. The cluster center is formed by a group of galaxies of about equal brightness. Two have the morphologies of regular ellipticals, while the NE galaxy has a more extended halo, not unlike a cD galaxy. The right panel of Figure 7.7 has the contrast adjusted to show morphological details in the central galaxies. The galaxy for which a spectroscopic redshift could be obtained exhibits a prominent single tidal arm, and a second weaker tidal feature on its west side. The core region shows a lot of dust extinction and is separated into two clearly distinct parts, both being clumpy and thus showing signs of recent or ongoing

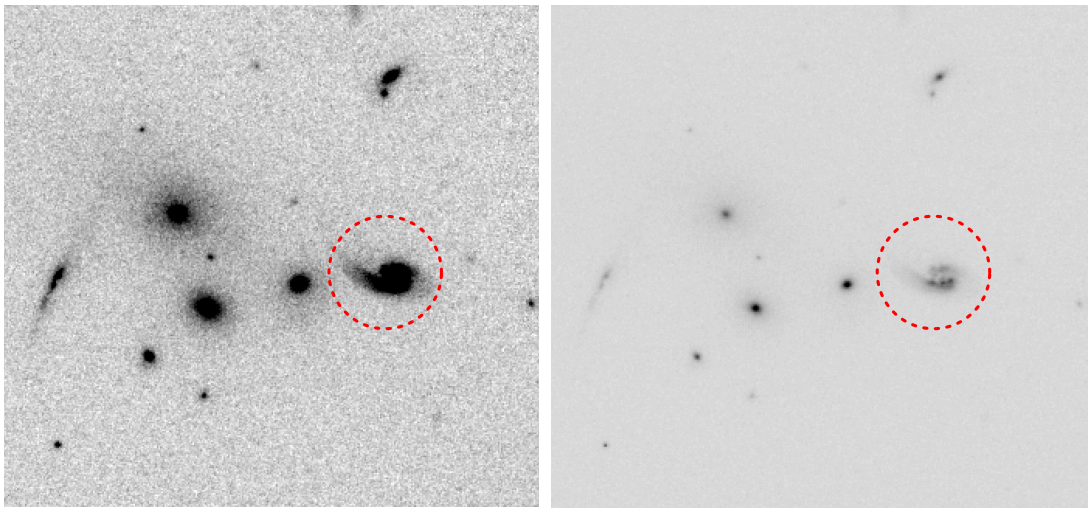

Figure 7.7: Central region of candidate #1. The galaxy with $z_{\mathrm{spec}} = 0.838$ is encircled. Left panel: Maximum contrast. Right panel: Contrast adjusted so that morphological details of the galaxies become visible. The object at the eastern border of the images is possibly a lensing arc of a more distant galaxy with a very blue SED. Its photo-z estimate is $z = 1.45 \pm 0.03$. Both images are $17.25''$ wide.

star formation. The overall shape of this galaxy shows clear sings of interaction and resembles a merger in its final state, similar to the Antennae galaxies. This is also supported by the fact that no "partner" galaxy with tidal features can be found in the vicinity as is to be expected for a recent close encounter. The galaxy's spectrum (right panel of Figure B.1) does show an OII emission line, but this is not very pronounced. In addition, the Balmer absorption lines suggest the galaxy to be past the peak of star formation activity. With just the two redshifts in hand, nothing about the cluster dynamics or the state of the other galaxies in its central region can be inferred. More spectroscopic redshifts will be needed to probe this further.

In candidate #2, three out of five galaxies are at redshifts around $z = 0.82$, again in very good agreement with the value of $z = 0.84$ estimated by the photo-zs of the overdense members. The spectra of the five galaxies for which redshifts could be obtained are plotted in Figure B.2. The maximum velocity separation calculated from the spectroscopic redshifts at $z_{\mathrm{spec}} = 0.817$, 0.820 and 0.821 is $v \approx 700\,\mathrm{km/s}$, consistent with the velocity dispersion of a large galaxy group or a galaxy cluster. This result is in agreement with the mass estimate based on the strong lensing arc for the cD galaxy close to the structure's center. Unfortunately, no spectroscopic redshift could be obtained for this galaxy, and the candidate's interesting substructure with the cD galaxy being off-centered can not be probed further with the data on hand. The three spectroscopic redshifts around $z = 0.82$ strongly point towards this candidate being a real galaxy group or cluster, even though it is lacking detectable X-ray emission. This could be due to for example a different assembly history or evolutionary state of this cluster, influencing the properties of its ICM. Unfortunately, the total number of spectroscopic redshifts

obtained to date is too low to confirm its status. Once confirmed, it would not be the first cluster of its kind. Hicks et al. (2005) have found red-sequence selected clusters of galaxies to be systematically less luminous than similarly rich, X-ray selected ones. In addition, the existence of X-ray dark groups and clusters has already been confirmed by Gilbank et al. (2004), so the assumption of this candidate being a real cluster is not too far-fetched.

For candidates #3 and #4, 8 resp. 3 spectroscopic redshifts could be determined. The according spectra are shown in Figures B.3-B.5. No groupings in redshift are apparent, and the results for these candidates remain inconclusive. They might be projection effects, false detections or sheets of LSS along the direction of view.

Table 7.2: Candidate members for which a redshift determination was possible. The position on the sky is given in columns #1 and #2. Column "z_{phot}" gives the photometric redshift estimate. The spectroscopic redshifts are listed in column "z_{spec}".

Candidate #1

RA [°]	DEC [°]	z_{phot}	z_{spec}
150.50114	2.23481	0.76	0.846
150.50313	2.22447	0.87	0.838

Candidate #2

RA [°]	DEC [°]	z_{phot}	z_{spec}
150.55059	2.21423	0.72	0.838
150.55571	2.21141	0.75	0.821
150.56619	2.21012	0.70	0.740
150.57310	2.20346	0.93	0.820
150.54013	2.20068	0.84	0.817
150.57354	2.19710	0.87	0.924

Candidate #3

RA [°]	DEC [°]	z_{phot}	z_{spec}
150.51300	2.14734	0.81	0.806
150.50726	2.14331	0.95	0.824
150.49423	2.13624	0.96	0.838
150.52079	2.13275	0.97	1.164
150.50215	2.12582	0.86	0.964
150.49766	2.12550	0.82	0.879
150.50450	2.12271	0.83	0.900
150.51031	2.11566	0.93	0.927

Candidate #4

RA [°]	DEC [°]	z_{phot}	z_{spec}
150.43833	2.15651	0.88	0.935
150.44431	2.13905	0.82	0.878
150.42946	2.13843	0.45	0.896

8 Summary, conclusions and outlook

8.1 Goals of the PhD project

The PhD project has been conducted within the MPIA extragalactic key project HIROCS (MANOS-wide). When the author joined the project in late 2005, the project had already been running for about three years, with the first data taken in 2002. Some of the software required for the survey had already been completed, like the program to calculate the local density which is the root for the cluster search, or the reduction pipeline for OMEGA2000 data. Most other software, like the data reduction package MPIAPHOT, was in late stages of development.

Data acquisition had not proceeded that well due to many technical and weather problems in the early stages. While H-band coverage with OMEGA2000 was exceeding 50% for several fields, the first successful observation run with LAICA had only been completed in 2005. With the 04/2006 public release of the first COSMOS data, deep optical data in many bands became available in this field. Near-infrared coverage of the COSMOS field at that time was restricted to the Ks-band, which was very shallow compared to the other bands. In the light of this, the PhD project was focused on the COSMOS field, with the goal to search for distant galaxy clusters, exploiting the time advantage HIROCS had for this over the COSMOS team due to proprietary H-band data. In the later course of the project, additional goals have been added: The development of a cluster selection procedure, and the evaluation of the cluster selection function using a simulation-based mock sky.

8.2 Contributions and results

In preparation for the cluster search in the COSMOS field, the H-band data available in 21 pointings were reduced and their quality assessed. The public optical and near-infrared COSMOS data stemming from ground-based observations as well as the Spitzer space telescope were integrated into the software framework used by HIROCS, and combined with the H-band data. Two object catalogs were extracted, one using wavebands from u^* to H, the other one based on object selection in the IRAC 4.5μ. The type and redshift of the catalog objects were estimated using multi-color classification procedures, in the case of the first object catalog with and without including IRAC 4.5μ-, and 5.8μ-bands in the photometry.

In the course of calibrating the data for the multi-color classification, stellar libraries from Pickles and Decin were evaluated, and their usefulness for the fine-tuning of the

calibration was tested. It was found that the Pickles library matches the PEGASE-generated galaxy templates better in the sense that better agreement between spectroscopic and photometric redshifts of a $0 \leq z \leq 3$ comparison sample can be achieved using the color shifts based on it. The initial hope that the Decin stellar library might be used to determine shifts for colors involving IRAC bands was not fulfilled. The reason for this is not only the mismatch between the Decin library and the galaxy templates which becomes even more pronounced in the IRAC wavelength regime, but a general disagreement between the galaxy templates' shapes and the observations which becomes more pronounced with increasing wavelength. Due to this, photometry for the HIROCS catalog was done without using the IRAC bands. After correcting the photo-zs of the passive galaxies which show a systematic offset, the accuracy of the photometric redshift in comparison with the spectroscopic sample is very good, with a scatter of $\sigma = 0.023$ and an offset of 0.01. There is good agreement between the HIROCS and COSMOS DR1 catalog photo-zs below $z \approx 1$.

Using the properties of the data determined during this analysis, a mock sky based on semi-analytical galaxy evolution models imprinted on the Millennium Simulation was prepared with the goal to resemble the real data as closely as possible. Special care was taken to accurately model the scatter and probability distribution of the photometric redshifts with respect to object brightness and redshift. Except for higher redshifts, where unfortunately an issue with the mock catalog's near-infrared colors causes differences between the mock sky and the real data, the mock sky resembles the real data well. Armed with this, and the cluster selection program based on a friends-of-friends selection on galaxies above the overdensity cut which has been implemented in the course of this thesis, the HIROCS cluster selection function was characterized.

It was found that the HIROCS search method can very reliably recover the positions and redshifts of massive halos. The selection of individual cluster members is less reliable, and there is a significant fraction of systems that suffers from contamination by fore- and background interlopers. As expected, the recovery fraction rises with increasing number of halo members as well as with increasing halo mass. Both of these quantities are correlated, albeit very loosely so that halo mass cannot be inferred reliably neither from the number of cluster members, nor from their total luminosity. As redshift increases, the lower mass threshold of the recovered systems rises due to increasing redshift uncertainties and fewer halo members above the brightness cut used to select the catalog. Above a certain number of halo members or halo mass (the exact values depend on redshift), the fraction of reliable position/redshift detections of halos is nearly constant at $\sim 80\%$ with the exception of the highest redshift bin, whereas the fraction of halos with a reliable identification of member galaxies rises monotonically with the number of members respectively the halo mass.

A parameter study on the effect of the overdensity threshold value on the cluster search result has been performed. It was shown that lowering the threshold leads to a higher fraction of recovered clusters, but going along with a higher fraction of contaminating false positives. The inverse is true for a higher overdensity threshold. Currently, the agreement between simulation-based models and the real data is not sufficient to use this parameter study to determine the optimal overdensity threshold choice for the real clusters search, but it could be shown that the HIROCS 3σ cut is a reasonable choice

for the simulated data.

The results from this chapter can be generalized to any search method based on finding galaxy overdensities using photometric redshifts as distance information (a similar study for red sequence based cluster finding algorithms has been conducted by Cohn et al. (2007)). One result with special implications is that if no additional cluster galaxy properties like color information is employed to select cluster members, the detected systems will suffer from significant contamination by projection effects. This complicates follow-up studies using multi-slit spectroscopy, because a high number of objects has to be targeted to gain a sufficient number of cluster member redshifts, as is also indicated by the spectroscopic follow-up observations of a group of four $z \approx 0.85$ HIROCS cluster candidates using MOSCA (see below). However, using secondary criteria to select cluster members introduces biases, especially at high redshifts where it is not yet clear how *e.g.* the cluster red sequence or the X-ray emission from the thermalized ICM develop. This will be discussed in more detail below.

As the main part of this thesis and fulfilling the major goal, a search for galaxy clusters on the COSMOS field was performed. Because of the lack of uniformly deep H-band data, the search had to be divided into two parts: A cluster search on the whole COSMOS field in the redshift range $0.5 \leq z < 1$, and a search for distant clusters at $z \geq 1$ in the 0.66 square degrees where the H-band reaches at least a depth of $H = 21.3^{\mathrm{mag}}$ (3σ). In total, 172 cluster candidates were found, 20 of which have redshifts $z \geq 1$. These findings were compared with the 72 clusters at $0.08 \leq z \leq 1.25$ identified by Finoguenov et al. (2007) due to their X-ray emission. 29 of their clusters are within the HIROCS redshift range and area covered. Of these, 21 can unambiguously identified with HIROCS candidates. Some of the remaining candidates are visible in the HIROCS data if the overdensity threshold is lowered below 3σ. Some X-ray clusters cannot be detected using the HIROCS search method; however, they are also not very prominent (if at all) in the COSMOS DR1 public catalog. Especially in the light of the findings of the cluster search on the simulation-based mock sky, especially the low contamination fraction by false-positive detections, the comparison of the HIROCS cluster search results with the X-ray clusters is very interesting. HIROCS detects about five times more candidates than there are X-ray detections, and according to the simulation results, most of them should be real. Comparisons of X-ray bright and X-ray dark clusters in the two redshift bins $0.5 \leq z < 0.75$ and $0.75 \leq z < 1$ shows that X-ray clusters are among the most luminous of all structures detected, and the fractions of their light contributed by red sequence galaxies are among the highest of all structures. This is in qualitative agreement with the results of Mulchaey et al. (1996), who find that local spiral-dominated groups show no X-ray emission connected to their ICM.

"Combined" clusters obtained by superposing X-ray bright respectively X-ray dark candidates in three redshift bins ($z \geq 1$ in addition to the two discussed above) show evolutionary effects in the fraction of their members with redshift. Towards higher redshifts, the fraction of red sequence members declines, but is still above the fractions obtained for field galaxies. At $z \geq 1.2$, the clear separation between the red sequence and the blue cloud in the color-magnitude diagram is no longer present. This result is in agreement with the findings of Cirasuolo et al. (2007), who also use photometric red-

shifts based on optical and near-infrared imaging. Perhaps, obtaining deeper imaging in more near-infrared bands would help to increase the contrast between the red sequence and the blue cloud by allowing for better photometric redshifts. Cooper et al. (2007) and Cucciati et al. (2006), both using spectroscopic data, find the red sequence and blue cloud still separated at $z = 1.3$, but without the strong correlation between passively evolving galaxies and overdense environments visible at $z \lesssim 1.1$. Their data show indications for "downsizing", meaning that star formation shifts to less massive systems with increasing cosmic time. The study by Eisenhardt et al. (2008), based on a cluster search on the IRAC shallow survey and tracing galaxy overdensities out to $z \approx 1.7$, adds to this picture. Their high-redshift candidates, all of them very massive, show a population of bright passive galaxies; however, this does not come as a surprise as they select their catalog on the IRAC 4.5μ-band. Putting all the information together, a possible scenario is that only the most massive distant clusters host a significant number of passive galaxies, whereas less massive distant clusters, being in earlier stages of their evolution, still have a high fraction of blue members. This is also shown by the HIROCS candidates, that, while featuring early type galaxies, have the majority of their members living in the blue cloud. Due to the limited field size, they are most likely less massive than the clusters found by Brodwin et al. (2008) on the IRAC shallow survey data. The mass limit below which clusters are "blue" is expected to increase towards higher redshifts. In addition, as the results by Mulchaey et al. (1996) mentioned above indicate, it might well be that the distant clusters with a high fraction of blue galaxies are not yet detectable in the X-rays due to their early evolutionary stage. All of this has important consequences for distant cluster searches. Most likely, cluster search methods trying to exploit this color-overdensity correlation, like red sequence cluster finders, or X-ray based distant cluster searches become more and more biased towards the tip of the cluster mass function with increasing redshifts, compared to color-blind overdensity searches. They might thus miss systems that are in earlier stages of evolution, still in the process of maturing, thus being valuable targets to witness cluster formation. Comparisons of overdensity-selected cluster samples with X-ray cluster searches at high redshifts will provide valuable insights into the formation history of galaxy clusters and the correlation between member galaxy and ICM evolution. Once data acquisition and reduction is complete, the COSMOS field with its large area and deep, panchromatic data will provide a wealth of information for future studies of these important issues.

A first attempt to study X-ray bright and dark HIROCS COSMOS clusters in detail has been undertaken in the context of this thesis. Spectra of a group of four $z \approx 0.85$ candidates, one of which is a cluster with very prominent X-ray emission were obtained using MOSCA at the Calar Alto 3.5 m telescope in early 2008 in multi-object spectroscopy mode. Targeting these faint objects already pushed the instrument to its limits, and despite a major effort by the Calar Alto observatory, observations could unfortunately not be completed due to a mix of bad weather and technical problems. In total, 19 spectroscopic redshifts could be derived. Two of them stem from the X-ray cluster. One of the three structures that is not detected in X-rays features three out of five spectroscopic redshifts within the range of velocity dispersions of large groups or clusters, providing a strong indication of its true cluster nature. The remaining redshifts of the two other candidates show no overdensities; these structures might be projection effects or suffer from contamination of their members by interlopers. It was clear from

the beginning that a spectrograph at a 3.5 m telescope would be pushed to its limit with these observations, and they would only be successful under excellent weather conditions. In the future, similar studies will be attempted using 8 m-class telescopes like the VLT or the LBT.

Another important, yet unsettled question of galaxy evolution is the one about the formation history of elliptical galaxies. As discussed in Section 1.1.2 of the introduction, it is still unclear whether or to which extent they are formed by a monolithic collapse, or by merger events according to the hierarchical structure formation scenario. To settle this question, both their star formation as well as their mass assembly history need to be probed. Samples of distant galaxy clusters provide the best leverage to put constraints on the star formation history. For this thesis, red sequence galaxies in the distant cluster candidates were selected by a rest-frame color-magnitude cut, and their colors averaged for each candidate. These average colors were compared with evolutionary tracks of passive galaxies from the multi-color classification template color library. For all candidates, formation redshifts of $z_f \geq 3$ are favored; for a large fraction even $z_f \geq 5$ provides the best fit. This agrees well with state-of-the-art semi-analytical simulations of galaxy evolution, which find the formation redshifts of early type galaxies to be $z_f \approx 5$ (*e.g.*, De Lucia et al. 2006). Furthermore, the HIROCS results are in very good agreement with the studies of Eisenhardt et al. (2008) and Faßbender (2007).

Current studies of the mass assembly history of early-type galaxies agree less well. Based on the evolution of the K-band luminosity function, de Propris et al. (1999) conclude that early-type galaxies were largely complete at $z \approx 1$, with only passive evolution afterwards. Others, like for example Faßbender (2007), find significant evolution in the brightness of cluster early-type galaxies since $z \approx 1.5$. The source for this evolution would be mergers, as postulated by numerical simulations (*e.g.* De Lucia and Blaizot 2007), who see a mass increase of BCGs by a factor of three between $z = 1$ and today. Support for this scenario, in which the stars in the BCG progenitors are formed at high redshifts, but the mass is assembled late, is found in the numerous interacting pairs found in the distant HIROCS candidates. 16% of their members are interacting pairs, and even four dry mergers could be identified based on their ACS-determined morphology together with the best-fitting template SED. Additional red sequence galaxies might be formed by some of the other interacting pairs if their gas is being removed during the merger, through the processes discussed in Section 1.1.2 of the Introduction.

8.3 Outlook

The redshift range above $z = 1$ in the second quarter of the universe is the time when galaxy clusters evolved most dramatically, both on the cluster scale as well as with respect to their members. This epoch witnesses the buildup of the color-magnitude relation and formation of the cluster red sequence, the thermalization of the ICM together with the onset of X-ray emission, and rapid mass accretion. HIROCS was set up with the goal to compile a statistically significant sample of clusters in this redshift range, together with one at $0.5 \leq z \leq 1$ to do comparison studies. In the course of this thesis and using public COSMOS together with proprietary HIROCS H-band data, the first

larger, uniformly selected sample of distant ($z \geq 1.2$) galaxy clusters in the literature could be published (Zatloukal et al. 2007). HIROCS is also progressing well in the other survey fields; data acquisition has been completed recently. The full 8 square degree multi-band dataset is currently being reduced, and an important milestone has been achieved: A final flux table with object classification for a complete 1 square degree subfield has been obtained, exclusively based on HIROCS observations, and the cluster search there is currently being conducted, with promising first results. In this field, a preliminary flux table based on incomplete data had been available already two years ago, yielding candidates at $z \approx 0.7$ (see Chapter 9 of Falter 2006), $z \approx 1.25$ and $z \approx 1.34$ picked by eye. With the final flux table based on deeper data and standardized clusters search methods in hand, a final cluster catalog based on a well-defined selection method can be obtained now.

In the COSMOS field, follow-up near-infrared Y-band imaging for the most distant cluster candidates has been taken in spring 2008. Unfortunately, it was beyond the scope of this thesis to reduce and integrate these data, but this will be one of the short term goals for the near future. With these deeper data in a filter targeting the rest-frame 4000 Å break it will be possible to constrain the photo$z$s better, probe deeper into the luminosity function of the candidates and select additional members.

The next public COSMOS data release will provide deep near-infrared data in additional bands and possibly also medium-band data in the optical, allowing significantly more precise photometric redshifts. With all the procedures and software to integrate COSMOS data into the MPIAPHOT framework already developed, adding these data to the HIROCS dataset should be relatively straightforward. This will be a major step forward for the cluster identification, and comparison with cluster samples identified using red sequence finders and X-ray emission with the complete XMM Newton data will provide important insights into the mechanisms of cluster formation and evolution.

With the commissioning of LUCIFER at the LBT in 2009, a powerful instrument for near-infrared multi-object spectroscopy on an 8 m-class telescope will become available, with privileged access for the MPIA. This is exactly the equipment needed to perform follow-up observations of the distant cluster candidate samples already available in the COSMOS field, and soon to be compiled for the first square degree of the 3 h field. Its useful wavelength range covers exactly the regime where the characteristic line emission and continuum features of galaxies at the redshifts of the most distant galaxy clusters. This instrument will certainly be a cornerstone of spectroscopic follow-up confirmation to confirm these candidates, assess star formation rates through measuring line emission and trace the total luminous matter of the galaxies using continuum properties.

Outside the HIROCS box, several teams are busy compiling high-redshift cluster samples comparable with the one to be expected from HIROCS. The cluster search on the 7.25 square degree IRAC shallow survey (Brodwin et al. 2008) needs to be mentioned here, that has already published first results based on a sample of 335 cluster candidates, 106 of which are at redshifts greater than one, with 12 spectroscopically confirmed $z > 1$ clusters. There is certainly a number of distant cluster candidates to be expected from the cluster search on the UKIDSS ultra-deep data (van Breukelen et al. 2006), and the various X-ray based cluster surveys (see Section 1.3 of the introduction for a brief list). The instrument of choice for next-generation X-ray cluster surveys is eROSITA

(extended ROentgen Survey with an Imaging Telescope Array Predehl et al. 2006) on board of the Russian satellite "Spectrum-X-Gamma", scheduled to launch in 2011. Its FoV will be about 3 times that of XMM-Newton, and currently an all-sky survey conducted with this instrument is being envisaged, with one of the major goals being to compile a sample of $50\,000 - 100\,000$ X-ray bright clusters.

In the lower-redshift regime, the 1000 square degree image-based red sequence cluster finding RCS-2 (RCS-2), is currently taking data. Together with the Sunyaev-Zel'dovich cluster surveys using the Planck satellite or the South Pole telescope in conjunction with wide-area ground-based imaging to provide redshift estimates, cluster samples of unprecedented richness will become available in the more distant future. The South Pole Telescope has seen first light in early 2007 and is currently preparing to commence survey observations. The area coverage of the South Pole Telescope Survey, once finished, will be 4000 square degrees (Ruhl et al. 2004). Due to their immense cluster sample size, these surveys will be able to put tight constraints on dark energy equation of state and evolution. Putting together results from cluster searches based on the SZE, X-rays, cluster red sequence tracking and galaxy overdensity selection from overlapping survey fields will ensure that the next decade will also be an extremely exciting period for cluster and galaxy evolution studies.

It will still be some years until results based on these data become available. In the meantime, many interesting questions, especially regarding galaxy and cluster evolution in the second quarter of the universe, can be explored using cluster catalogs compiled by surveys like HIROCS. An important beginning step was taken by compiling the first uniformly selected sample of distant galaxy clusters (Zatloukal et al. 2007) published in the literature; the soon to be available cluster catalog for the 3h field will be another important addition. Together with cluster samples from all other surveys listed above, the period of most active cluster development will finally become accessible.

A List of HIROCS cluster candidates in the COSMOS field

Table A.1: List of all $0.42 \leq z \leq 1.59$ HIROCS cluster candidates in the COSMOS field. The table is sorted with respect to z. RA and DEC are given in degrees. Column z gives the mean redshift of the members, σz is the error of the mean. $N_{>3\sigma}$ contains the number of members above the 3σ overdensity cutoff; $M_{V,\mathrm{tot}}$ is their total rest-frame V_J-band luminosity. If a cluster is detected in the X-rays, an "X" is listed in the column "Confirmation". "S" means spectroscopic confirmation of at least 3 members, whereas clusters that coincide with overdensities of objects whose $(z^+ - 3.6\mu)$ colors are consistent with passively evolving galaxies in the same redshift range are marked with "I".

Cluster ID	RA [deg]	DEC [deg]	z	σz	$N_{3\sigma}$	$M_{V,\mathrm{tot}}$	Confirmation
1	150.144	1.565	0.50	0.01	47	-22.96	
2	150.576	2.163	0.51	0.04	9	-21.75	
3	149.758	1.979	0.51	0.02	16	-21.52	
4	150.215	1.825	0.51	0.01	26	-23.00	X
5	150.327	1.504	0.51	0.01	26	-22.78	
6	150.192	1.496	0.52	0.01	25	-22.28	
7	150.361	1.633	0.52	0.04	10	-22.11	
8	150.007	1.606	0.52	0.02	12	-22.40	
9	149.761	1.605	0.52	0.03	12	-22.10	
10	150.293	1.684	0.52	0.01	58	-23.89	
11	150.131	1.857	0.53	0.01	20	-22.37	
12	150.059	1.518	0.53	0.03	13	-21.55	
13	150.048	1.626	0.53	0.01	21	-21.34	
14	149.817	1.815	0.53	0.01	43	-23.48	
15	149.665	1.887	0.55	0.03	10	-20.43	
16	149.491	1.794	0.55	0.01	128	-24.42	
17	149.518	1.878	0.56	0.01	112	-24.15	
18	149.418	1.849	0.56	0.01	26	-22.28	
19	149.434	1.704	0.57	0.01	205	-25.00	
20	149.806	1.580	0.57	0.01	41	-22.82	
21	149.560	1.710	0.58	0.01	64	-23.27	
22	149.625	1.735	0.58	0.01	84	-24.03	
23	149.727	1.745	0.58	0.05	9	-20.92	

Cluster ID	RA [deg]	DEC [deg]	z	σz	$N_{3\sigma}$	$M_{V,\mathrm{tot}}$	Confirmation
24	149.904	1.686	0.59	0.03	10	-21.35	
25	149.680	1.549	0.59	0.01	84	-24.00	
26	149.596	1.517	0.59	0.02	15	-22.46	
27	149.571	1.782	0.59	0.01	47	-23.39	
28	149.522	1.498	0.60	0.01	24	-22.33	
29	150.486	2.755	0.60	0.01	21	-21.83	X
30	149.491	1.633	0.61	0.05	8	-21.42	
31	149.800	1.722	0.62	0.01	57	-23.46	
32	149.767	1.904	0.62	0.03	11	-21.65	
33	150.256	1.974	0.62	0.01	21	-23.00	
34	150.150	2.910	0.62	0.02	18	-22.57	
35	149.608	1.890	0.62	0.01	138	-24.41	
36	149.762	2.027	0.63	0.03	12	-22.52	
37	150.196	2.157	0.64	0.01	34	-23.61	
38	150.277	1.583	0.64	0.03	9	-22.36	
39	150.455	1.893	0.65	0.03	10	-22.32	
40	150.059	2.800	0.65	0.06	8	-21.03	
41	150.639	2.716	0.65	0.03	12	-22.95	
42	150.751	1.524	0.66	0.01	37	-24.45	X
43	149.720	1.845	0.66	0.01	55	-23.50	S
44	150.780	2.805	0.66	0.03	11	-22.26	
45	150.502	2.447	0.66	0.01	26	-22.86	
46	150.710	2.762	0.67	0.05	9	-21.84	
47	149.948	2.098	0.67	0.01	24	-23.24	S
48	150.090	2.195	0.68	0.01	45	-24.15	
49	150.163	2.506	0.68	0.01	83	-24.49	X
50	150.099	2.269	0.68	0.06	7	-21.14	
51	150.054	2.311	0.68	0.01	40	-23.90	
52	150.050	2.452	0.69	0.02	14	-22.10	
53	149.654	2.825	0.69	0.02	14	-23.30	
54	150.272	2.535	0.69	0.02	19	-23.15	
55	150.141	2.067	0.69	0.01	46	-23.59	S
56	149.967	2.662	0.69	0.01	54	-23.56	
57	149.960	2.541	0.70	0.01	991	-27.30	X
58	150.062	2.413	0.70	0.01	25	-22.81	
59	149.843	2.400	0.70	0.01	31	-23.68	
60	149.836	1.681	0.70	0.03	12	-23.65	
61	150.286	2.386	0.70	0.01	60	-24.06	
62	150.027	2.354	0.70	0.03	13	-22.88	
63	149.785	2.465	0.70	0.02	18	-22.85	
64	150.088	2.460	0.70	0.01	31	-23.47	
65	149.899	2.393	0.71	0.01	33	-23.68	
66	150.107	2.719	0.71	0.01	32	-24.12	
67	150.147	2.603	0.71	0.01	35	-23.83	
68	150.242	1.495	0.72	0.05	6	-22.47	

Cluster ID	RA [deg]	DEC [deg]	z	σz	$N_{3\sigma}$	$M_{V,\mathrm{tot}}$	Confirmation
69	150.594	2.127	0.72	0.04	10	-22.92	
70	150.317	2.871	0.72	0.02	16	-23.49	
71	150.745	2.408	0.73	0.01	31	-23.74	
72	149.678	2.264	0.74	0.01	51	-23.87	
73	149.405	2.453	0.74	0.03	10	-21.36	
74	149.552	2.009	0.76	0.01	80	-24.23	
75	149.667	2.372	0.76	0.03	12	-21.90	
76	150.806	2.409	0.77	0.01	43	-23.66	
77	150.403	2.776	0.79	0.01	125	-25.05	
78	150.663	2.818	0.79	0.01	37	-23.63	
79	150.829	2.467	0.80	0.02	15	-22.84	
80	150.214	2.853	0.80	0.03	14	-22.06	
81	149.916	2.509	0.81	0.01	28	-23.24	
82	150.469	2.554	0.81	0.02	26	-22.94	
83	150.299	2.481	0.81	0.01	21	-22.94	
84	150.526	2.139	0.81	0.01	61	-24.57	
85	150.702	2.422	0.82	0.02	16	-21.61	
86	149.718	2.709	0.83	0.02	20	-23.13	
87	150.345	2.843	0.83	0.01	42	-23.46	
88	150.734	2.666	0.83	0.01	32	-23.53	
89	149.623	2.403	0.83	0.01	47	-24.58	
90	149.431	2.190	0.84	0.04	8	-20.77	
91	149.610	2.860	0.84	0.07	6	-22.13	X
92	150.533	2.205	0.84	0.01	117	-25.25	X, S
93	150.065	2.407	0.84	0.05	7	-22.27	
94	150.365	2.004	0.84	0.01	25	-23.81	X
95	150.550	2.590	0.84	0.02	20	-23.39	
96	150.431	1.965	0.85	0.01	36	-23.76	
97	150.514	1.992	0.85	0.02	20	-23.61	
98	150.091	2.527	0.86	0.01	62	-24.45	
99	150.588	2.871	0.86	0.01	22	-23.22	X
100	150.494	2.439	0.86	0.05	6	-21.20	
101	150.448	2.140	0.87	0.01	113	-25.14	
102	150.040	2.552	0.87	0.01	47	-23.72	X
103	150.181	2.588	0.87	0.01	119	-24.97	X
104	150.073	2.636	0.87	0.01	37	-24.12	
105	150.666	2.380	0.87	0.01	20	-23.36	
106	150.212	2.286	0.88	0.01	34	-24.23	X
107	150.506	2.569	0.88	0.03	12	-22.11	
108	150.395	2.707	0.88	0.01	20	-22.72	
109	149.958	2.664	0.88	0.01	21	-22.70	
110	149.434	1.965	0.88	0.01	25	-23.72	
111	150.207	2.812	0.88	0.01	44	-23.86	
112	149.652	2.363	0.88	0.01	93	-24.77	X
113	149.922	2.631	0.88	0.01	37	-24.04	

Cluster ID	RA [deg]	DEC [deg]	z	σz	$N_{3\sigma}$	$M_{V,\mathrm{tot}}$	Confirmation
114	150.268	2.086	0.89	0.04	10	-22.12	
115	150.419	2.515	0.89	0.01	47	-24.24	X
116	150.148	2.197	0.89	0.01	41	-24.28	
117	149.966	2.349	0.89	0.01	212	-25.87	X
118	150.210	2.399	0.89	0.01	32	-24.10	X
119	150.430	1.851	0.89	0.01	80	-25.00	X
120	150.258	1.894	0.89	0.03	10	-22.43	
121	149.401	2.412	0.90	0.01	22	-23.44	
122	150.259	2.899	0.90	0.01	59	-24.67	
123	150.149	2.795	0.90	0.03	15	-22.89	
124	150.094	2.201	0.90	0.01	23	-23.65	X
125	150.320	2.883	0.90	0.01	44	-23.66	
126	149.909	2.702	0.90	0.01	47	-24.64	
127	149.509	2.262	0.90	0.07	6	-22.76	
128	150.025	2.203	0.91	0.01	57	-24.76	
129	149.457	2.591	0.91	0.02	16	-22.84	
130	149.493	2.014	0.91	0.01	57	-24.06	
131	150.230	1.752	0.92	0.02	13	-22.79	
132	149.843	2.574	0.92	0.01	23	-23.50	
133	149.845	2.144	0.92	0.05	7	-22.26	
134	149.660	2.234	0.93	0.01	116	-25.57	X
135	150.114	2.129	0.93	0.02	13	-22.82	
136	150.321	2.280	0.94	0.02	17	-23.62	
137	149.752	2.279	0.94	0.01	30	-23.91	
138	150.750	2.469	0.94	0.01	21	-23.64	
139	149.898	2.335	0.94	0.01	25	-23.65	
140	150.006	2.154	0.94	0.01	35	-24.33	
141	150.161	2.686	0.95	0.05	7	-22.02	
142	150.128	2.716	0.95	0.01	25	-23.20	
143	150.077	2.684	0.96	0.02	22	-23.45	
144	150.081	2.733	0.96	0.04	9	-23.14	
145	150.806	2.202	0.96	0.01	26	-23.89	
146	149.974	1.662	0.97	0.01	24	-23.76	
147	149.428	1.883	0.98	0.04	11	-22.41	
148	149.739	2.351	0.98	0.02	19	-23.80	
149	149.483	2.221	0.98	0.01	20	-23.14	
150	150.807	2.688	0.99	0.02	18	-23.02	
151	149.973	1.560	0.99	0.04	9	-22.04	
152	149.949	1.507	1.00	0.01	31	-24.07	
153	150.121	2.002	1.00	0.05	11	-23.92	
154	149.762	2.283	1.03	0.06	8	-23.46	
155	150.685	2.285	1.04	0.06	8	-23.40	
156	150.703	2.347	1.06	0.01	16	-23.50	
157	150.568	2.497	1.08	0.04	11	-24.03	
158	150.080	2.042	1.10	0.06	7	-22.92	

Cluster ID	RA [deg]	DEC [deg]	z	σz	$N_{3\sigma}$	$M_{V,\mathrm{tot}}$	Confirmation
159	150.439	2.751	1.12	0.08	6	-23.45	
160	150.606	1.751	1.17	0.06	7	-22.85	
161	150.585	2.094	1.20	0.06	8	-23.38	I
162	150.148	2.061	1.22	0.07	7	-23.48	
163	150.551	1.816	1.24	0.05	9	-24.11	
164	150.227	2.784	1.26	0.07	7	-23.84	I
165	150.099	2.699	1.27	0.03	15	-24.20	I
166	150.589	2.180	1.32	0.07	7	-23.88	
167	149.990	2.691	1.32	0.07	7	-23.83	I
168	149.958	2.336	1.43	0.02	14	-24.61	I
169	149.976	2.490	1.45	0.06	8	-24.18	
170	149.911	2.327	1.51	0.04	13	-24.34	I
171	149.910	2.358	1.55	0.07	7	-24.21	
172	149.773	2.455	1.59	0.05	11	-24.13	(I)

B Spectra of the four $z \approx 0.85$ cluster candidates

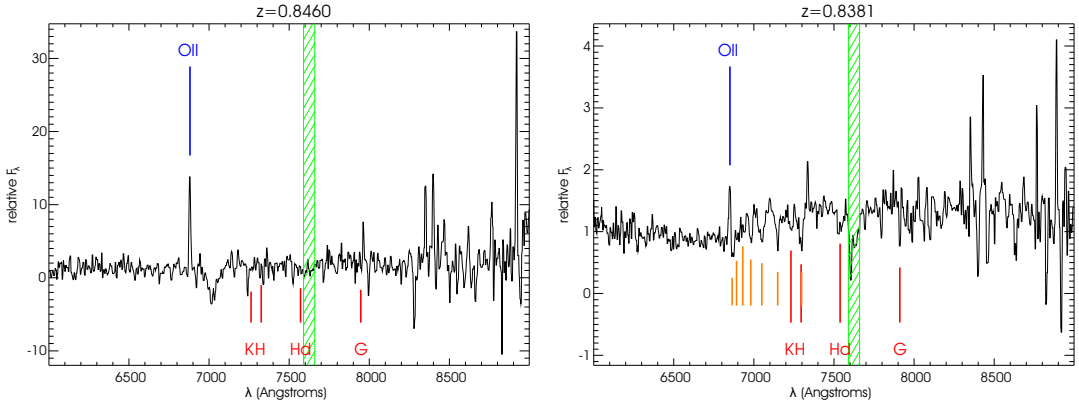

Figure B.1: Spectra for members of candidate #1. The positions of emission / absorption lines is indicated in light blue / red, respectively, with their labels given. The green bands show the position of strong night sky emission lines that cannot be fully removed from the spectra. In the right panel, additional Balmer absorption features detected in this galaxy are marked in orange.

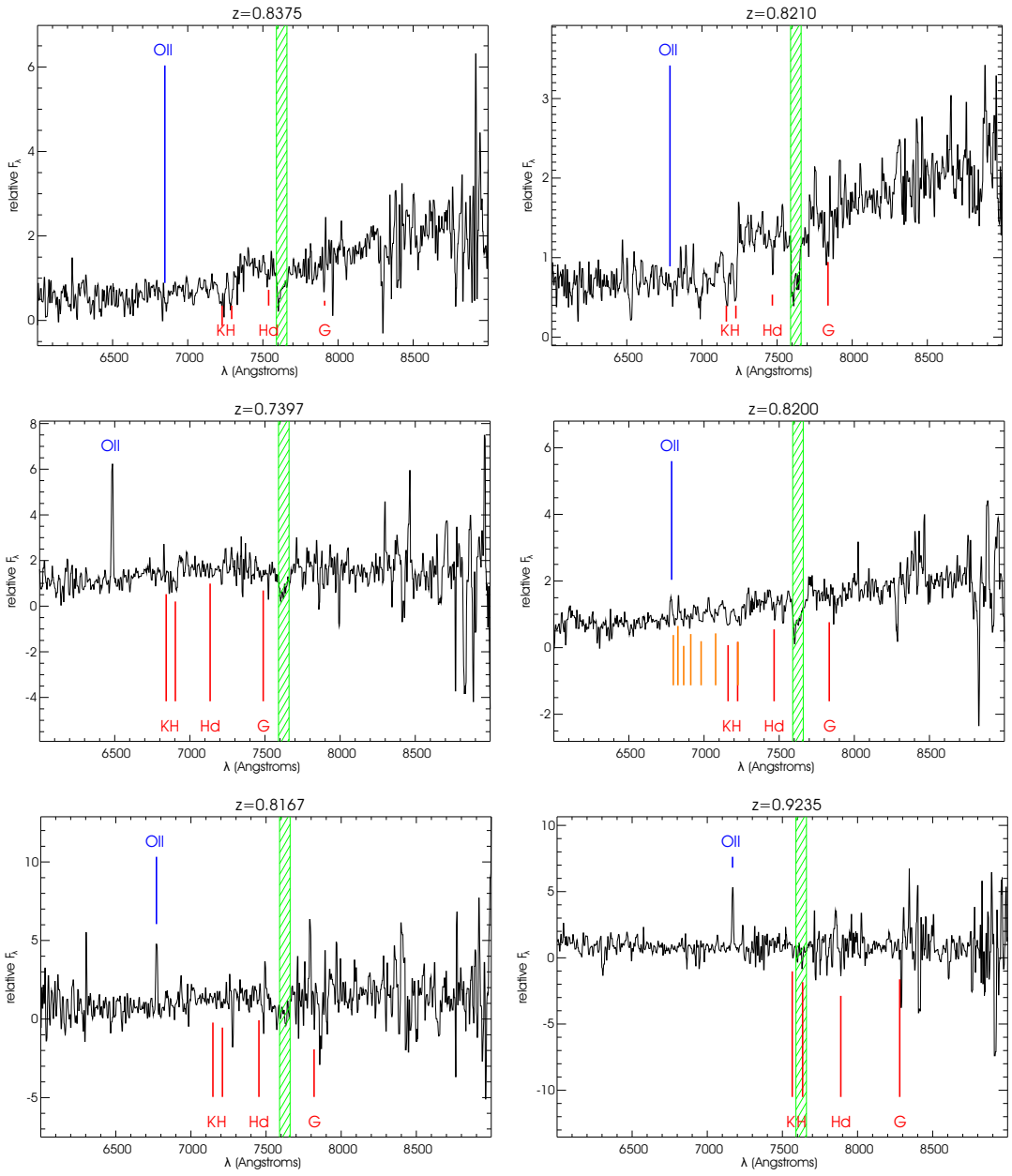

Figure B.2: Spectra for members of candidate #2. See Figure B.1 for explanation of the colors and labels.

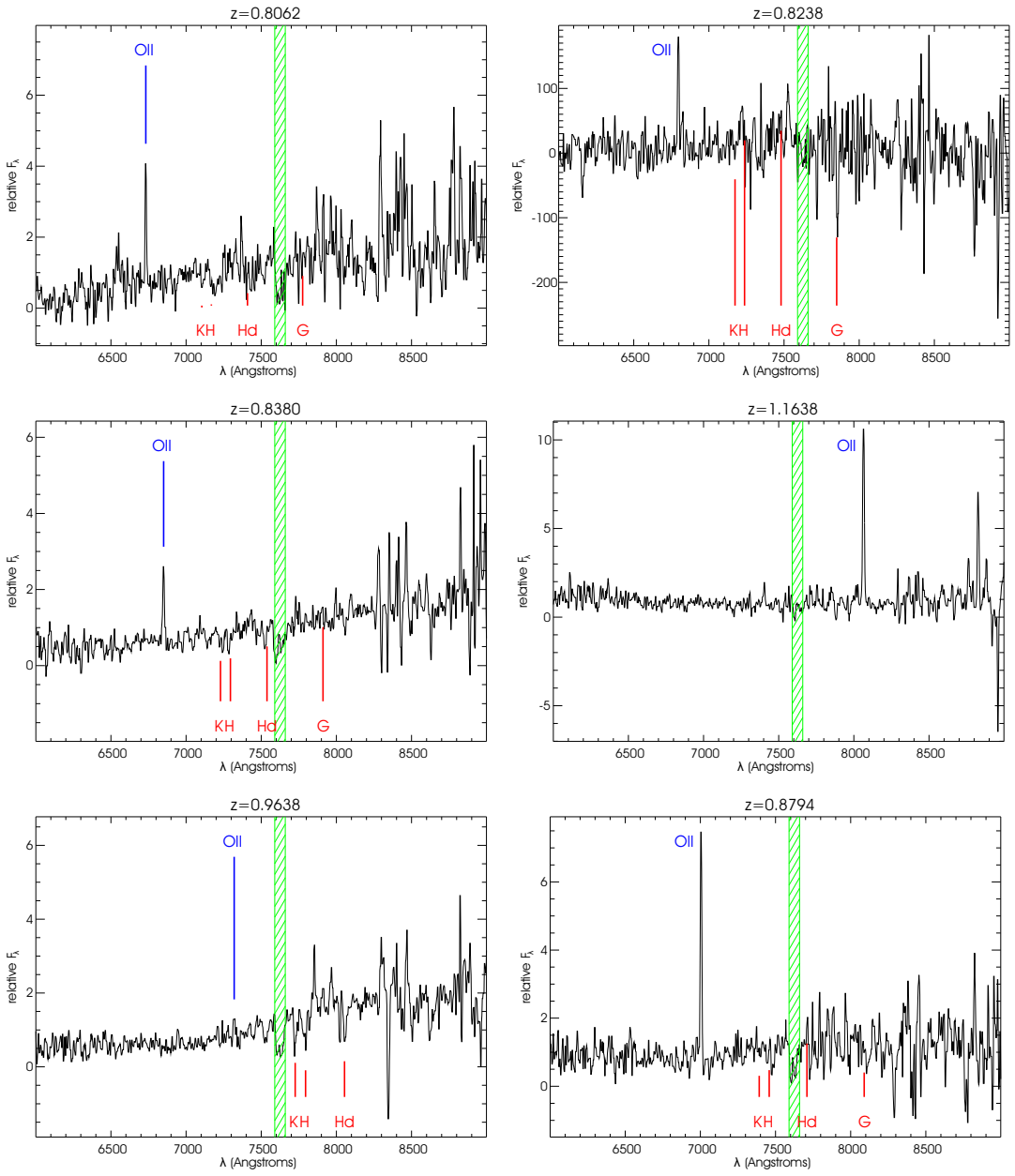

Figure B.3: Spectra for members of candidate #3. See Figure B.1 for explanation of the colors and labels.

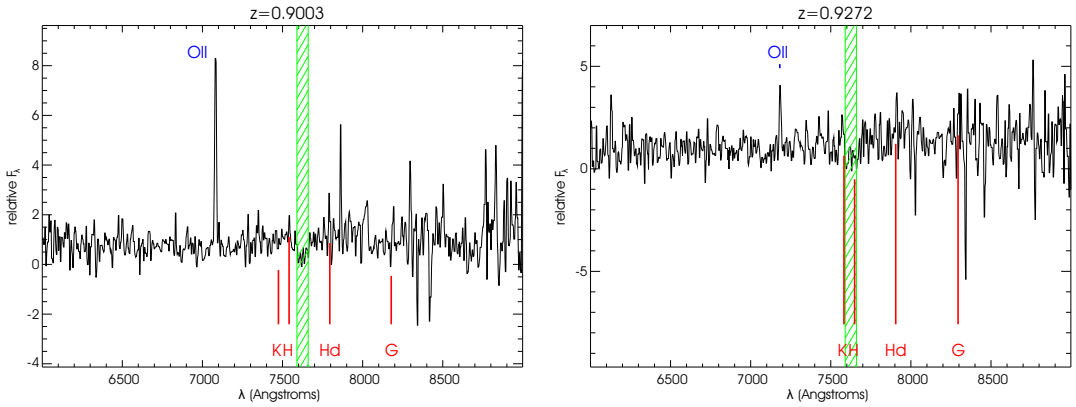

Figure B.4: Spectra for members of candidate #3, continued. See Figure B.1 for explanation of the colors and labels.

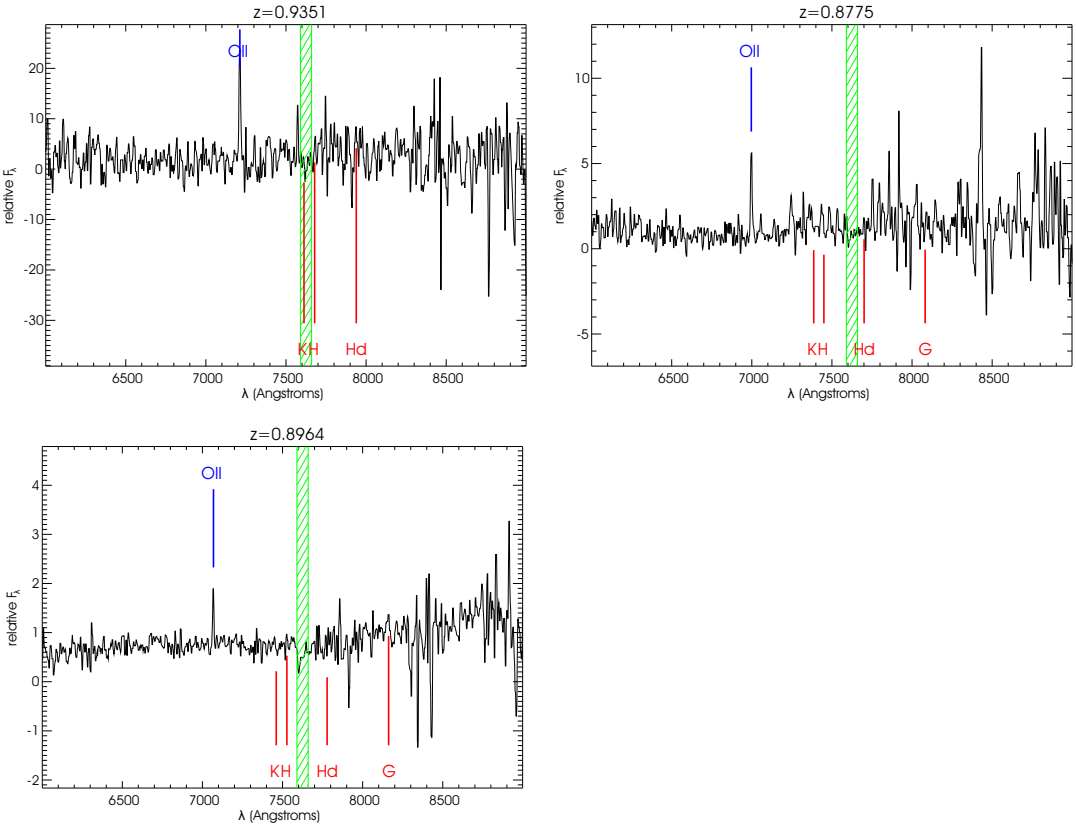

Figure B.5: Spectra for members of candidate #4. See Figure B.1 for explanation of the colors and labels.

C SExtractor configuration

C.1 Object detection on the optical COSMOS mosaics

```
DETECT_TYPE CCD
DETECT_MINAREA 29
DETECT_THRESH 5.0
ANALYSIS_THRESH 5.0
FILTER N
DEBLEND_NTHRESH 64
#DEBLEND_MINCONT 0.001
DEBLEND_MINCONT 0.005
CLEAN Y
CLEAN_PARAM 1.0
MASK_TYPE CORRECT
PHOT_APERTURES 10
PHOT_AUTOPARAMS 2.5,3.5
SATUR_LEVEL 600100000
MAG_ZEROPOINT 27.92
MAG_GAMMA 4.0
GAIN 1.0
PIXEL_SCALE 0.150
SEEING_FWHM 0.95
STARNNW_NAME $PM/../tbl/sex_default.nnw
BACK_SIZE 256
BACK_FILTERSIZE 3
BACKPHOTO_TYPE GLOBAL
MEMORY_OBJSTACK 5000
MEMORY_PIXSTACK 200000
MEMORY_BUFSIZE 1024
THRESH_TYPE RELATIVE
```

C.2 Object detection on the IRAC 4.5μ mosaic

```
DETECT_TYPE CCD
DETECT_MINAREA 3
DETECT_THRESH 5.0
ANALYSIS_THRESH 5.0
FILTER N
DEBLEND_NTHRESH 64
DEBLEND_MINCONT 0.005
CLEAN Y
CLEAN_PARAM 1.0
MASK_TYPE CORRECT
PHOT_APERTURES 10
PHOT_AUTOPARAMS 2.5,3.5
SATUR_LEVEL 10000000
MAG_ZEROPOINT 27.92
MAG_GAMMA 4.0
GAIN 1.0
PIXEL_SCALE 0.150
```

```
SEEING_FWHM 1.75
STARNNW_NAME $PM/../tbl/sex_default.nnw
BACK_SIZE 256
BACK_FILTERSIZE 3
BACKPHOTO_TYPE GLOBAL
WEIGHT_IMAGE wght_ch2_sci.fits
WEIGHT_TYPE MAP_VAR
CHECKIMAGE_TYPE NONE
MEMORY_OBJSTACK 5000
MEMORY_PIXSTACK 200000
MEMORY_BUFSIZE 1024
VERBOSE_TYPE NORMAL
THRESH_TYPE RELATIVE
```

C.3 Object detection on the H-band mosaic

```
DETECT_TYPE CCD
DETECT_MINAREA 3
DETECT_THRESH 2.838
ANALYSIS_THRESH 2.838
FILTER Y
FILTER_NAME $PM/../tbl/sex_psf10.conv
DEBLEND_NTHRESH 64
DEBLEND_MINCONT 0.001
CLEAN Y
CLEAN_PARAM 1.0
MASK_TYPE CORRECT
PHOT_APERTURES 10
PHOT_AUTOPARAMS 2.5,3.5
SATUR_LEVEL 2922000.000
MAG_ZEROPOINT 34.22
MAG_GAMMA 4.0
GAIN 1.0
PIXEL_SCALE 0.300
SEEING_FWHM 0.471
STARNNW_NAME $PM/../tbl/sex_default.nnw
BACK_SIZE 32
BACK_FILTERSIZE 3
BACKPHOTO_TYPE GLOBAL
CHECKIMAGE_TYPE APERTURES
MEMORY_OBJSTACK 5000
MEMORY_PIXSTACK 200000
MEMORY_BUFSIZE 1024
THRESH_TYPE RELATIVE
WEIGHT_TYPE MAP_WEIGHT
WEIGHT_GAIN N
WEIGHT_IMAGE wght10h_H.fits
```

D Evaluate configuration

D.1 For photometry mode

D.1.1 Photometry with a common PSF of $1.8''$

```
0   1    1                  > (TEST,CLEAN,POL), FIX_PSF, EFF_PSF
22  0 0.00    7.22 0.00 0.  > SUM: (W_SUM,SUM),-, APERT, PSF: SX,SY,AL
0  20 1.00 0000.0           > FIT: (yes/no),ITER,CHI_LIM, SATUR/1000.
0  48 3.00 0.0000           > BACK:(FIT,L,HST),HW, S_LOC, B_OFF/1000.
1   4 1.30    4.00 0.00     > PROF:(DIRT,EDGE),R_MIN, X, S_DIRT, -
11 20    4    0.00 0.00     > MAXI:( FIX,FILT), ITER,RAD, X0, Y0, -
```

D.1.2 Photometry with a common PSF of $2.1''$

```
0   1    1                  > (TEST,CLEAN,POL), FIX_PSF, EFF_PSF
22  0 0.00    9.23 0.00 0.  > SUM: (W_SUM,SUM),-, APERT, PSF: SX,SY,AL
0  20 1.00 0000.0           > FIT: (yes/no),ITER,CHI_LIM, SATUR/1000.
0  48 3.00 0.0000           > BACK:(FIT,L,HST),HW, S_LOC, B_OFF/1000.
1   4 1.30    4.00 0.00     > PROF:(DIRT,EDGE),R_MIN, X, S_DIRT, -
11 20    4    0.00 0.00     > MAXI:( FIX,FILT), ITER,RAD, X0, Y0, -
```

D.2 For object shape fitting

```
100 0    0                  > (TEST,CLEAN,POL), FIX_PSF, EFF_PSF
01  0 0.00    0.00 0.00 0.  > SUM: (W_SUM,SUM),-, APERT, PSF: SX,SY,AL
2  20 10.00   0.0000        > FIT: (yes/no),ITER,CHI_LIM, SATUR/1000.
1  40 3.00    0.0000        > BACK:(FIT,L,HST),HW, S_LOC, B_OFF/1000.
1  8  1.50      4.00 0.00   > PROF:(DIRT,EDGE),R_MIN, X, S_DIRT, -
11 20    5      0.00 0.00   > MAXI:( FIX,FILT), ITER,RAD, X0, Y0, -
```

E HIROCS field layout

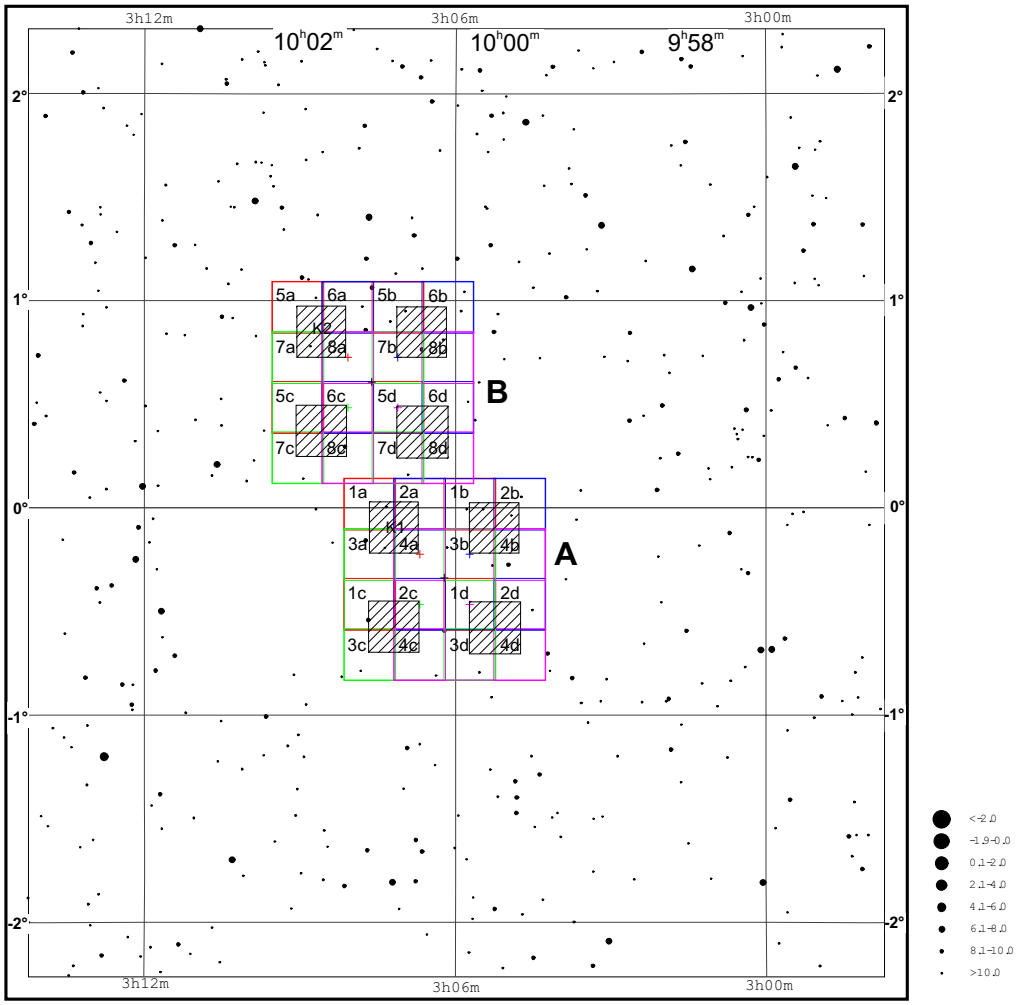

Figure E.1: HIROCS 03h field. All fields are composed of 1 square degree patches, consisting of 16 subfields. These represent the OMEGA2000 FoV of $15.4' \times 15.4'$ (identical to that of one of the four LAICA chips). LAICA images four of these squares in one exposure; they are labeled accordingly (*e.g.*, 1a, 1b, 1c, 1d). The WFI FoV corresponds to four adjacent OMEGA2000 FoVs. In the hatched pointings, short exposures are taken using OMEGA2000 and LAICA, overlapping four science pointings to "connect" the science images for mosaic correction. In the 03h field, the 1 square degree patches are not aligned to avoid bright stars.

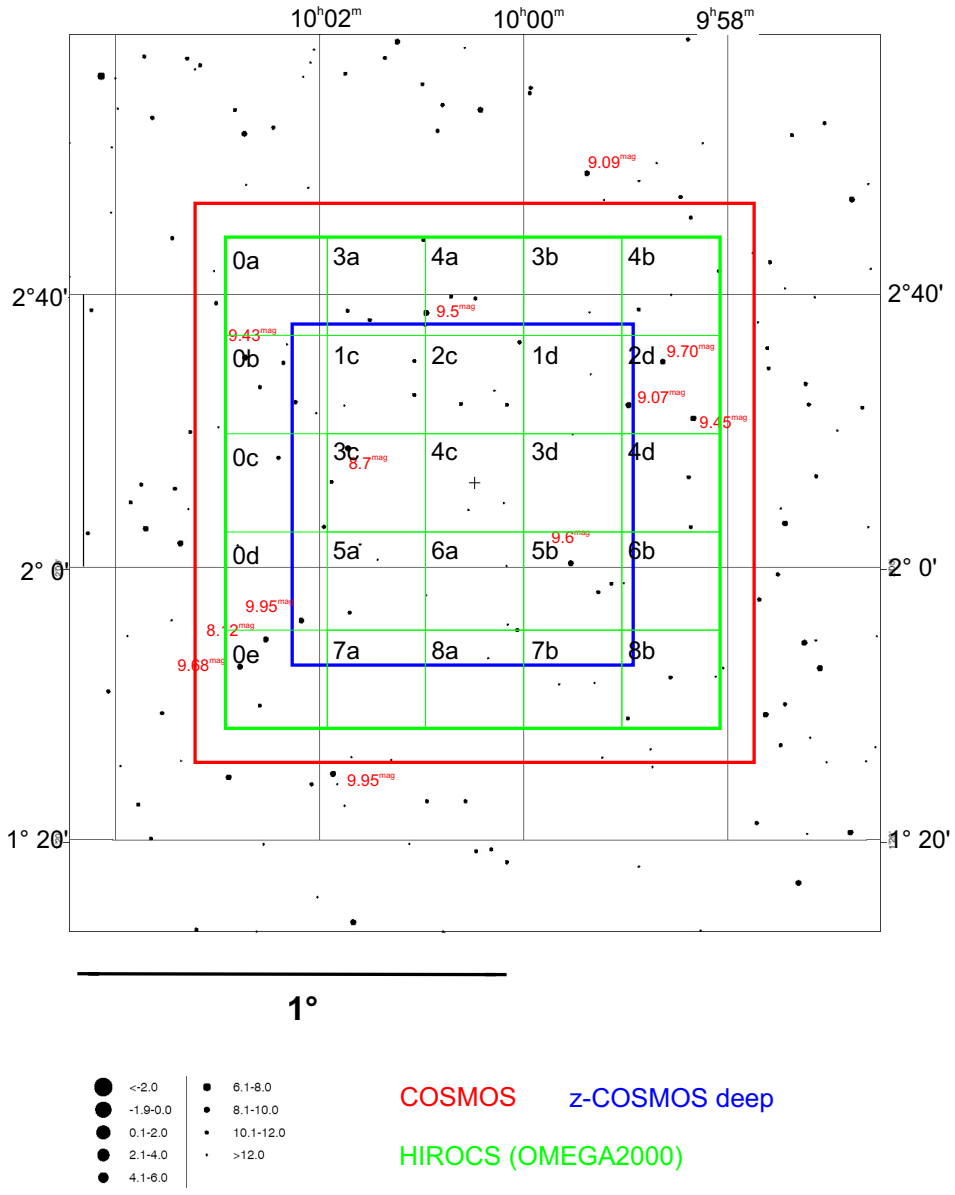

Figure E.2: HIROCS 10h (COSMOS) field. The red line indicates the 2 square degree field covered by most datasets of the COSMOS survey. HIROCS OMEGA2000 pointings are shown as green squares. The area covered by HIROCS H-band observations also corresponds to HST ACS imaging and z-COSMOS bright spectroscopy coverage. The area inside the blue square is to be covered by z-COSMOS deep spectroscopy; see Chapter 3 for details.

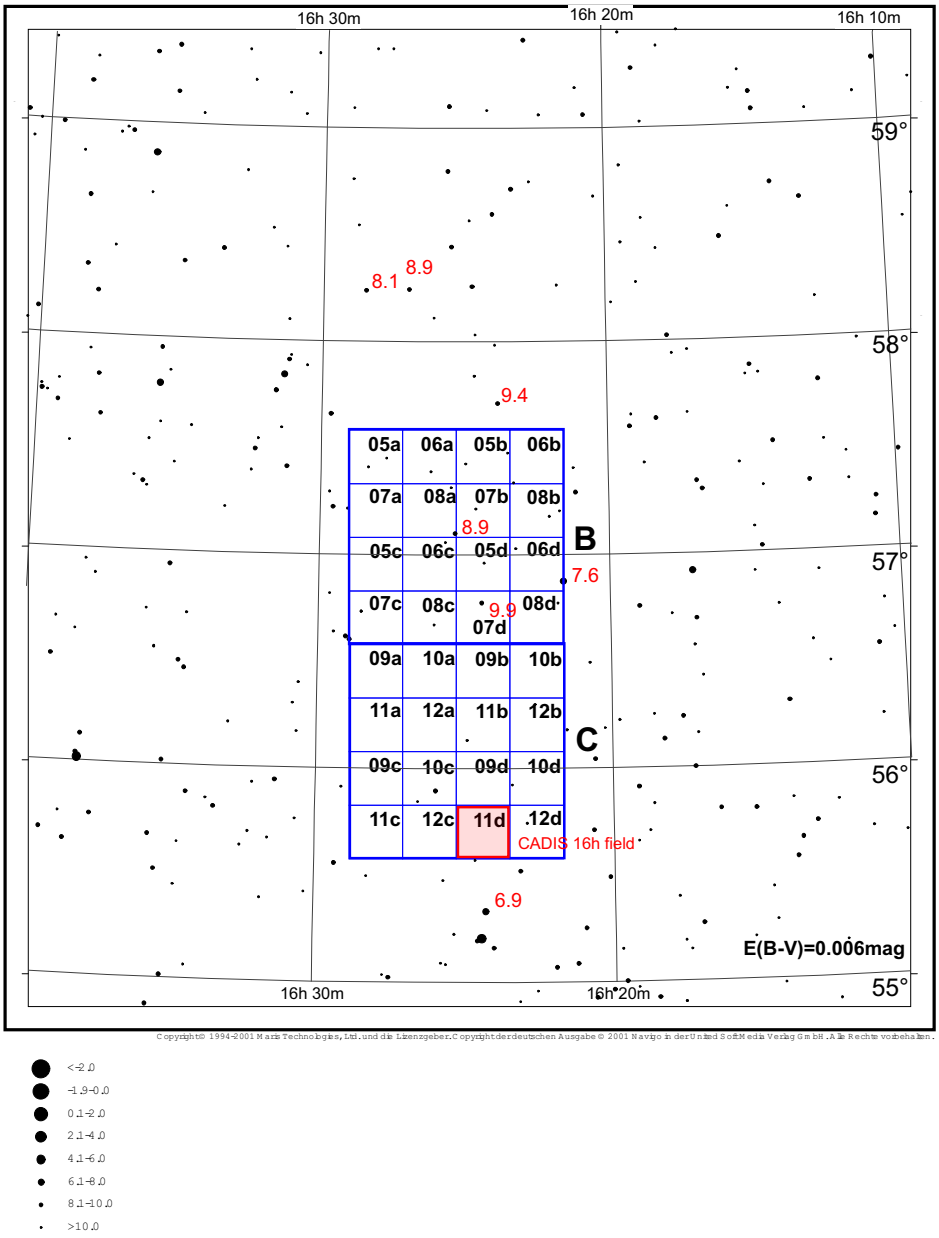

Figure E.3: HIROCS 16h field. See Figure E.1 for details. The red square at the bottom indicates the position of the original CADIS 16h field.

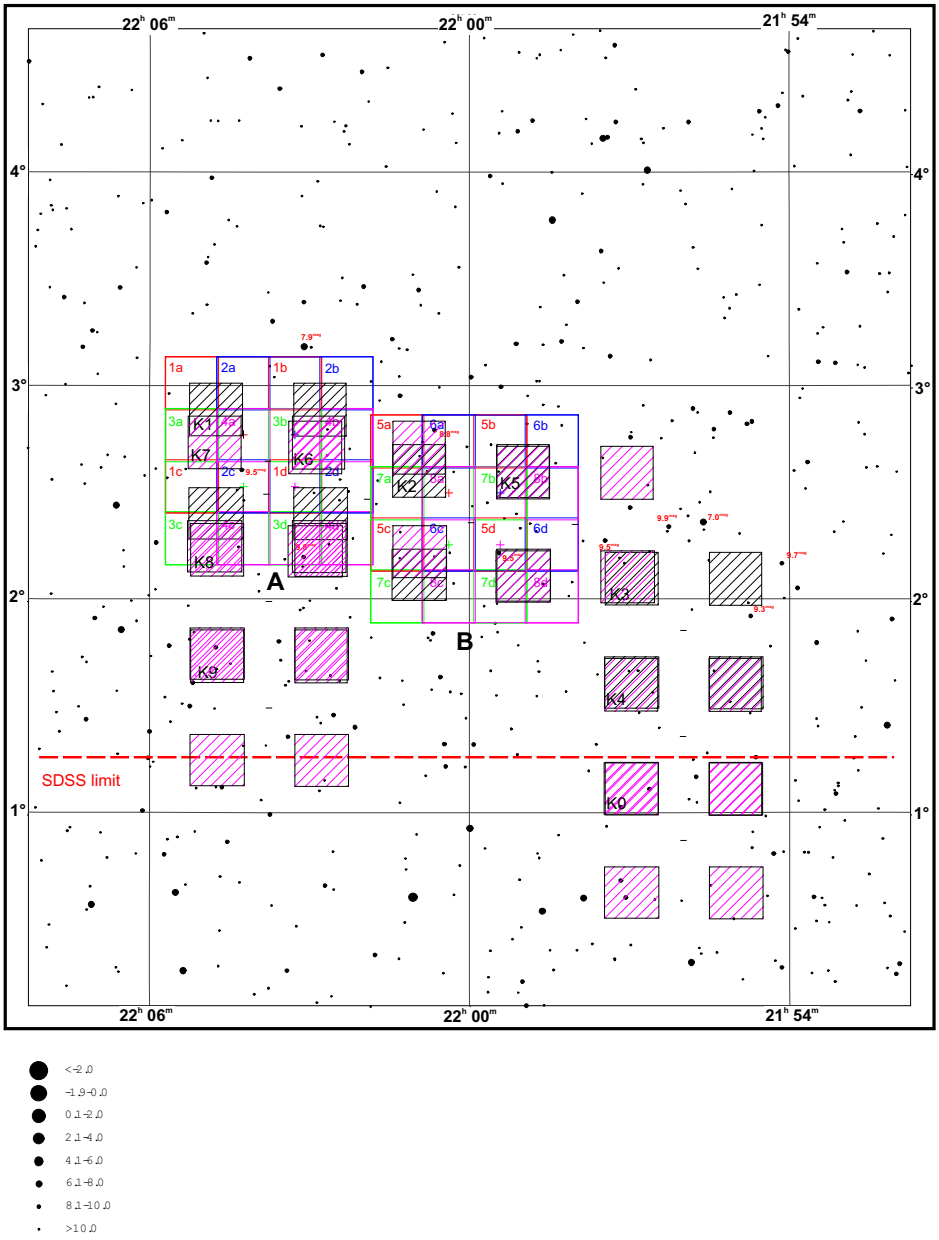

Figure E.4: HIROCS 22h field. See Figure E.1 for details. The dashed red line gives the limit of the SDSS survey area. To be able to use SDSS stars for calibration, the overlapping, purple hatched pointings are taken with LAICA, thus connecting the SDSS photometry to the HIROCS 22h field by matching stars in the overlap regions.

F List of acronyms

2MASS	Two Micron All Sky Survey
AGN	Active Galactic Nucleus
BCG	Brightest Cluster Galaxy
BPM	Bad Pixel Mask
CMB	Cosmic Microwave Background
COSMOS	COSMic evOlution Survey
DR1	Data Release One
FoF	Friends-of-Friends
FoV	Field of View
FWHM	Full Width at Half Maximum
HIROCS	Heidelberg InfraRed / Optical Cluster Survey
ICM	IntraCluster Medium
ΛCDM	Λ Cold Dark Matter
PSF	Point Spread Function
QSO	Quasi-Stellar Object
RMS	Root Mean Square
SDSS	Sloan Digital Sky Survey
SED	Spectral Energy Distribution
SNR	Signal-to-Noise Ratio

List of Figures

List of Tables

Bibliography

George O. Abell. The Distribution of Rich Clusters of Galaxies. *Astrophysical Journal Supplement Series*, 3:211, 1958.

S. Andreon. Galaxy luminosity evolution: How much is due to model choice? *Astronomy and Astrophysics*, 416:865–873, 2004.

S. Andreon, C. Lobo, and A. Iovino. Extending the Butcher-Oemler effect up to z $\sim$ 0.7. *Monthly Notices of the Royal Astronomical Society*, 349:889–898, 2004.

S. Andreon, H. Quintana, M. Tajer, G. Galaz, and J. Surdej. The Butcher-Oemler effect at z $\sim$ 0.35: a change in perspective. *Monthly Notices of the Royal Astronomical Society*, 365:915–928, 2006.

Michael L. Balogh and Simon L. Morris. Halpha photometry of Abell 2390. *Monthly Notices of the Royal Astronomical Society*, 318:703–714, 2000.

W. A. Barkhouse, P. J. Green, A. Vikhlinin, D. W. Kim, D. Perley, R. Cameron, J. Silverman, A. Mossman, R. Burenin, B. T. Jannuzi, M. Kim, M. G. Smith, R. C. Smith, H. Tananbaum, and B. J. Wilkes. ChaMP Serendipitous Galaxy Cluster Survey. *Astrophysical Journal*, 645:955–976, 2006.

M. Bartelmann and S. D. M. White. Cluster detection from surface-brightness fluctuations in SDSS data. *Astronomy and Astrophysics*, 388:732–740, 2002.

Eric F. Bell, Christian Wolf, Klaus Meisenheimer, Hans-Walter Rix, Andrea Borch, Simon Dye, Martina Kleinheinrich, Lutz Wisotzki, and Daniel H. McIntosh. Nearly 5000 Distant Early-Type Galaxies in COMBO-17: A Red Sequence and Its Evolution since z$\sim$1. *Astrophysical Journal*, 608:752–767, 2004.

E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117:393–404, 1996.

H. Böhringer, W. Voges, J. P. Huchra, B. McLean, R. Giacconi, P. Rosati, R. Burg, J. Mader, P. Schuecker, D. Simiç, S. Komossa, T. H. Reiprich, J. Retzlaff, and J. Trümper. The Northern ROSAT All-Sky (NORAS) Galaxy Cluster Survey. I. X-Ray Properties of Clusters Detected as Extended X-Ray Sources. *Astrophysical Journal Supplement Series*, 129:435–474, 2000.

H. Böhringer, P. Schuecker, L. Guzzo, C. A. Collins, W. Voges, R. G. Cruddace, A. Ortiz-Gil, G. Chincarini, S. De Grandi, A. C. Edge, H. T. MacGillivray, D. M. Neumann, S. Schindler, and P. Shaver. The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue. *Astronomy and Astrophysics*, 425:367–383, 2004.

C. S. Botzler, J. Snigula, R. Bender, and U. Hopp. Finding structures in photometric redshift galaxy surveys: an extended friends-of-friends algorithm. *Monthly Notices of the Royal Astronomical Society*, 349:425–439, 2004.

Mark Brodwin, Peter R. Eisenhardt, Anthony H. Gonzalez, Adam Stanford, Daniel Stern, Leonidas A. Moustakas, Michael J. I. Brown, Ranga-Ram Chary, and Audrey Galametz. A Large Population of High Redshift Galaxy Clusters in the IRAC Shallow Cluster Survey. *ArXiv e-prints*, astro-ph/08050813, 2008.

G. Bruzual and S. Charlot. Stellar population synthesis at the resolution of 2003. *Monthly Notices of the Royal Astronomical Society*, 344:1000–1028, 2003.

H. Butcher and Jr. Oemler, A. The evolution of galaxies in clusters. V - A study of populations since Z approximately equal to 0.5. *Astrophysical Journal*, 285:426–438, 1984.

P. Capak, H. Aussel, M. Ajiki, H. J. McCracken, B. Mobasher, N. Scoville, P. Shopbell, Y. Taniguchi, D. Thompson, S. Tribiano, S. Sasaki, A. W. Blain, M. Brusa, C. Carilli, A. Comastri, C. M. Carollo, P. Cassata, J. Colbert, R. S. Ellis, M. Elvis, M. Giavalisco, W. Green, L. Guzzo, G. Hasinger, O. Ilbert, C. Impey, K. Jahnke, J. Kartaltepe, J. P. Kneib, J. Koda, A. Koekemoer, Y. Komiyama, A. Leauthaud, O. Lefevre, S. Lilly, C. Liu, R. Massey, S. Miyazaki, T. Murayama, T. Nagao, J. A. Peacock, A. Pickles, C. Porciani, A. Renzini, J. Rhodes, M. Rich, M. Salvato, D. B. Sanders, C. Scarlata, D. Schiminovich, E. Schinnerer, M. Scodeggio, K. Sheth, Y. Shioya, L. A. M. Tasca, J. E. Taylor, L. Yan, and G. Zamorani. The First Release COSMOS Optical and Near-IR Data and Catalog. *Astrophysical Journal Supplement Series*, 172:99–116, 2007.

John E. Carlstrom, Gilbert P. Holder, and Erik D. Reese. Cosmology with the Sunyaev-Zel'dovich Effect. *Annual Review of Astronomy and Astrophysics*, 40:643–680, 2002.

M. Castellano, S. Salimbeni, D. Trevese, A. Grazian, L. Pentericci, F. Fiore, A. Fontana, E. Giallongo, P. Santini, S. Cristiani, M. Nonino, and E. Vanzella. A Photometrically Detected Forming Cluster of Galaxies at Redshift 1.6 in the GOODS Field. *Astrophysical Journal*, 671:1497–1502, 2007.

M. Cirasuolo, R. J. McLure, J. S. Dunlop, O. Almaini, S. Foucaud, Ian Smail, K. Sekiguchi, C. Simpson, S. Eales, S. Dye, M. G. Watson, M. J. Page, and P. Hirst. The evolution of the near-infrared galaxy luminosity function and colour bimodality up to z $\approx$ 2 from the UKIDSS Ultra Deep Survey Early Data Release. *Monthly Notices of the Royal Astronomical Society*, 380:585–595, 2007.

J. D. Cohn, A. E. Evrard, M. White, D. Croton, and E. Ellingson. Red-sequence cluster finding in the Millennium Simulation. *Monthly Notices of the Royal Astronomical Society*, 382:1738–1750, 2007.

Michael C. Cooper, Jeffrey A. Newman, Alison L. Coil, Darren J. Croton, Brian F. Gerke, Renbin Yan, Marc Davis, S. M. Faber, Puragra Guhathakurta, David C. Koo, Benjamin J. Weiner, and Christopher N. A. Willmer. The DEEP2 galaxy redshift survey: evolution of the colour-density relation at $0.4 < z < 1.35$. *Monthly Notices of the Royal Astronomical Society*, 376:1445–1459, 2007.

Lennox L. Cowie, Antoinette Songaila, Ester M. Hu, and J. G. Cohen. New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields. *Astronomical Journal*, 112:839, 1996.

Darren J. Croton, Volker Springel, Simon D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. The many lives

of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *Monthly Notices of the Royal Astronomical Society*, 365:11–28, 2006.

O. Cucciati, A. Iovino, C. Marinoni, O. Ilbert, S. Bardelli, P. Franzetti, O. Le Fèvre, A. Pollo, G. Zamorani, A. Cappi, L. Guzzo, H. J. McCracken, B. Meneux, R. Scaramella, M. Scodeggio, L. Tresse, E. Zucca, D. Bottini, B. Garilli, V. Le Brun, D. Maccagni, J. P. Picat, G. Vettolani, A. Zanichelli, C. Adami, M. Arnaboldi, S. Arnouts, M. Bolzonella, S. Charlot, P. Ciliegi, T. Contini, S. Foucaud, I. Gavignaud, B. Marano, A. Mazure, R. Merighi, S. Paltani, R. Pellò, L. Pozzetti, M. Radovich, M. Bondi, A. Bongiorno, G. Busarello, S. de La Torre, L. Gregorini, F. Lamareille, G. Mathez, Y. Mellier, P. Merluzzi, V. Ripepi, D. Rizzo, S. Temporin, and D. Vergani. The VIMOS VLT Deep Survey: the build-up of the colour-density relation. *Astronomy and Astrophysics*, 458:39–52, 2006.

Gabriella De Lucia and Jérémy Blaizot. The hierarchical formation of the brightest cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 375:2–14, 2007.

Gabriella De Lucia, Volker Springel, Simon D. M. White, Darren Croton, and Guinevere Kauffmann. The formation history of elliptical galaxies. *Monthly Notices of the Royal Astronomical Society*, 366:499–509, 2006.

Roberto de Propris, S. A. Stanford, Peter R. Eisenhardt, Mark Dickinson, and Richard Elston. The K-Band Luminosity Function in Galaxy Clusters to z $\sim$ 1. *Astronomical Journal*, 118:719–729, 1999.

J. P. Dietrich, T. Erben, G. Lamer, P. Schneider, A. Schwope, J. Hartlap, and M. Maturi. BLOX: the Bonn lensing, optical, and X-ray selected galaxy clusters. I. Cluster catalog construction. *Astronomy and Astrophysics*, 470:821–834, 2007.

A. Dressler. Galaxy morphology in rich clusters - Implications for the formation and evolution of galaxies. *Astrophysical Journal*, 236:351–365, 1980.

N. Drory, G. Feulner, R. Bender, C. S. Botzler, U. Hopp, C. Maraston, C. Mendes de Oliveira, and J. Snigula. The Munich Near-Infrared Cluster Survey - I. Field selection, object extraction and photometry. *Monthly Notices of the Royal Astronomical Society*, 325:550–562, 2001.

J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, D. Larson, G. Hinshaw, L. Page, C. L. Bennett, B. Gold, N. Jarosik, J. L. Weiland, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data. *ArXiv e-prints*, astro-ph/08030586, 2008.

O. J. Eggen, D. Lynden-Bell, and A. R. Sandage. Evidence from the motions of old stars that the Galaxy collapsed. *Astrophysical Journal*, 136:748, 1962.

Peter R. M. Eisenhardt, Mark Brodwin, Anthony H. Gonzalez, S. Adam Stanford, Daniel Stern, Pauline Barmby, Michael J. I. Brown, Kyle Dawson, Arjun Dey, Mamoru Doi, Audrey Galametz, B. T. Jannuzi, C. S. Kochanek, Joshua Meyers, Tomoki Morokuma, and Leonidas A. Moustakas. Clusters of Galaxies in the First Half of the Universe from the IRAC Shallow Survey. *ArXiv e-prints*, astro-ph/08044798, 2008.

Richard J. Elston, Anthony H. Gonzalez, Eric McKenzie, Mark Brodwin, Michael J. I. Brown, Gustavo Cardona, Arjun Dey, Mark Dickinson, Peter R. Eisenhardt, Buell T. Jannuzi, Yen-Ting Lin, Joseph J. Mohr, S. Nicholas Raines, S. A. Stanford, and Daniel Stern. The FLAMINGOS Extragalactic Survey. *Astrophysical Journal*, 639: 816–826, 2006.

S. M. Faber, C. N. A. Willmer, C. Wolf, D. C. Koo, B. J. Weiner, J. A. Newman, M. Im, A. L. Coil, C. Conroy, M. C. Cooper, M. Davis, D. P. Finkbeiner, B. F. Gerke, K. Gebhardt, E. J. Groth, P. Guhathakurta, J. Harker, N. Kaiser, S. Kassin, M. Kleinheinrich, N. P. Konidaris, R. G. Kron, L. Lin, G. Luppino, D. S. Madgwick, K. Meisenheimer, K. G. Noeske, A. C. Phillips, V. L. Sarajedini, R. P. Schiavon, L. Simard, A. S. Szalay, N. P. Vogt, and R. Yan. Galaxy Luminosity Functions to z$\sim$1 from DEEP2 and COMBO-17: Implications for Red Galaxy Formation. *Astrophysical Journal*, 665:265–294, 2007.

Siegfried Falter. *Searching for distant galaxy clusters: Optical observations, cluster search algorithms and the cluster selection function.* Phd thesis, Ruprecht-Karls-Universität, 2006.

René Faßbender. *Studying Cosmic Evolution with the XMM-Newton Distant Cluster Project: X-ray Luminous Galaxy Clusters at $z \gtrsim 1$ and their Galaxy Populations.* Phd thesis, Ludwig-Maximilians-Universität, 2007.

A. Finoguenov, L. Guzzo, G. Hasinger, N. Z. Scoville, H. Aussel, H. Böhringer, M. Brusa, P. Capak, N. Cappelluti, A. Comastri, S. Giodini, R. E. Griffiths, C. Impey, A. M. Koekemoer, J. P. Kneib, A. Leauthaud, O. Le Fèvre, S. Lilly, V. Mainieri, R. Massey, H. J. McCracken, B. Mobasher, T. Murayama, J. A. Peacock, I. Sakelliou, E. Schinnerer, J. D. Silverman, V. Smolcic, Y. Taniguchi, L. Tasca, J. E. Taylor, J. R. Trump, and G. Zamorani. The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies. *Astrophysical Journal Supplement Series*, 172:182–195, 2007.

M. Fioc and B. Rocca-Volmerange. PEGASE: a UV to NIR spectral evolution model of galaxies. Application to the calibration of bright galaxy counts. *Astronomy and Astrophysics*, 326:950–962, 1997.

A. Gabasch, Y. Goranova, U. Hopp, S. Noll, and M. Pannella. A deep i-selected multiwaveband galaxy catalogue in the COSMOS field. *Monthly Notices of the Royal Astronomical Society*, 383:1319–1335, 2008.

J. S. Gallagher. Possible optical evidence for ram-pressure-sweeping in the Hydra I cluster of galaxies. *Astrophysical Journal*, 223:386–390, 1978.

G. Gavazzi and M. Scodeggio. The mass dependence of the star formation history of disk galaxies. *Astronomy and Astrophysics*, 312:L29–L32, 1996.

David G. Gilbank, Richard G. Bower, F. J. Castander, and B. L. Ziegler. Exploring the selection of galaxy clusters and groups: an optical survey for X-ray dark clusters. *Monthly Notices of the Royal Astronomical Society*, 348:551–580, 2004.

Michael D. Gladders and H. K. C. Yee. A New Method For Galaxy Cluster Detection. I. The Algorithm. *Astronomical Journal*, 120:2148–2162, 2000.

Anthony H. Gonzalez, Dennis Zaritsky, Julianne J. Dalcanton, and Amy Nelson. The Las Campanas Distant Cluster Survey: The Catalog. *Astrophysical Journal Supplement Series*, 137:117–138, 2001.

Tomotsugu Goto, Maki Sekiguchi, Robert C. Nichol, Neta A. Bahcall, Rita S. J. Kim, James Annis, Zeljko Ivezic, J. Brinkmann, Gregory S. Hennessy, Gyula P. Szokoly, and Douglas L. Tucker. The Cut-and-Enhance Method: Selecting Clusters of Galaxies from the Sloan Digital Sky Survey Commissioning Data. *Astronomical Journal*, 123: 1807–1825, 2002.

B. Gustafsson, B. Edvardsson, K. Eriksson, U. G. Jørgensen, Å Nordlund, and B. Plez. A grid of MARCS model atmospheres for late-type stars. I. Methods and general properties. *Astronomy and Astrophysics*, 486:951–970, 2008.

L. Guzzo, P. Cassata, A. Finoguenov, R. Massey, N. Z. Scoville, P. Capak, R. S. Ellis, B. Mobasher, Y. Taniguchi, D. Thompson, M. Ajiki, H. Aussel, H. Böhringer, M. Brusa, D. Calzetti, A. Comastri, A. Franceschini, G. Hasinger, M. M. Kasliwal, M. G. Kitzbichler, J. P. Kneib, A. Koekemoer, A. Leauthaud, H. J. McCracken, T. Murayama, T. Nagao, J. Rhodes, D. B. Sanders, S. Sasaki, Y. Shioya, L. Tasca, and J. E. Taylor. The Cosmic Evolution Survey (COSMOS): A Large-Scale Structure at z = 0.73 and the Relation of Galaxy Morphologies to Local Environment. *Astrophysical Journal Supplement Series*, 172:254–269, 2007.

G. Hasinger, N. Cappelluti, H. Brunner, M. Brusa, A. Comastri, M. Elvis, A. Finoguenov, F. Fiore, A. Franceschini, R. Gilli, R. E. Griffiths, I. Lehmann, V. Mainieri, G. Matt, I. Matute, T. Miyaji, S. Molendi, S. Paltani, D. B. Sanders, N. Scoville, L. Tresse, C. M. Urry, P. Vettolani, and G. Zamorani. The XMM-Newton Wide-Field Survey in the COSMOS Field. I. Survey Description. *Astrophysical Journal Supplement Series*, 172:29–37, 2007.

M. A. Hausman and J. P. Ostriker. Galactic cannibalism. III - The morphological evolution of galaxies and clusters. *Astrophysical Journal*, 224:320–336, 1978.

A. K. Hicks, E. Ellingson, M. Bautz, H. K. C. Yee, M. Gladders, and G. Garmire. Chandra X-ray observations of newly discovered, z $\sim$ 1 clusters from the red-sequence cluster survey. *Advances in Space Research*, 36:706–709, 2005.

Henk Hoekstra. A comparison of weak-lensing masses and X-ray properties of galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 379:317–330, 2007.

Rita Seung Jung Kim, Jeremy V. Kepner, Marc Postman, Michael A. Strauss, Neta A. Bahcall, James E. Gunn, Robert H. Lupton, James Annis, Robert C. Nichol, Francisco J. Castander, J. Brinkmann, Robert J. Brunner, Andrew Connolly, Istvan Csabai, Robert B. Hindsley, Zeljko Ivezic, Michael S. Vogeley, and Donald G. York. Detecting Clusters of Galaxies in the Sloan Digital Sky Survey. I. Monte Carlo Comparison of Cluster Detection Algorithms. *Astronomical Journal*, 123:20–36, 2002.

Ivan R. King. The Probable Ages of the Oldest Open Clusters. *Astrophysical Journal*, 151:L59, 1968.

M. G. Kitzbichler and S. D. M. White. The high-redshift galaxy population in hierarchical galaxy formation models. *Monthly Notices of the Royal Astronomical Society*, 376:2–12, 2007.

Rüdiger Kneissl, Michael E. Jones, Richard Saunders, Vincent R. Eke, Anthony N. Lasenby, Keith Grainge, and Garret Cotter. Surveying the sky with the Arcminute MicroKelvin Imager: expected constraints on galaxy cluster evolution and cosmology. *Monthly Notices of the Royal Astronomical Society*, 328:783–794, 2001.

T. Kodama and N. Arimoto. Origin of the colour-magnitude relation of elliptical galaxies. *Astronomy and Astrophysics*, 320:41–53, 1997.

Tadayuki Kodama, Ichi Tanaka, Masaru Kajisawa, Jaron Kurk, Bram Venemans, Carlos De Breuck, Joël Vernet, and Chris Lidman. The first appearance of the red sequence of galaxies in proto-clusters at $2 \lesssim z \lesssim 3$. *Monthly Notices of the Royal Astronomical Society*, 377:1717–1725, 2007.

A. M. Koekemoer, H. Aussel, D. Calzetti, P. Capak, M. Giavalisco, J. P. Kneib, A. Leauthaud, O. Le Fèvre, H. J. McCracken, R. Massey, B. Mobasher, J. Rhodes, N. Scoville, and P. L. Shopbell. The COSMOS Survey: Hubble Space Telescope Advanced Camera for Surveys Observations and Data Processing. *Astrophysical Journal Supplement Series*, 172:196–202, 2007.

Benjamin P. Koester, Timothy A. McKay, James Annis, Risa H. Wechsler, August E. Evrard, Eduardo Rozo, Lindsey Bleem, Erin S. Sheldon, and David Johnston. MaxBCG: A Red-Sequence Galaxy Cluster Finder. *Astrophysical Journal*, 660:221–238, 2007.

Bernd von Kuhlmann. *Commissioning and performance tests of MOSCA a focal reducer for the 3.5-meter-telescope at Calar Alto*. Diploma thesis, Ruprecht-Karls-Universität, 1997.

Katy Lancaster, Ricardo Genova-Santos, Nelson Falcòn, Keith Grainge, Carlos Gutièrrez, Rüdiger Kneissl, Phil Marshall, Guy Pooley, Rafael Rebolo, Jose-Alberto Rubiño-Martin, Richard D. E. Saunders, Elizabeth Waldram, and Robert A. Watson. Very Small Array observations of the Sunyaev-Zel'dovich effect in nearby galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 359:16–30, 2005.

O. Le Fèvre, G. Vettolani, B. Garilli, L. Tresse, D. Bottini, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, A. Zanichelli, C. Adami, M. Arnaboldi, S. Arnouts, S. Bardelli, M. Bolzonella, A. Cappi, S. Charlot, P. Ciliegi, T. Contini, S. Foucaud, P. Franzetti, I. Gavignaud, L. Guzzo, O. Ilbert, A. Iovino, H. J. McCracken, B. Marano, C. Marinoni, G. Mathez, A. Mazure, B. Meneux, R. Merighi, S. Paltani, R. Pellò, A. Pollo, L. Pozzetti, M. Radovich, G. Zamorani, E. Zucca, M. Bondi, A. Bongiorno, G. Busarello, F. Lamareille, Y. Mellier, P. Merluzzi, V. Ripepi, and D. Rizzo. The VIMOS VLT deep survey. First epoch VVDS-deep survey: 11 564 spectra with $17.5 \leq \mathrm{IAB} \leq 24$, and the redshift distribution over $0 \leq z \leq 5$. *Astronomy and Astrophysics*, 439:845–862, 2005.

S. J. Lilly, O. Le Fèvre, A. Renzini, G. Zamorani, M. Scodeggio, T. Contini, C. M. Carollo, G. Hasinger, J. P. Kneib, A. Iovino, V. Le Brun, C. Maier, V. Mainieri, M. Mignoli, J. Silverman, L. A. M. Tasca, M. Bolzonella, A. Bongiorno, D. Bottini, P. Capak, K. Caputi, A. Cimatti, O. Cucciati, E. Daddi, R. Feldmann, P. Franzetti, B. Garilli, L. Guzzo, O. Ilbert, P. Kampczyk, K. Kovac, F. Lamareille, A. Leauthaud, J. F. Le Borgne, H. J. McCracken, C. Marinoni, R. Pello, E. Ricciardelli, C. Scarlata, D. Vergani, D. B. Sanders, E. Schinnerer, N. Scoville, Y. Taniguchi,

S. Arnouts, H. Aussel, S. Bardelli, M. Brusa, A. Cappi, P. Ciliegi, A. Finoguenov, S. Foucaud, R. Franceschini, C. Halliday, C. Impey, C. Knobel, A. Koekemoer, J. Kurk, D. Maccagni, S. Maddox, B. Marano, G. Marconi, B. Meneux, B. Mobasher, C. Moreau, J. A. Peacock, C. Porciani, L. Pozzetti, R. Scaramella, D. Schiminovich, P. Shopbell, I. Smail, D. Thompson, L. Tresse, G. Vettolani, A. Zanichelli, and E. Zucca. zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field. *Astrophysical Journal Supplement Series*, 172:70–85, 2007.

Yen-Ting Lin and Joseph J. Mohr. K-band Properties of Galaxy Clusters and Groups: Brightest Cluster Galaxies and Intracluster Light. *Astrophysical Journal*, 617:879–895, 2004.

Robert H. Lupton, James E. Gunn, and Alexander S. Szalay. A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements. *Astronomical Journal*, 118:1406–1410, 1999.

MANOS-Proposal. http://www.mpia.de/HIROCS/MANOS-proposal.pdf.

V. E. Margoniner, R. R. de Carvalho, R. R. Gal, and S. G. Djorgovski. The Butcher-Oemler Effect in 295 Clusters: Strong Redshift Evolution and Cluster Richness Dependence. *Astrophysical Journal*, 548:L143–L146, 2001.

Patrick J. McCarthy, Haojing Yan, Roberto G. Abraham, Erin Mentuch, Karl Glazebrook, Lin Yan, Hsiao-Wen Chen, S. Eric Persson, Preethi Nair, Sandra Savaglio, David Crampton, Stephanie Juneau, Damien Le Borgne, R. G. Carlberg, Ronald O. Marzke, Inger Jørgensen, Kathy Roth, and Richard Murowinski. A Compact Cluster of Massive Red Galaxies at a Redshift of 1.5. *Astrophysical Journal*, 664:L17–L21, 2007.

Yannick Mellier. Probing the Universe with Weak Lensing. *Annual Review of Astronomy and Astrophysics*, 37:127–189, 1999.

B. Mobasher, P. Capak, N. Z. Scoville, T. Dahlen, M. Salvato, H. Aussel, D. J. Thompson, R. Feldmann, L. Tasca, O. Lefevre, S. Lilly, C. M. Carollo, J. S. Kartaltepe, H. McCracken, J. Mould, A. Renzini, D. B. Sanders, P. L. Shopbell, Y. Taniguchi, M. Ajiki, Y. Shioya, T. Contini, M. Giavalisco, O. Ilbert, A. Iovino, V. Le Brun, V. Mainieri, M. Mignoli, and M. Scodeggio. Photometric Redshifts of Galaxies in COSMOS. *Astrophysical Journal Supplement Series*, 172:117–131, 2007.

M. Moles, N. Benítez, J. A. L. Aguerri, E. J. Alfaro, T. Broadhurst, J. Cabrera-Caño, F. J. Castander, J. Cepa, M. Cerviño, D. Cristóbal-Hornillos, A. Fernández-Soto, R. M. González Delgado, L. Infante, I. Márquez, V. J. Martínez, J. Masegosa, A. del Olmo, J. Perea, F. Prada, J. M. Quintana, and S. F. Sánchez. The ALHAMBRA Project: A large area multi medium-band optical and NIR photometric survey. *ArXiv e-prints*, astro-ph/08063021, 2008.

Ben Moore, Neal Katz, George Lake, Alan Dressler, and Augustus Oemler. Galaxy harassment and the evolution of clusters of galaxies. *Nature*, 379:613–616, 1996.

Stephen Muchovej, Tony Mroczkowski, John E. Carlstrom, John Cartwright, Christopher Greer, Ryan Hennessy, Michael Loh, Clem Pryke, Ben Reddall, Marcus Runyan, Matthew Sharp, David Hawkins, James W. Lamb, David Woody, Marshall Joy,

Erik M. Leitch, and Amber D. Miller. Observations of High-Redshift X-Ray Selected Clusters with the Sunyaev-Zel'dovich Array. *Astrophysical Journal*, 663:708–716, 2007.

John S. Mulchaey, David S. Davis, Richard F. Mushotzky, and David Burstein. The Intragroup Medium in Poor Groups of Galaxies. *Astrophysical Journal*, 456:80, 1996.

C. R. Mullis, P. Rosati, G. Lamer, H. Böhringer, A. Schwope, P. Schuecker, and R. Fassbender. Discovery of an X-Ray-luminous Galaxy Cluster at z = 1.4. *Astrophysical Journal*, 623:L85–L88, 2005.

Danial Obreschkow. in preparation.

Jeremiah P. Ostriker, Lori M. Lubin, and Lars Hernquist. Using X-rays to determine which compact groups are illusory. *Astrophysical Journal*, 444:L61–L64, 1995.

Yichuan C. Pei. Interstellar dust from the Milky Way to the Magellanic Clouds. *Astrophysical Journal*, 395:130–139, 1992.

A. J. Pickles. A Stellar Spectral Flux Library: 1150-25000 Å. *Publications of the Astronomical Society of the Pacific*, 110:863–878, 1998.

M. Pierre, F. Pacaud, P. A. Duc, J. P. Willis, S. Andreon, I. Valtchanov, B. Altieri, G. Galaz, A. Gueguen, J. P. Le Fèvre, O. Le Fèvre, T. Ponman, P. G. Sprimont, J. Surdej, C. Adami, A. Alshino, M. Bremer, L. Chiappetti, A. Detal, O. Garcet, E. Gosset, C. Jean, D. Maccagni, C. Marinoni, A. Mazure, H. Quintana, and A. Read. The XMM Large-Scale Structure survey: a well-controlled X-ray cluster sample over the D1 CFHTLS area. *Monthly Notices of the Royal Astronomical Society*, 372:591–608, 2006.

M. Postman, M. Franx, N. J. G. Cross, B. Holden, H. C. Ford, G. D. Illingworth, T. Goto, R. Demarco, P. Rosati, J. P. Blakeslee, K. V. Tran, N. Benítez, M. Clampin, G. F. Hartig, N. Homeier, D. R. Ardila, F. Bartko, R. J. Bouwens, L. D. Bradley, T. J. Broadhurst, R. A. Brown, C. J. Burrows, E. S. Cheng, P. D. Feldman, D. A. Golimowski, C. Gronwall, L. Infante, R. A. Kimble, J. E. Krist, M. P. Lesser, A. R. Martel, S. Mei, F. Menanteau, G. R. Meurer, G. K. Miley, V. Motta, M. Sirianni, W. B. Sparks, H. D. Tran, Z. I. Tsvetanov, R. L. White, and W. Zheng. The Morphology-Density Relation in z $\sim$ 1 Clusters. *Astrophysical Journal*, 623:721–741, 2005.

P. Predehl, G. Hasinger, H. Böhringer, U. Briel, H. Brunner, E. Churazov, M. Freyberg, P. Friedrich, E. Kendziorra, D. Lutz, N. Meidinger, M. Pavlinsky, E. Pfeffermann, A. Santangelo, J. Schmitt, P. Schuecker, A. Schwope, M. Steinmetz, L. Strüder, R. Sunyaev, and J. Wilms. eROSITA. In *Space Telescopes and Instrumentation II: Ultraviolet to Gamma Ray. Edited by Turner, Martin J. L.; Hasinger, Günther. Proceedings of the SPIE*, volume 6266, page 19, 2006.

M. K. M. Prescott, C. D. Impey, R. J. Cool, and N. Z. Scoville. Quasars in the COSMOS Field. *Astrophysical Journal*, 644:100–115, 2006.

Massimo Ramella, Margaret J. Geller, Armando Pisani, and Luiz N. da Costa. The UZC-SSRS2 Group Catalog. *Astronomical Journal*, 123:2976–2984, 2002.

RCS-2. http://www.rcs2.org/summary.html.

A. Kathy Romer, Pedro T. P. Viana, Andrew R. Liddle, and Robert G. Mann. A Serendipitous Galaxy Cluster Survey with XMM: Expected Catalog Properties and Scientific Applications. *Astrophysical Journal*, 547:594–608, 2001.

Piero Rosati, Roberto della Ceca, Colin Norman, and Riccardo Giacconi. The ROSAT Deep Cluster Survey: The X-Ray Luminosity Function Out to z = 0.8. *Astrophysical Journal*, 492:L21, 1998.

H. J. Röser and K. Meisenheimer. The synchrotron light from the jet of 3C 273. *Astronomy and Astrophysics*, 252:458–474, 1991.

H. J. Röser, H. H. Hippelein, and C. Wolf. The Heidelberg InfraRed Optical Cluster Survey (HIROCS). In *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, page 46, 2004.

John Ruhl, Peter A. R. Ade, John E. Carlstrom, Hsiao-Mei Cho, Thomas Crawford, Matt Dobbs, Chris H. Greer, Nils w Halverson, William L. Holzapfel, Trevor M. Lanting, Adrian T. Lee, Erik M. Leitch, Jon Leong, Wenyang Lu, Martin Lueker, Jared Mehl, Stephan S. Meyer, Joe J. Mohr, Steve Padin, T. Plagge, Clem Pryke, Marcus C. Runyan, Dan Schwan, M. K. Sharp, Helmuth Spieler, Zak Staniszewski, and Antony A. Stark. The South Pole Telescope. In *Millimeter and Submillimeter Detectors for Astronomy II. Edited by Jonas Zmuidzinas, Wayne S. Holland and Stafford Withington Proceedings of the SPIE*, volume 5498, pages 11–29, 2004.

D. B. Sanders, M. Salvato, H. Aussel, O. Ilbert, N. Scoville, J. A. Surace, D. T. Frayer, K. Sheth, G. Helou, T. Brooke, B. Bhattacharya, L. Yan, J. S. Kartaltepe, J. E. Barnes, A. W. Blain, D. Calzetti, P. Capak, C. Carilli, C. M. Carollo, A. Comastri, E. Daddi, R. S. Ellis, M. Elvis, S. M. Fall, A. Franceschini, M. Giavalisco, G. Hasinger, C. Impey, A. Koekemoer, O. Le Fèvre, S. Lilly, M. C. Liu, H. J. McCracken, B. Mobasher, A. Renzini, M. Rich, E. Schinnerer, P. L. Shopbell, Y. Taniguchi, D. J. Thompson, C. M. Urry, and J. P. Williams. S-COSMOS: The Spitzer Legacy Survey of the Hubble Space Telescope ACS 2 deg^2 COSMOS Field I: Survey Strategy and First Analysis. *Astrophysical Journal Supplement Series*, 172: 86–98, 2007.

E. Schinnerer, V. Smolcic, C. L. Carilli, M. Bondi, P. Ciliegi, K. Jahnke, N. Z. Scoville, H. Aussel, F. Bertoldi, A. W. Blain, C. D. Impey, A. M. Koekemoer, O. Le Fevre, and C. M. Urry. The VLA-COSMOS Survey. II. Source Catalog of the Large Project. *Astrophysical Journal Supplement Series*, 172:46–69, 2007.

Peter Schücker. New Cosmology with Clusters of Galaxies. In *Reviews in Modern Astronomy*, volume 18, pages 76–105, 2005.

N. Scoville, R. G. Abraham, H. Aussel, J. E. Barnes, A. Benson, A. W. Blain, D. Calzetti, A. Comastri, P. Capak, C. Carilli, J. E. Carlstrom, C. M. Carollo, J. Colbert, E. Daddi, R. S. Ellis, M. Elvis, S. P. Ewald, M. Fall, A. Franceschini, M. Giavalisco, W. Green, R. E. Griffiths, L. Guzzo, G. Hasinger, C. Impey, J. P. Kneib, J. Koda, A. Koekemoer, O. Lefevre, S. Lilly, C. T. Liu, H. J. McCracken, R. Massey, Y. Mellier, S. Miyazaki, B. Mobasher, J. Mould, C. Norman, A. Refregier, A. Renzini, J. Rhodes, M. Rich, D. B. Sanders, D. Schiminovich, E. Schinnerer, M. Scodeggio, K. Sheth, P. L. Shopbell, Y. Taniguchi, N. D. Tyson, C. M. Urry, L. Van Waerbeke, P. Vettolani, S. D. M. White, and L. Yan. COSMOS: Hubble

Space Telescope Observations. *Astrophysical Journal Supplement Series*, 172:38–45, 2007a.

N. Scoville, H. Aussel, A. Benson, A. Blain, D. Calzetti, P. Capak, R. S. Ellis, A. El-Zant, A. Finoguenov, M. Giavalisco, L. Guzzo, G. Hasinger, J. Koda, O. Le Fèvre, R. Massey, H. J. McCracken, B. Mobasher, A. Renzini, J. Rhodes, M. Salvato, D. B. Sanders, S. S. Sasaki, E. Schinnerer, K. Sheth, P. L. Shopbell, Y. Taniguchi, J. E. Taylor, and D. J. Thompson. Large Structures and Galaxy Evolution in COSMOS at z $<$ 1.1. *Astrophysical Journal Supplement Series*, 172:150–181, 2007b.

N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J. P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson. The Cosmic Evolution Survey (COSMOS): Overview. *Astrophysical Journal Supplement Series*, 172:1–8, 2007c.

S. A. Shectman. Clusters of galaxies from the Shane-Wirtanen counts. *Astrophysical Journal Supplement Series*, 57:77–90, 1985.

M. F. Skrutskie, R. M. Cutri, R. Stiening, M. D. Weinberg, S. Schneider, J. M. Carpenter, C. Beichman, R. Capps, T. Chester, J. Elias, J. Huchra, J. Liebert, C. Lonsdale, D. G. Monet, S. Price, P. Seitzer, T. Jarrett, J. D. Kirkpatrick, J. E. Gizis, E. Howard, T. Evans, J. Fowler, L. Fullmer, R. Hurt, R. Light, E. L. Kopan, K. A. Marsh, H. L. McCallon, R. Tam, S. Van Dyk, and S. Wheelock. The Two Micron All Sky Survey (2MASS). *Astronomical Journal*, 131:1163–1183, 2006.

Volker Springel, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, John A. Peacock, Shaun Cole, Peter Thomas, Hugh Couchman, August Evrard, Jörg Colberg, and Frazer Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, 2005.

S. A. Stanford, Peter R. Eisenhardt, Mark Brodwin, Anthony H. Gonzalez, Daniel Stern, B. T. Jannuzi, Arjun Dey, Michael J. I. Brown, Eric McKenzie, and Richard Elston. An IR-selected Galaxy Cluster at z = 1.41. *Astrophysical Journal*, 634:L129–L132, 2005.

R. A. Sunyaev and Ia B. Zeldovich. Microwave background radiation as a probe of the contemporary structure and history of the universe. *Annual Review of Astronomy and Astrophysics*, 18:537–560, 1980.

Alexander S. Szalay, Takahiko Matsubara, and Stephen D. Landy. Redshift-Space Distortions of the Correlation Function in Wide-Angle Galaxy Surveys. *Astrophysical Journal*, 498:L1, 1998.

M. I. Takahashi, Y. Shioya, Y. Taniguchi, T. Murayama, M. Ajiki, S. S. Sasaki, O. Koizumi, T. Nagao, N. Z. Scoville, B. Mobasher, H. Aussel, P. Capak, C. Carilli, R. S. Ellis, B. Garilli, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, M. G. Kitzbichler, A. Koekemoer, O. Le Fèvre, S. J. Lilly, D. Maccagni, A. Renzini, M. Rich, D. B. Sanders, E. Schinnerer, M. Scodeggio, P. Shopbell, V. Smolcic, S. Tribiano, Y. Ideue, and S. Mihara. The [O II] lambda3727 Luminosity Function and Star Formation Rate at z $\sim$ 1.2 in the COSMOS 2 Square Degree Field and the Subaru Deep Field. *Astrophysical Journal Supplement Series*, 172:456–467, 2007.

A. N. Taylor, D. J. Bacon, M. E. Gray, C. Wolf, K. Meisenheimer, S. Dye, A. Borch, M. Kleinheinrich, Z. Kovacs, and L. Wisotzki. Mapping the 3D dark matter with weak lensing in COMBO-17. *Monthly Notices of the Royal Astronomical Society*, 353:1176–1196, 2004.

C. van Breukelen, L. Clewley, D. G. Bonfield, S. Rawlings, M. J. Jarvis, J. M. Barr, S. Foucaud, O. Almaini, M. Cirasuolo, G. Dalton, J. S. Dunlop, A. C. Edge, P. Hirst, R. J. McLure, M. J. Page, K. Sekiguchi, C. Simpson, I. Smail, and M. G. Watson. Galaxy clusters at $0.6 < z < 1.4$ in the UKIDSS Ultra Deep Survey Early Data Release. *Monthly Notices of the Royal Astronomical Society*, 373:L26–L30, 2006.

Caroline van Breukelen, Garret Cotter, Steve Rawlings, Tony Readhead, David Bonfield, Lee Clewley, Rob Ivison, Matt Jarvis, Chris Simpson, and Mike Watson. Spectroscopic follow-up of a cluster candidate at $z = 1.45$. *Monthly Notices of the Royal Astronomical Society*, 382:971–984, 2007.

P. G. van Dokkum and M. Franx. The Fundamental Plane in CL 0024 at $z = 0.4$: implications for the evolution of the mass-to-light ratio. *Monthly Notices of the Royal Astronomical Society*, 281:985–1000, 1996.

S. D. M. White and M. J. Rees. Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. *Monthly Notices of the Royal Astronomical Society*, 183:341–358, 1978.

Gillian Wilson, Adam Muzzin, Mark Lacy, Howard Yee, Jason Surace, Carol Lonsdale, Henk Hoekstra, Subhabrata Majumdar, David Gilbank, and Mike Gladders. Clusters of Galaxies at $1 < z < 2$: The Spitzer Adaptation of the Red-Sequence Cluster Survey. *ArXiv e-prints*, astro-ph/06044289, 2006.

C. Wolf, K. Meisenheimer, and H. J. Röser. Object classification in astronomical multi-color surveys. *Astronomy and Astrophysics*, 365:660–680, 2001a.

C. Wolf, K. Meisenheimer, H. J. Röser, S. V. W. Beckwith, Jr. Chaffee, F. H., J. Fried, H. Hippelein, J. S. Huang, M. Kümmel, B. von Kuhlmann, C. Maier, S. Phleps, H. W. Rix, E. Thommes, and D. Thompson. Multi-color classification in the Calar Alto Deep Imaging Survey. *Astronomy and Astrophysics*, 365:681–698, 2001b.

C. Wolf, K. Meisenheimer, H. W. Rix, A. Borch, S. Dye, and M. Kleinheinrich. The COMBO-17 survey: Evolution of the galaxy luminosity function from 25 000 galaxies with $0.2 < z < 1.2$. *Astronomy and Astrophysics*, 401:73–98, 2003.

C. Wolf, K. Meisenheimer, M. Kleinheinrich, A. Borch, S. Dye, M. Gray, L. Wisotzki, E. F. Bell, H. W. Rix, A. Cimatti, G. Hasinger, and G. Szokoly. A catalogue of the Chandra Deep Field South with multi-colour classification and photometric redshifts from COMBO-17. *Astronomy and Astrophysics*, 421:913–936, 2004.

C. Wolf, M. E. Gray, and K. Meisenheimer. Red-sequence galaxies with young stars and dust: the cluster Abell 901/902 seen with COMBO-17. *Astronomy and Astrophysics*, 443:435–449, 2005.

Christian Wolf. *Vielfarbenklassifikation in CADIS und die Suche nach Quasaren*. Phd thesis, Ruprecht-Karls-Universität, 1999.

Donald G. York, J. Adelman, Jr. Anderson, John E., Scott F. Anderson, James Annis, Neta A. Bahcall, J. A. Bakken, Robert Barkhouser, Steven Bastian, Eileen Berman,

William N. Boroski, Steve Bracker, Charlie Briegel, John W. Briggs, J. Brinkmann, Robert Brunner, Scott Burles, Larry Carey, Michael A. Carr, Francisco J. Castander, Bing Chen, Patrick L. Colestock, A. J. Connolly, J. H. Crocker, István Csabai, Paul C. Czarapata, John Eric Davis, Mamoru Doi, Tom Dombeck, Daniel Eisenstein, Nancy Ellman, Brian R. Elms, Michael L. Evans, Xiaohui Fan, Glenn R. Federwitz, Larry Fiscelli, Scott Friedman, Joshua A. Frieman, Masataka Fukugita, Bruce Gillespie, James E. Gunn, Vijay K. Gurbani, Ernst de Haas, Merle Haldeman, Frederick H. Harris, J. Hayes, Timothy M. Heckman, G. S. Hennessy, Robert B. Hindsley, Scott Holm, Donald J. Holmgren, Chi-hao Huang, Charles Hull, Don Husby, Shin-Ichi Ichikawa, Takashi Ichikawa, Zeljko Ivezic, Stephen Kent, Rita S. J. Kim, E. Kinney, Mark Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, John Korienek, Richard G. Kron, Peter Z. Kunszt, D. Q. Lamb, B. Lee, R. French Leger, Siriluk Limmongkol, Carl Lindenmeyer, Daniel C. Long, Craig Loomis, Jon Loveday, Rich Lucinio, Robert H. Lupton, Bryan MacKinnon, Edward J. Mannery, P. M. Mantsch, Bruce Margon, Peregrine McGehee, Timothy A. McKay, Avery Meiksin, Aronne Merelli, David G. Monet, Jeffrey A. Munn, Vijay K. Narayanan, Thomas Nash, Eric Neilsen, Rich Neswold, Heidi Jo Newberg, R. C. Nichol, Tom Nicinski, Mario Nonino, Norio Okada, Sadanori Okamura, Jeremiah P. Ostriker, Russell Owen, A. George Pauls, et al. The Sloan Digital Sky Survey: Technical Summary. *Astronomical Journal*, 120:1579–1587, 2000.

M. Zatloukal, H. J. Röser, C. Wolf, H. Hippelein, and S. Falter. Distant galaxy clusters in the COSMOS field found by HIROCS. *Astronomy and Astrophysics*, 474:L5–L8, 2007.

F. Zwicky. Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, 6:110–127, 1933.

Fritz Zwicky, E. Herzog, and P. Wild. *Catalogue of galaxies and of clusters of galaxies*. Pasadena: California Institute of Technology (CIT), 1961-1968. 1968.

Acknowledgments

It is my pleasure to thank my supervisor, Hermann-Josef Röser, to give me the opportunity to work on this interesting project, to be accessible and helpful at any time, be it at home or on vacation, give advice and support when needed, and always respect my own wishes and needs.

My sincere thanks go to Matthias Bartelmann, who kindly agreed to be member of my IMPRS thesis committee and second thesis referee. His enthusiasm for astronomy and positive attitude were always very encouraging for me.

I would like to thank Eva Schinnerer for valuable advice as member of my IMPRS thesis committee and outside, especially on career planning, and helping me to deal with COSMOS data.

Many thanks go to Hans Hippelein, who always had an open ear for any of my questions, helping me to deal with many different kinds of software and data reduction problems, for reading parts of my thesis, and for his unique sense of humor.

It was a true pleasure to work with Christian Wolf. Besides valuable advice on many aspects of my thesis, his ability to let others participate in his excitement for astronomy helped me in many ways, and his introduction of the La Silla night sky is one of the most inspiring experiences during my PhD. Thank you for reminding me that there's more to astronomy than black dots on a computer screen.

I am grateful to Sadegh Khochfar for his help during the writing of the paper, pointing out the importance of using simulations to characterize the cluster survey, spending much time on issues related to the mock sky, and his help with any theory-related questions.

The positive personality of Kris Blindert cheered me up during many long proposal-writing sessions, and I am very grateful for her reduction of the follow-up observation MOSCA spectra and her correction of the corresponding chapter of my thesis.

I received a lot of support from Klaus Meisenheimer, be it with respect to the multi-color classification, or by granting financial support for the interesting conferences and summer schools I could attend, which I am very thankful for.

Many thanks go to Siegfried Falter, who helped me a lot during my first time at the institute, and later especially with the writing of the paper.

Danail Obreschkow prepared the mock sky for the characterization of the cluster selection function, which I am very thankful for.

Frank Richter, Uli Hiller and Marco Piroth ensured that I could work smoothly at any time by taking care of all computer problems, which I am very thankful for.

I am very indebted to the staff at Calar Alto Observatory and La Silla Observatory who supported me a lot during my observations.

Special thanks go to my office mates in those three years: Sebastian Haan, Markus Janson, Jan Pitann, Martin Hennemann, Felix Hormuth, and most of all Marie-Helene Nicol who shared all MPIAPHOT and MIDAS problems with me. Thanks to Isabel, Aday, Vernesa and Dominik for relaxing evenings after work.

Erklärung

Ich versichere, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Declaration

I confirm that I produced this work by my own hand, and that I did not use any other resources and help except for those mentioned.

Heidelberg, 1. September 2008 .....

Unterschrift / Signature