Hendrik Jürgen Hildebrandt

Application of photometric redshifts on the correlation properties of galaxies and matter

Dissertation

Hendrik Jürgen Hildebrandt

Application of photometric redshifts on the correlation properties of galaxies and matter

Mathematisch-Naturwissenschaftliche Fakultät
Rheinische Friedrich-Wilhelms-Universität Bonn
2007

Erster Gutachter:	Prof. Dr. Peter Schneider
Zweiter Gutachter:	PD Dr. Bodo Ziegler
Tag der Prüfung:	20.08.2007

Bonn 2007

Contents

Introduction

In the past years cosmology, the science of the universe as a whole, has seen tremendous progress. The Lambda-Cold-Dark-Matter (ΛCDM) scenario is widely accepted as the standard model of cosmology describing the evolution of the universe and its main constituents. Most cosmological parameters are known to a few percent accuracy now. This concordance cosmological model together with the basic theory of cosmological physics is presented in Chap. 1.

Albeit, the main contributors to the energy density of the universe, Dark Matter (DM) and Dark Energy (DE), have not been observed so far in the laboratory. We can now predict the expansion history, the age, the energy density, etc. of the universe but we do not know the physical origin of the majority of the ingredients driving this cosmic evolution.

There is hope that dark matter particles will be detected in the laboratory soon, either via passive detectors that measure the properties of these hypothetically weakly interacting massive particles (WIMPs) penetrating through the Earth, or actively via events created in the next generation of particle accelerators like the Large Hadron Collider.

For DE, however, the situation is different. No concept was presented yet to detect this component, which is responsible for the accelerated expansion of the universe, in the laboratory. With its physical origin being completely unknown the only way to learn more about its properties, like e.g. its equation of state describing the time-evolution, will be from observational cosmology.

Different methods are proposed to shed light on the nature of DE. One of them is the measurement of coherent distortions in the shape of galaxies due to the gravitational deflection of light by the large-scale structure of the universe, called cosmic shear. Another one is the detection of baryonic acoustic oscillations in the two-point correlation function of galaxies. The accuracy of both of these promising methods for constraining DE properties depends heavily on the measurement of redshifts, hence distances, of many million galaxies. This cannot be done in the traditional way by taking spectra for these large samples. Rather, approximate redshifts, called photometric redshifts, must be estimated from the colours of the galaxies.

While the determination of cosmological parameters like the equation of state of DE is not the subject of this thesis, the photometric redshift technique is introduced and analysed in great detail in Chaps. 4 & 5. Understanding the efficacy of this tool and its shortcomings is essential for many large future survey projects tackling the questions above.

Besides these purely cosmological questions which are hoped to be answered by the measurements of galaxy properties, it is the galaxy population itself we are interested in. In particular, we still have no precise picture about how galaxies form and how they evolve. The behaviour of the DM component which is the dominant driver of cosmological structure formation seems to be well understood through large N-body simulations, although we do not know about the nature of the DM particles. In contrast to this, the formation and evolution of galaxies involve mainly well-known baryonic physics. But the processes involved like star-formation, hydrodynamics, radiative feedback, etc. are so complicated that we are still far from a coherent description of galaxy formation and evolution.

This ignorance is partly caused by the fact that we cannot observe a galaxy form and evolve directly because of the very long timescales for these processes. Due to the finite speed of light, looking at increasingly distant/redshifted galaxies means looking at younger objects. One main task to understand the physics of galaxy evolution is to identify which objects at an early cosmic epoch evolve into which type of galaxies observed today. Again, photometric redshift and similar techniques like the Lyman-break technique can be applied to select galaxies at different epochs. The properties of these samples, e.g. their clustering, can be studied and compared to numerical simulations. By doing so one gets insight into the relationship between the properties of luminous matter in form of galaxies and the underlying structure in form of DM halos. If this is done for several cosmic epochs, the evolution of galaxies can be understood in more detail because the evolution of the halo population is well-known from simulations.

We contribute to the field of galaxy formation and evolution in this thesis by analysing the clustering properties of an unprecedented large sample of galaxies at redshift $z \sim 3$. The selection of these Lyman-break galaxies (LBGs), the simulation of their properties, and the measurement of their two-point correlation function is described in Chap. 6. As a result we obtain estimates for the masses of the halos that host these galaxies, and these masses are compared to estimates at different redshifts from other studies to detect evolutionary trends in the galaxy-DM relationship.

Neither the photometric redshift analyses nor the study of $z \sim 3$ Lyman-break galaxies would be possible without high-quality imaging data from a modern multi-chip CCD camera. The general concepts of the complex processing of the raw data to reach scientifically exploitable images, also called data reduction, is presented in Chap. 2. These techniques are applied to a specific dataset, the optical data of the ESO Deep Public Survey (DPS), which forms the basis of most analyses presented in this thesis.

1

Cosmological framework

This thesis deals with measurements of galaxy properties at large cosmological distances. Over these distances space behaves differently from the well-known Euclidean space we are used to in everyday life. An apparently simple quantity like, for example, distance does not have the same properties as it has in the familiar environment on Earth. Moreover, we now have ample evidence that the universe mainly consists of components which have not been observed directly in laboratories and which have an unknown physical origin. In this chapter the cosmological framework that is necessary for the correct interpretation of the observations presented in later chapters is laid down. An extensive coverage of the topic can be found e.g. in Peacock (1999).

1.1 The homogeneous universe

1.1.1 The field equation

A theory of the universe as a whole must be a theory of gravity because all other fundamental physical forces do not play a role on the large scales considered here. Gravity is described by the theory of General Relativity (Einstein 1916) as being a property of a four-dimensional space-time which is directly influenced by the presence of matter. In general, this connection is described by Einstein's field equation,

$$G_{\mu\nu} = -\frac{8\pi G}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu} \,, \tag{1.1}$$

with $G_{\mu\nu}$ being the Einstein tensor which represents the geometry of space-time, $T_{\mu\nu}$ is the energy-momentum tensor which represents the energy content, Λ is the cosmological constant, and $g_{\mu\nu}$ is the metric tensor.

1.1.2 The cosmological principle

The observed universe, although showing lots of structure on small scales, seems to be homogeneous and isotropic on large scales. It is fundamental to postulate these properties, which are also known as the cosmological principle, to find one particular simple solution for the field equation.

Considering a homogeneously expanding (or contracting) sphere with a spatially uniform density $\rho(t)$, a matter element which is characterised by its position $\boldsymbol{x}$ at a chosen time t_0 will have the position

$$\boldsymbol{r}(t) = a(t)\,\boldsymbol{x} \tag{1.2}$$

at time t. $a(t)$ is the time-dependent scale-factor, normalised to the present value $a(t_0) \equiv a_0 = 1$, and $\boldsymbol{x}$ is the time-independent comoving coordinate. Thus, the world-line $(\boldsymbol{r}, t)$ of a particle in such a homogeneously expanding universe is fully characterised by $\boldsymbol{x}$ and $a(t)$.

Assuming that "fundamental" observers exist, who all experience the same history of the universe and observe the universe as being isotropic, Robertson (1935) and Walker (1936) independently showed that the following metric written in comoving coordinates solves the field equation,

$$\mathrm{d}s^2 = c^2\mathrm{d}t^2 - a^2(t)[\mathrm{d}w^2 + f_K^2(w)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\varphi^2)]\,, \tag{1.3}$$

with w being the radial comoving coordinate and θ and ϕ the angular coordinates. $f_K(w)$ is the comoving angular diameter distance and it relates comoving transverse separations to angular sizes. It is a function of the curvature of space, K,

$$f_K(w) = \begin{cases} \frac{1}{\sqrt{K}}\sin(\sqrt{K}w) & K > 0 \\ w & K = 0 \\ \frac{1}{\sqrt{-K}}\sinh(\sqrt{-K}w) & K < 0 \end{cases} . \tag{1.4}$$

In such a universe, fundamental observer are identical to comoving observers which are characterised by constant spatial coordinates (w, θ, ϕ).

1.1.3 The Friedmann equations

Inserting the metric (1.3) into Eq. (1.1) yields an energy-momentum tensor, $T_{\mu\nu}$, that is of the form of a perfect fluid with density $\rho(t)$ and pressure $p(t)$. The field equation reduces to two differential equations (Friedmann equations), describing the temporal evolution of the scale factor,

$$
\begin{aligned}
H^2(t) = \left(\frac{\dot{a}}{a}\right)^2 &= \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} + \frac{\Lambda}{3}\,, \qquad (1.5)\\
\frac{\ddot{a}}{a} &= -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda}{3}\,. \qquad (1.6)
\end{aligned}
$$

$H(t) = \left(\frac{\dot{a}}{a}\right)$ is called the Hubble parameter and its present value, $H_0 \equiv H(t_0)$, the Hubble constant which is one of the fundamental cosmological parameters representing the expansion rate of the universe. Its is difficult to measure H_0 accurately and this is reflected in the fact that it is usually parametrised in the form $H_0 = 100\frac{\mathrm{km}}{\mathrm{s \cdot Mpc}} \cdot h$, with h expressing our ignorance about the true value. In the following we will assume a concordance value of $h = 0.7$. In Eqs. (1.5) & (1.6) the cosmological constant Λ could also be included in ρ and p since it can be described as a vacuum energy density with negative pressure as will become clear soon.

The two Friedmann equations (1.5) & (1.6) can be combined to yield the adiabatic equation,

$$
\frac{\mathrm{d}\left(a^3\rho c^2\right)}{\mathrm{d}t} = -p\frac{\mathrm{d}a^3}{\mathrm{d}t}\,. \qquad (1.7)
$$

This is nothing but the cosmological representation of the first law of thermodynamics. The temporal change of energy in a comoving volume, $\frac{\mathrm{d}\left(a^3\rho c^2\right)}{\mathrm{d}t}$, equals the expansion or contraction work carried out in this volume, $-p\frac{\mathrm{d}a^3}{\mathrm{d}t}$. In order to derive a theoretical prediction of the temporal evolution of the scale factor one further ingredient is needed, the equation of state (EOS) relating pressure to density for the various components that are contributing to the total energy density. In cosmology one usually considers three different components dominating at different times with the following EOS:

$$\begin{aligned} \text{matter:} \quad & p_\mathrm{m} = 0 && \Rightarrow \rho_\mathrm{m} \propto a^{-3}\,; \\ \text{radiation:} \quad & p_\mathrm{r} = \rho_\mathrm{r} c^2/3 && \Rightarrow \rho_\mathrm{r} \propto a^{-4}\,; \\ \text{vacuum energy:} \quad & \rho_\Lambda = \text{const.} && \Rightarrow p_\Lambda = -\rho_\Lambda\,. \end{aligned} \tag{1.8}$$

Due to the different dependence on scale factor the different components dominate the total density at different cosmic epochs. Therefore, one particular important point is the time of *matter-radiation equality* when matter and radiation had the same density.

At cosmic epochs when only one of these components dominates and under the assumption of vanishing curvature ($K = 0$) Eq. (1.5) can be solved easily,

$$\begin{aligned} \text{matter:} \quad & a \propto t^{2/3}\,, \\ \text{radiation:} \quad & a \propto t^{1/2}\,, \\ \text{vacuum energy:} \quad & a \propto \exp\left(\sqrt{\tfrac{8\pi G}{3}\rho}\, t\right)\,. \end{aligned} \tag{1.9}$$

The densities ρ are often normalised to the critical density ρ_cr at $t = t_0$ which is required to make the universe flat today. For $K = 0$ it follows from Eq. (1.5) that $\rho_\mathrm{cr} = \frac{3H_0^2}{8\pi G}$. The first Friedmann equation in terms of these density parameters $\Omega_i = \frac{\rho_i}{\rho_\mathrm{cr}}$ then reads

$$H^2(t) = H_0^2[a^{-4}\Omega_\mathrm{r} + a^{-3}\Omega_\mathrm{m} + a^{-2}(1-\Omega_0) + \Omega_\Lambda]\,, \tag{1.10}$$

with $\Omega_0 = \Omega_\mathrm{r} + \Omega_\mathrm{m} + \Omega_\Lambda$ being the total density parameter today.

The last decade has seen tremendous progress in the determination of these cosmological density parameters. Different kinds of observations now yield values that agree well within errors and form the basis of the concordance Lambda-Cold-Dark-Matter (ΛCDM) model.

In the following we assume $\Omega_\mathrm{m} = 0.3$ and $\Omega_\Lambda = 0.7$ which are not exactly the best-fit values obtained from recent observations (see e.g. Spergel et al. 2006) but are rather widely used in a lot of studies over the past years. Since we want to compare our findings to these former studies we decided to stick to the values given above.

1.2 Redshift and distances

1.2.1 Cosmological redshift

Due to the expansion of the universe the light emitted by distant sources is redshifted when observed today. The redshift is defined as

$$z = \frac{\lambda_\mathrm{a} - \lambda_\mathrm{e}}{\lambda_\mathrm{e}} = \frac{a(t_\mathrm{a})}{a(t_\mathrm{e})} - 1 \overset{t_\mathrm{a}=t_0}{=} \frac{1}{a} - 1\,, \tag{1.11}$$

with indices 'e' denoting quantities at the emission and 'a' at the absorption of the photons.

For a homogeneously expanding universe the redshift is directly related to a distance. However, distance does not have an unambiguous meaning in cosmology. Rather, different distance measures are used for different applications.

1.2.2 Distance measures

Comoving distance (radial)

The fundamental distance measure in cosmology is the comoving distance, w, which remains constant between two comoving observers. It can be interpreted as the spatial distance between the intersections of the world-lines of the two observers with the spatial hypersurface at $t = t_0$, because for this hypersurface spatial and comoving coordinates coincide due to $a(t_0) = 1$.

For radial light rays ($\mathrm{d}\theta = 0$ and $\mathrm{d}\varphi = 0$) that travel on null geodesics ($\mathrm{d}s^2{=}0$) the metric becomes

$$c\,\mathrm{d}t = -a\,\mathrm{d}w\,.$$

After inserting the definition of the Hubble parameter, integrating over the scale parameter, and transforming coordinates from a to z the comoving distance reads

$$w = \frac{c}{H_0} \int_0^z \frac{\mathrm{d}z'}{E(z')} \,, \tag{1.12}$$

with

$$E(z) = \sqrt{\Omega_\mathrm{r}(1+z)^4 + \Omega_\mathrm{m}(1+z)^3 + (1-\Omega_0)(1+z)^2 + \Omega_\Lambda} \,. \tag{1.13}$$

The limit of w for $z \rightarrow \infty$ is called the horizon and marks the largest comoving distance that can be in causal contact to the observer. The horizon size certainly increases with time and structures or length scales that are larger than the horizon at a given time will *enter* the horizon at later times.

Transverse comoving separation

In the following we are not only interested in radial comoving distances but also in transverse comoving separations to estimate comoving volumes at high redshift. The transverse comoving separation of an object of angular size θ is given by

$$D_\mathrm{trans} = f_K(w)\,\theta \,, \tag{1.14}$$

with $f_K(w)$ being the comoving angular diameter distance from Eq. (1.4). In a flat space-time this reduces to the intuitive geometric relation $D_\mathrm{trans} = w\,\mathrm{d}\theta$. Comoving volumes are estimated in general then by integrating over the radial comoving distance and the transverse comoving separation.

Angular diameter distance

The angular diameter distance that relates the physical (instead of the comoving) size of an object to its angular size on the sky is given by multiplying D_trans by the scale factor (see Eq. 1.2),

$$D_\mathrm{ang} = a\,D_\mathrm{trans} = \frac{D_\mathrm{trans}}{1+z} \,. \tag{1.15}$$

An interesting fact about D_{ang} is that it does not increase to infinity with redshift. It has a maximum, e.g. at $z \sim 1.5$ for ΛCDM, and then decreases again. Thus, objects of the same physical size appear larger to the observer if they are located at $z = 3$ compared to $z = 1$.

Luminosity distance

In Euclidean space the flux of an object decreases with the inverse square of the distance. The distances defined above do not obey this rule because space expands during the travel time of the photons and the flux is further diluted. It can be shown (Etherington 1933) that the luminosity distance, D_{lum}, obeys the inverse square law in general and that it is related to the other distance measures in the following way,

$$D_{\text{lum}} = \frac{D_{\text{ang}}}{a^2} = \frac{D_{\text{trans}}}{a} \,. \tag{1.16}$$

This equation is correct for bolometric quantities. For fluxes at specific wavelengths or frequencies a K-correction that depends on the spectrum of the source must be applied to account for the cosmological redshift.

Distance modulus

Since optical astronomers usually do not use fluxes but magnitudes, another common quantity, the distance modulus DM, is defined by

$$DM = 5 \log \frac{D_{\text{lum}}}{10\text{pc}} \,. \tag{1.17}$$

1.3 Structure formation in the universe

The homogeneous universe described in Sect. 1.1 is a good approximation for very large scales ($\gtrsim 200$Mpc). On smaller scales, however, a wealth of structures in the form of galaxy clusters, voids, and filaments has developed. The following section deals with the theoretical description of the formation of these cosmological structures.

1.3.1 Linear perturbation theory

The Robertson-Walker metric (1.3) inserted into Einstein's field equation (1.1) yields an energy-momentum tensor that is of the form of a perfect fluid as described above. If we now consider length-scales which are well below the horizon size, the Newtonian description of gravity is a valid approximation. In the case of pressure-less matter (dust) the behaviour of the fluid is described by the following three equations:

$$\frac{\partial \rho}{\partial t} + \nabla_r \cdot [\rho \boldsymbol{u}] = 0 \qquad \text{Continuity equation,} \tag{1.18}$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + (\boldsymbol{u} \cdot \nabla_r)\boldsymbol{u} = -\nabla_r \phi \qquad \text{Euler equation,} \tag{1.19}$$

$$\nabla_r^2 \phi = 4\pi G \rho - \Lambda \qquad \text{Poisson equation.} \tag{1.20}$$

The case of a homogeneous and isotropic universe is a solution to these equations. The Poisson equation was modified with the Λ term although it is not theoretically motivated in the Newtonian description. This modification was done in such a way that the solution with a homogeneously and isotropically expanding universe [$\rho(\boldsymbol{r}, t) = \bar{\rho}(t)$ and $\boldsymbol{u}(\boldsymbol{r}, t) = \frac{\dot{a}}{a}\boldsymbol{r}$] agrees with the Friedmann equations (1.5 & 1.6).

In the following we are interested in small deviations from this trivial case which can be expressed by density contrasts $\delta(\boldsymbol{r}, t) = \frac{\rho(\boldsymbol{r},t)-\bar{\rho}(t)}{\bar{\rho}(t)}$ and peculiar velocities $\boldsymbol{v} = \boldsymbol{u} - (\dot{a}/a)\boldsymbol{r}$. The fluid equations perturbed to first order in density and velocity, and transformed to comoving coordinates [$\boldsymbol{x} = \boldsymbol{r}/a(t)$] then become

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\nabla_x \cdot \boldsymbol{v} = 0 \qquad \text{Continuity equation,} \tag{1.21}$$

$$\frac{\partial \boldsymbol{v}}{\partial t} + \frac{\dot{a}}{a}\boldsymbol{v} = -\frac{1}{a}\nabla_x \Phi \qquad \text{Euler equation,} \tag{1.22}$$

$$\nabla_x^2 \Phi = \frac{3H_0^2\Omega_{\mathrm{m}}}{2a}\delta \qquad \text{Poisson equation,} \tag{1.23}$$

with $\Phi = \phi(a\boldsymbol{x}, t)) + (\ddot{a}a/2)|\boldsymbol{x}|^2$ being the comoving potential. This result is not explicitly dependent on Λ confirming the intuitive guess that a uniform energy background does not influence structure formation directly. However, the cosmological constant Λ does have an influence on the temporal evolution of the scale factor, a, and therefore on the evolution of a density contrast δ with redshift.

Combining different derivatives of (1.21)-(1.23) yields a homogeneous, second-order differential equation for δ independent of $\boldsymbol{v}$ that describes the temporal behaviour of density perturbations in the linear regime,

$$\ddot{\delta} + \frac{2\dot{a}}{a}\dot{\delta} = \frac{3H_0^2\Omega_\mathrm{m}}{2a^3}\delta \,. \tag{1.24}$$

Separating the spatial dependence we get a solution of the following form:

$$\delta(\boldsymbol{x},t) = D_+(t)\Delta_+(\boldsymbol{x}) + D_-(t)\Delta_-(\boldsymbol{x}) \,. \tag{1.25}$$

The terms D_+ and D_- represent a growing and a decaying mode, respectively. The mode D_-, decaying with increasing scale-factor, will have died out at low redshift and plays no role anymore. Thus,

$$\delta(\boldsymbol{x},t) = D_+(t)\delta_0(\boldsymbol{x}) \tag{1.26}$$

with $\delta_0(\boldsymbol{x}) = \delta(\boldsymbol{x},t_0)$ and $D_+(t_0) = 1$. The actual form of the growing mode D_+, which is also called growth factor, is dependent on the cosmological density parameters and can be calculated explicitly for different cosmologies. For example, in the Einstein-de Sitter (EdS) case with $\Omega_\mathrm{m} = 1$ and $\Omega_\Lambda = 0$ it equals the scale-factor. If $\Omega_\mathrm{m} < 1$ then $D_+(t) > a(t)$.

1.3.2 Description of cosmological density fields

Cosmology cannot be aimed at predicting the matter structures in our universe exactly. That would require precise knowledge about the complete density field at some earlier epoch. Rather, cosmology should produce predictions about the statistical properties of this density field; in the following it will be described how these statistical properties can be quantified.

Random fields

The density fluctuations in the universe, $\delta(\boldsymbol{x}, t)$, are considered to be the realisation of a random field. Such a random field represents the probability distribution for the occurrence of such a specific realisation, $\delta(\boldsymbol{x}, t)$. Usually, the properties of a random field are inferred from the ensemble average over many realisations. Lacking an ensemble of universes it is commonly assumed that the ensemble average can be replaced by a volume average for sufficiently large volumes, i.e. it is assumed that different well-separated patches of the universe can be regarded as different realisations of the underlying random field.

The correlation function

The two-point correlation function of a density field $\delta(\boldsymbol{x})$ with $\langle\delta\rangle = 0$ and which is homogeneous and isotropic on large scales is defined by

$$\xi(|\boldsymbol{x} - \boldsymbol{y}|) = \xi(r) = \langle\delta(\boldsymbol{x})\delta^*(\boldsymbol{y})\rangle \,, \tag{1.27}$$

with $\langle ... \rangle$ denoting an ensemble average (or volume average, see above). The correlation function ξ of an arbitrary density field could depend on the vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ explicitly. A homogeneous density field, however, can only depend on $\boldsymbol{x} - \boldsymbol{y}$, and the additional assumption of isotropy reduces this to a dependence on $r = |\boldsymbol{x} - \boldsymbol{y}|$.

The power spectrum

It can be useful to quantify the density field in Fourier space,

$$\delta(\boldsymbol{x}, t) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 k}{(2\pi)^3} \tilde{\delta}(\boldsymbol{k}, t) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \,, \tag{1.28}$$

because it can be shown that the different Fourier modes $\tilde{\delta}$ evolve mutually independently in the linear regime. From (1.26) it follows that

$$\tilde{\delta}(\boldsymbol{k}, t) = D_+(t)\tilde{\delta}_0(\boldsymbol{k}) \,. \tag{1.29}$$

One defines the power spectrum as the Fourier transform of the two-point correlation function,

$$P(|\boldsymbol{k}|) = \int_{\mathbb{R}^3} \mathrm{d}^3 x \, \mathrm{e}^{-\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \xi(|\boldsymbol{x}|) \,, \tag{1.30}$$

which obeys the following relation

$$\langle \tilde{\delta}(\boldsymbol{k}) \tilde{\delta}^*(\boldsymbol{k}') \rangle = (2\pi)^3 \delta_{\mathrm{D}}(\boldsymbol{k} - \boldsymbol{k}') P(|\boldsymbol{k}|) \,, \tag{1.31}$$

with δ_{D} being the three-dimensional Dirac delta function.

Gaussian random fields

Equation (1.28) can be discretised by considering a finite volume (e.g. a cube with side-length L) at fixed time, t,

$$\begin{aligned} \delta(\boldsymbol{x}) &= \int_{\mathbb{R}^3} \frac{\mathrm{d}^3 k}{(2\pi)^3} \tilde{\delta}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \\ &\approx \left(\frac{\Delta k}{2\pi}\right)^3 \sum_{\boldsymbol{k}} \tilde{\delta}(\boldsymbol{k}) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}} \\ &= \sum_{\boldsymbol{k}} \delta_{\boldsymbol{k}} \mathrm{e}^{\mathrm{i}\boldsymbol{x}\cdot\boldsymbol{k}} \,, \end{aligned} \tag{1.32}$$

with $\Delta k = 2\pi/L$ and $\delta_{\boldsymbol{k}} = (\Delta k/2\pi)^3 \tilde{\delta}(\boldsymbol{k})$.

A special class of random fields which is important in cosmology is characterised by the two following properties:

1. The Fourier components $\delta_{\boldsymbol{k}}$ are mutually statistically independent.
2. The probability distribution of one particular Fourier component $\delta_{\boldsymbol{k}}$ is Gaussian.

The second property follows from the first through the central limit theorem. This can be understood by refining the discretisation in (1.32) by a factor n and expressing one particular Fourier component

$\delta_{\boldsymbol{k}}$ as the sum of the n^3 Fourier components on the finer grid. The central limit theorem then states that the probability density of such a sum of independent variables approaches a Gaussian if the mean and variance of the probability densities of the independent variables are finite. A random field that has these properties is called a *Gaussian random field.* It can be shown that such a Gaussian random field is completely characterised by its correlation function ξ or, equivalently, by its power spectrum P.

1.3.3 The dark matter density field

Today, we have good evidence for the existence of dark matter. The dominant kind of matter is of non-baryonic form and cosmic structure formation is driven by this component.

The primordial power spectrum

Inflationary theories of the very early universe suggest that the primordial spectrum of density fluctuations can be described as the realisation of a Gaussian random field. Since different Fourier modes evolve independently in the linear regime, this Gaussianity is preserved until non-linear structures form and different modes couple. Thus, the primordial density field is fully characterised by its power spectrum. But how does this primordial power spectrum $P_0(k)$ look like?

At early cosmic epochs all length scales of interest are much larger than the horizon so that no particular length scale is preferred. This suggests that $P_0(k)$ has the form of a power law which is the only mathematical function that does not depend on a characteristic scale,

$$P_0(k) \propto k^{n_\mathrm{S}} \,. \tag{1.33}$$

If one further requires that fluctuations entering the horizon all have same amplitude (*scale-invariant*) then the spectral index is constrained to $n_\mathrm{S} = 1$ (Harrison 1970; Zeldovich 1972). A spectral index very close to, but slightly smaller than $n_\mathrm{S} = 1$ is also predicted by inflationary models.

The actual amplitude A of the primordial power spectrum cannot be deduced from theory but must be determined from observations,

$$P_0(k) = Ak^{n_\mathrm{s}} \,. \tag{1.34}$$

The time evolution follows from, (1.29) & (1.31)

$$P(k,t) = D_+^2(t)Ak^{n_\mathrm{s}} \,. \tag{1.35}$$

The transfer function

Equation (1.34) only holds in the matter-dominated epoch. For the derivation of D_+ via the fluid equations, pressure was neglected (see Sect. 1.3.1). In the radiation-dominated epoch this assumption is no longer valid. Furthermore, there are always Fourier modes in the density field which are larger than the horizon and enter the horizon at a later time. These super-horizon perturbations cannot be described by Newtonian gravity as was done above but they require a full general relativistic treatment. These deviations from the pressure-less Newtonian case can be accounted for in the following way:

$$P_0(k) = Ak^{n_\mathrm{s}}T^2(k) \,, \tag{1.36}$$

with $T(k)$ being the transfer function. The actual shape of the transfer function depends on the nature of the dark matter particles. Hot dark matter (HDM) is characterised by relativistic particles at the time of matter-radiation equality. It would lead to a transfer function with small-scale perturbations being suppressed by the free streaming of the relativistic particles. In this case large structures would form first and later these would fragment into smaller units (top-down). Several observations rule out that HDM is the dominating dark matter component.

Cold dark matter (CDM) scenarios yield a much better agreement with observations. In the following we will use the fitting formula for the CDM transfer function by Bardeen et al. (1986),

$$T(q) = \frac{\ln(1+2.34q)}{2.34q}\left[1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4\right]^{-1/4} \,, \tag{1.37}$$

with $q = \frac{k}{\Gamma}\frac{\mathrm{Mpc}}{h}$ and $\Gamma = \Omega_{\mathrm{m}}\, h$ being the shape parameter. Γ depends also weakly on the baryon density Ω_{b} but we will neglect this small effect here. This CDM power spectrum results in small-scale structures forming first, later merging to larger structures (bottom-up, hierarchical scenario).

Nonlinear evolution

When the density contrast of a particular Fourier mode approaches unity, linear theory breaks down. In order to study the evolution of these non-linear structures further, which are the seeds for galaxies and clusters, large numerical simulations are carried out. Starting from a Gaussian random field with the properties described above (scale-independent power spectrum) one realisation is evolved in time on the computer. The resulting dark matter structures can be investigated to find fitting formulae for the non-linear power spectrum as it was done by e.g. Peacock & Dodds (1996) and Smith et al. (2003).

1.3.4 The spherical collapse model

In contrast to numerical simulations, some properties of non-linear structures in the density field can be described by semi-analytical models that give some additional insight into the processes involved.

A better understanding of the formation of dark matter halos and a rough estimate of their density, size, etc. can be obtained by considering a spherical overdensity with $\rho(t) = [1+\delta(t)]\bar{\rho}(t)$, with $\bar{\rho}(t)$ being the average density of the universe. Though not very realistic, this model can be solved analytically and gives fair estimates of the overdensity and size of e.g. galaxy clusters.

Due to symmetry reasons this over-dense sphere evolves like a separate Friedmann universe, however with decreased expansion rate. Similar to the model of the homogeneous universe a critical density exist inside the over-dense sphere, and for densities larger than this local critical density the expansion will come to a halt and the sphere will re-collapse. Small-scale perturbations and imperfect radial orbits of the particles will lead to a collapse that does not end up in a single point but in an over-dense, virialised structure with a radius which is half as large as the radius of maximum expansion of the sphere. Such a collapsed, virialised structure is called a "halo" and it is characterised by its virial radius, r_{vir}. The value for the overdensity inside the sphere with radius $r = r_{\mathrm{vir}}$ can be calculated theoretically and is approximately $1 + \delta_{\mathrm{vir}} \approx 178\Omega_{\mathrm{m}}^{-0.7}$. If one uses linear perturbation theory, regardless of the fact that it has long before ceased to be valid, and sets $\Omega_{\mathrm{m}} = \Omega_0 = 1$ (EdS) it yields a value of $\delta_{\mathrm{lin,c}} \approx 1.69$ for a perturbation that will have re-collapsed at $t = t_0$. For larger redshifts re-collapse will occur if

$$\delta > \delta_{\mathrm{lin,c}}(1 + z)\,. \tag{1.38}$$

Thus, all matter concentrations with a linearly extrapolated density contrast of $\delta(t_0) \geq 1.69$ will have collapsed today and form virialised structures with a true density contrast of $\delta \approx 177$. For convenience one usually considers a sphere with an overdensity of $\delta = 200$. The corresponding radius of this sphere called r_{200} is often used to measure e.g. the size of a galaxy cluster.

1.3.5 The abundance of dark matter halos

Given the results of the spherical collapse model and taking the initial Gaussian density field as a basis one can calculate the number density of collapsed objects of a given mass at a given cosmic epoch. This was first done by Press & Schechter (1974) and the resulting mass function was widely used until only recently more accurate fitting formulae to numerical simulations became available.

The argument runs as follows. It is assumed that all structures that at redshift z have reached a linearly extrapolated density contrast of $\delta > \delta_{\mathrm{lin,c}}(1 + z)$ can be considered as collapsed halos. The initial density field at some early time $t = t_{\mathrm{i}}$ can be smoothed with a filter W_R of characteristic size R. The variance of the smoothed density field then becomes

$$\sigma^2(R, t_{\mathrm{i}}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \left|\tilde{W}_R(k)\right|^2 P(k, t_{\mathrm{i}})\,, \tag{1.39}$$

with $\tilde{W}_R$ being the Fourier transform of the filter function. For a Gaussian initial density field δ the smoothed version δ_R is also Gaussian. The probability for a point in space to be located inside a halo of characteristic mass $M_R = \frac{4}{3}\pi R^3 \bar{\rho}$ or larger and with a density contrast of $\delta > \delta_{\mathrm{c}}$ reads

$$\begin{aligned} P(\delta_{\mathrm{c}}, t_{\mathrm{i}}, R) &= \int\limits_{\delta_{\mathrm{c}}}^{\infty} \frac{\mathrm{d}\delta'}{\sqrt{2\pi}\sigma(R, t_{\mathrm{i}})} \mathrm{e}^{-\frac{\delta'^2}{2\sigma^2(R,t_{\mathrm{i}})}} && (1.40) \\ &= \frac{1}{\sqrt{\pi}} \int\limits_{\frac{1}{\sqrt{2}}\delta_{\mathrm{c}}/\sigma(R,t_{\mathrm{i}})}^{\infty} \mathrm{d}x\, \mathrm{e}^{-x^2} \\ &= \frac{1}{2} \operatorname{erfc}\left(\frac{1}{\sqrt{2}} \frac{\delta_{\mathrm{c}}}{\sigma(R, t_{\mathrm{i}})} \right), && (1.41) \end{aligned}$$

with $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int\limits_{x}^{\infty} \mathrm{d}x' \mathrm{e}^{-x'^2}$ being the complementary Gauss error function.

The required density contrast at time t_{i} for collapse before t reads

$$\delta_{\mathrm{c}}(t, t_{\mathrm{i}}) = D_+(t_{\mathrm{i}}) \delta_{\mathrm{c}0}(t)\,. \tag{1.42}$$

Hence, we define $P(M, t) = P[\delta_{\mathrm{c}}(t, t_{\mathrm{i}}), t_{\mathrm{i}}, R]$ representing the fraction of volume in collapsed halos of mass $M \geq M_R$ at time t. Differentiation of this expression with respect to M yields the fraction of volume in collapsed objects with a mass in a small interval $\mathrm{d}M$ around M,

$$\begin{aligned} p(M, t) &= -\left(\frac{\partial P}{\partial M} \right) \mathrm{d}M = -\left(\frac{\partial P}{\partial \sigma} \frac{\partial \sigma}{\partial M} \right) \mathrm{d}M \\ &= -\left(\frac{1}{\sqrt{2\pi}} \frac{\delta_{\mathrm{c}}}{\sigma^2(R, t_{\mathrm{i}})} \mathrm{e}^{-\frac{\delta_{\mathrm{c}}^2}{2\sigma^2(R,t_{\mathrm{i}})}} \frac{\partial \sigma}{\partial M} \right) \mathrm{d}M\,, && (1.43) \end{aligned}$$

where we used the definition of the complementary Gauss error function in the last step to get $\frac{\mathrm{d}\operatorname{erfc}(x)}{\mathrm{d}x} = -\frac{2}{\sqrt{\pi}} \mathrm{e}^{-x^2}$.

The total mass inside a given volume V is $V\Omega_{\mathrm{m}}\rho_{\mathrm{cr}}$. We define the mass function $\frac{\mathrm{d}n}{\mathrm{d}M}$ representing the differential comoving number density of halos of mass M at time t in the following way:

$$\frac{\mathrm{d}n}{\mathrm{d}M}\mathrm{d}M\, VM = V\Omega_{\mathrm{m}}\rho_{\mathrm{cr}}\left(\frac{\partial P}{\partial M}\right)\mathrm{d}M\,. \tag{1.44}$$

Inserting the result from (1.43) we get

$$\begin{aligned}\frac{\mathrm{d}n}{\mathrm{d}M} &= -\frac{\Omega_{\mathrm{m}}\rho_{\mathrm{cr}}}{\sqrt{2\pi}M}\,\frac{\delta_{\mathrm{c}0}(t)}{\sigma^2(R)}\,\frac{\mathrm{d}\sigma(R)}{\mathrm{d}M}\mathrm{e}^{-\frac{\delta_{\mathrm{c}0}^2(t)}{2\sigma^2(M)}} \\ &= -\frac{\Omega_{\mathrm{m}}\rho_{\mathrm{cr}}}{\sqrt{2\pi}M^2}\,\frac{\delta_{\mathrm{c}0}(t)}{\sigma(R)}\,\frac{\mathrm{d}\ln\sigma(R)}{\mathrm{d}\ln M}\mathrm{e}^{-\frac{\delta_{\mathrm{c}0}^2(t)}{2\sigma^2(M)}}\,. \end{aligned}\tag{1.45}$$

The normalisation of the mass function is, however, not correct. In the derivation above, the initially under-dense regions are neglected and Press & Schechter (1974) argue that this can be accounted for by multiplying the right-hand side of (1.45) by a factor of 2.

We define the quantity $f(\sigma)$ as

$$f(\sigma) = \frac{M}{\rho_0}\frac{\mathrm{d}n}{\mathrm{d}\ln\sigma^{-1}} = -\frac{M}{\rho_0}M\frac{\mathrm{d}\ln M}{\mathrm{d}\ln\sigma}\frac{\mathrm{d}n}{\mathrm{d}M} = \sqrt{\frac{2}{\pi}}\frac{\delta_{\mathrm{c}}}{\sigma}\mathrm{e}^{-\frac{\delta_{\mathrm{c}}^2}{2\sigma^2}} \tag{1.46}$$

As mentioned above, today more accurate formula for the dark matter halo mass function exist derived from empirical fitting of large N-body simulation data. In the following we will use the one by Sheth & Tormen (1999),

$$f(\sigma) = A\sqrt{\frac{2a}{\pi}}\left[1+\left(\frac{\sigma^2}{a\delta_{\mathrm{c}}^2}\right)^p\right]\frac{\delta_{\mathrm{c}}}{\sigma}\mathrm{e}^{-\frac{a\delta_{\mathrm{c}}^2}{2\sigma^2}}\,, \tag{1.47}$$

with the numerical constants $A = 0.3222$, $a = 0.707$, and $p = 0.3$ determined from a fit to simulations by the GIF/Virgo collaboration (Kauffmann et al. 1999) for different cosmologies.

1.3.6 Halo biasing

Halos of a given mass can be more or less strongly clustered than the underlying mass distribution. This *biasing* is expressed by the mass-dependent halo bias factor $b(M)$. Let $P_{\mathrm{halo}}(M)$ denote the power-spectrum of halos of mass M (computed from the positions of the halos via the correlation function). Then the bias factor is defined by $b^2(M) = P_{\mathrm{halo}}(M)/P_{\mathrm{matter}}$. In general, the bias factor is also scale dependent which is clearly shown by simulations. However, on large scales a scale independent bias factor is a good approximation. For the halo bias factor we again adopt the fitting formula by Sheth & Tormen (1999),

$$b(M, z) = 1 + \frac{a\nu - 1}{\delta_{\mathrm{c}}(z)} + \frac{2p/\delta_{\mathrm{c}}(z)}{1 + (a\nu)^p}\,, \tag{1.48}$$

with $\nu = \left(\frac{\delta_{\mathrm{c}}(z)}{\sigma(M)}\right)^2$.

1.3.7 The halo mass profile

From different numerical simulations it is found that dark matter halos follow a universal radial density profile regardless of their mass,

$$\rho(r) = \frac{\rho_{\mathrm{s}}}{(r/r_{\mathrm{s}})^{\alpha}(1 + r/r_{\mathrm{s}})^{3-\alpha}}\,, \tag{1.49}$$

with r_{s} and ρ_{s} being a characteristic radius and a characteristic density depending on the halo mass and redshift, and α being a constant derived from the particular simulation. We use the classical value of $\alpha = 1$ by Navarro et al. (1996, 1997).

1.4 Galaxy clustering

In the preceding sections the structure formation of the dominating dark matter component was discussed. As its name suggests this component is not observable directly. Rather, the properties of the dark matter must be inferred from observations of visible matter in the form of galaxies. One powerful tool to study the underlying dark matter distribution is the clustering of galaxies. However, unlike other methods like e.g. cosmic shear which probe the matter directly regardless of its nature, the conclusions from galaxy clustering measurements rely on several assumptions about the relation between galaxies and the underlying dark matter halos. In the following a model is described which predicts the clustering properties of galaxies on the basis of the theoretical model of dark matter clustering presented in Sect. 1.3.1-1.3.4.

1.4.1 The halo model

It turns out to be useful to regard the structures in the universe as being made up of distinct halos of different sizes and masses. In order to get a complete picture of massive structures three ingredients are needed:

1. A model for the distribution of halo masses (mass function)
2. A description of the spatial distribution of halos (power spectrum)
3. The mass profile of halos

These three ingredients were described in the preceding sections. We apply the mass function by Sheth & Tormen (1999), the linear power spectrum given by (1.36) (incorporating the transfer function by Bardeen et al. 1986), and the NFW mass profile by Navarro et al. (1996, 1997).

The halo occupation distribution

Unlike the dark matter field, which is usually treated as a continuous quantity, galaxies are discrete units and are considered as points in clustering studies. It is therefore necessary to relate the properties of a dark matter halo, such as its mass or its angular momentum, to the number of galaxies that are hosted by this halo. Models of galaxy formation and evolution are quite complicated including gas physics, star-formation, and feedback processes, and they require such a large number of parameters to be adjusted that to date they do not deliver accurate answers to these questions. Thus, simple toy models are constructed to relate the clustering of dark matter to the clustering of galaxies and the validity of these toy models is tested by observations. In the following we adopt the halo model described in Hamana et al. (2004). A more complete coverage of this topic can be found in the extensive review by Cooray & Sheth (2002).

All galaxies are assumed to be hosted by dark matter halos since according to current understanding only in these places the environment is such that baryons can cool. There are no galaxies formed outside dark matter halos but there may be dark matter halos that do not form a galaxy because their mass is to low. Therefore, we introduce a lower cut-off mass, M_{min}, representing the minimal mass required for a halo to host a galaxy. Furthermore, since halos can host sub-halos of lower mass it is reasonable to assume that massive halos can also host multiple galaxies. We choose a simple power-law behaviour for the number of galaxies scaling with mass described by the slope α and the normalisation mass M_1.

The above considerations lead to the following recipe for the halo-occupation-distribution (HOD) $\langle N_{\mathrm{g}}|M\rangle$, describing the mean number of galaxies in a halo of mass M,

$$\langle N_{\mathrm{g}}|M\rangle = \begin{cases} (M/M_1)^{\alpha} & \text{for } M > M_{\mathrm{min}} \\ 0 & \text{for } M < M_{\mathrm{min}} \end{cases} , \tag{1.50}$$

with M_{min}, M_1, and α being the three parameters of the model.

The mean number of pairs of galaxies, $\langle N_{\mathrm{g}}(N_{\mathrm{g}}-1)|M\rangle$, does not only depend on the mean number of galaxies in a halo but also on the probability distribution around this mean, $p(N_{\mathrm{g}}|M)$. Simulations show, that a Poisson distribution with $\langle N_{\mathrm{g}}(N_{\mathrm{g}}-1)|M\rangle = \langle N_{\mathrm{g}}|M\rangle^2$ is a good approximation for large masses associated with $\langle N_{\mathrm{g}}|M\rangle > 1$. For smaller masses and smaller values of $\langle N_{\mathrm{g}}|M\rangle$ the Poisson distribution is not adequate anymore because it allows for an arbitrary large number of galaxies due to its tail. A binomial distribution leads to a much better fit to simulations with its scatter parametrised by $\langle N_{\mathrm{g}}(N_{\mathrm{g}}-1)|M\rangle = \langle N_{\mathrm{g}}|M\rangle^2 \log\left[4\langle N_{\mathrm{g}}|M\rangle\right]/\log 4$. Thus, we apply the following HOD for pairs of galaxies:

$$\langle N_\mathrm{g}(N_\mathrm{g}-1)|M\rangle \\ = \begin{cases} \langle N_\mathrm{g}|M\rangle^2 & \text{if } \langle N_\mathrm{g}|M\rangle > 1 \\ \langle N_\mathrm{g}|M\rangle^2 \log\left[4\,\langle N_\mathrm{g}|M\rangle\right]/\log 4 & \text{if } 1 > \langle N_\mathrm{g}|M\rangle > 0.25 \\ 0 & \text{otherwise} \end{cases} . \tag{1.51}$$

Galaxy number density

It is now easy to calculate the comoving number density of galaxies at redshift z,

$$n_\mathrm{g}(z) = \int\limits_{M_\mathrm{min}}^{\infty} \mathrm{d}M\, n_\mathrm{halo}(M,z)\ \langle N_\mathrm{g}|M\rangle\ , \tag{1.52}$$

where $n_\mathrm{halo}(M,z)$ represents the mass function by Sheth & Tormen (1999) written now in dependence on redshift z instead of time t.

Galaxy power spectrum

The galaxy power spectrum can be predicted by summing up the contributions from galaxies in different halos (2-halo term) and from galaxy pairs in the same halo (1-halo term), the former being dominant on large and the latter on small scales,

$$P_\mathrm{g}(k) = P_\mathrm{g}^\mathrm{1h}(k) + P_\mathrm{g}^\mathrm{2h}(k)\,. \tag{1.53}$$

In the following we assume that if a halo hosts exactly one galaxy it is located in the centre of the halo and if a halo hosts more than one galaxy, one of them is located in the centre and the others are distributed (statistically) according to the halo mass profile.

The 2-halo term can be calculated from the linear CDM power spectrum,

$$P_{\mathrm{g}}^{2\mathrm{h}}(k) = P_{\mathrm{lin}}(k) \left[\frac{1}{n_{\mathrm{g}}} \int \mathrm{d}M \, n_{\mathrm{halo}}(M) \, \langle N_{\mathrm{g}} | M \rangle \, b(M) \, y(k, M) \right]^2 , \tag{1.54}$$

with $b(M)$ the halo bias factor described above and $y(k, M)$ the normalised Fourier transform of the halo density profile,

$$y(k, M) = \frac{1}{M} \int 4\pi r^2 \mathrm{d}r \, \rho(r, M) \frac{\sin(kr)}{kr} . \tag{1.55}$$

Using the linear power spectrum here under-predicts the power on intermediate and small scales. Fortunately, the small-scale behaviour is completely determined by the 1-halo term. Thus, an inaccurate 2-halo term on small scales has no effect on the overall accuracy of the galaxy power spectrum. On very large scales where the spatial extent of halos is negligible the 2-halo term equals the linear power spectrum times the square of the bias factor.

Similarly, the 1-halo term can be written as

$$P_{\mathrm{g}}^{1\mathrm{h}}(k) = \frac{1}{(2\pi)^3 \, n_{\mathrm{g}}^2} \int \mathrm{d}M \, n_{\mathrm{halo}}(M) \, \langle N_{\mathrm{g}}(N_{\mathrm{g}} - 1) | M \rangle \, |y(k, M)|^p . \tag{1.56}$$

The value for the parameter p depends on the assumption how the galaxies are spatially distributed in the halo. It is natural to assume that one galaxy is located at the centre of the halo. All other galaxies are assumed to be distributed corresponding to the halo density profile. If there is approximately one pair in the halo, the term $y(k, M)$ occurs only once (for the satellite and not for the central galaxy which contributes a factor 1) and hence $p = 1$. In the limit of many pairs the fact that the central galaxy sits at the centre can be neglected and it is assumed that each galaxy contributes a factor $y(k, M)$ and hence $p = 2$ which we adopt for $\langle N_{\mathrm{g}}(N_{\mathrm{g}} - 1) | M \rangle > 1$.

The galaxy two-point correlation function, $\xi_{\mathrm{g}}(r)$, can be obtained via Fourier transformation of $P_{\mathrm{g}}(k)$.

1.4.2 Projection onto the sky

Until here we have only dealt with the correlation function in 3-dimensional space and with the power spectrum in 3-dimensional Fourier space. From astronomical observations it may be difficult to measure e.g. $\xi(r)$ of galaxies directly because this would require precise 3D information about every single galaxy. While in large spectroscopic redshift surveys like the Sloan Digital Sky Survey (SDSS, York et al. 2000) and the 2dF Galaxy Redshift Survey (2dFGRS, Colless et al. 2001) the correlation function can be measured in redshift space directly (which is still influenced by distortion effects originating from peculiar velocities when compared to real space) for many applications only angular positions on the sky are available. However, if, instead of precise individual redshifts, knowledge about the overall redshift distribution of the galaxies is available the correlation function can still be estimated.

The angular correlation function $\omega(\theta)$ can be measured from the angular positions of galaxies on the sky. It is defined in analogy to the real-space correlation function by

$$\omega(\theta) = \langle N(\boldsymbol{\theta}') \, N(\boldsymbol{\theta} + \boldsymbol{\theta}') \rangle \,, \tag{1.57}$$

with $N(\boldsymbol{\theta})$ being the galaxy over-/underdensity at position $\boldsymbol{\theta}$, and $\theta = |\boldsymbol{\theta}|$. ω can be related to the galaxy number density contrast δn in real-space by the following projection integral:

$$\begin{aligned} \omega(\theta) \;\; = \;\; & \int_0^\infty \mathrm{d}w \, p(w) \int_0^\infty \mathrm{d}w' \, p(w') \\ & \times \langle \delta n(f_K(w)\boldsymbol{\theta}', w) \, \delta n(f_K(w')(\boldsymbol{\theta} + \boldsymbol{\theta}'), w') \,, \end{aligned} \tag{1.58}$$

where $p(w)$ represents the galaxy selection function, i.e. the probability distribution in comoving distance. Assuming that the real-space correlation function does not evolve over the look-back time interval where $p(w)$ is non-zero, this reduces to

$$\omega(\theta) = \int_0^\infty \mathrm{d}\bar{w} \int_{-\infty}^\infty \mathrm{d}\Delta w \, p\left(\bar{w} + \frac{\Delta w}{2}\right) p\left(\bar{w} - \frac{\Delta w}{2}\right) \times \xi_{\mathrm{g}} \left(\sqrt{f_K^2(\bar{w})\theta^2 + \Delta w^2}, \bar{w} \right) , \tag{1.59}$$

with $\bar{w} := \frac{w+w'}{2}$ and $\Delta w := w - w'$, and the second argument of ξ_{g} specifying the time at which the real-space correlation function is observed.

Another important assumption is needed to reach an even simpler relation between ω and ξ_{g} that can be solved analytically. If the weight function p is changing slowly on scales of the size of the structures which are described by ξ_{g} one can approximate $p\left(\bar{w} + \frac{\Delta w}{2}\right) p\left(\bar{w} - \frac{\Delta w}{2}\right) \approx p^2(\bar{w})$. This leads to the famous relativistic Limber equation (Limber 1953; Peebles 1980),

$$\omega(\theta) = \int_0^\infty \mathrm{d}\bar{w} \, p^2(\bar{w}) \int_{-\infty}^\infty \mathrm{d}\Delta w \, \xi \left(\sqrt{f_K^2(\bar{w})\theta^2 + \Delta w^2}, \bar{w} \right) , \tag{1.60}$$

which is widely used to relate the observed angular correlation function to the real-space correlation function if the redshift selection function p is known. It should be noted that the assumptions which lead to Limber's equation are non-trivial. It turns out that the crucial parameter determining its accuracy is $y = \sigma / r_{\mathrm{m}}$, with σ representing the width of p and r_{m} representing the mean of p. The accuracy of Limber's equation decreases with decreasing y and increasing θ. For quantitative estimates of the expected accuracy see Simon (2006).

1.4.3 Shape of the galaxy correlation function

An interesting result of Limber's equation is that if ξ is a power law, $\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma}$, the second integral can be solved analytically and it turns out that ω is also a power law,

$$\omega(\theta) = \theta^{1-\gamma} \, r_0^\gamma \cdot \frac{\Gamma(1/2)\,\Gamma(\gamma/2 - 1/2)}{\Gamma(\gamma/2)} \int_0^\infty \mathrm{d}\bar{w} \, p^2(\bar{w}) \, [f_K(\bar{w})]^{1-\gamma} , \tag{1.61}$$

with Γ being the Euler Gamma function.

Hence, the slope δ of ω is related to the slope γ of ξ through $-\delta = 1 - \gamma$ and the amplitude

$$A_\theta = r_0^\gamma \cdot \frac{\Gamma(1/2)\,\Gamma(\gamma/2 - 1/2)}{\Gamma(\gamma/2)} \int_0^\infty d\bar{w}\, p^2(\bar{w})\, [f_K(\bar{w})]^{1-\gamma} \,. \quad (1.62)$$

Thus, measuring the amplitude A_θ of the angular correlation function gives an estimate for the correlation length r_0 of the real-space correlation function, provided the distribution $p(w)$ is known.

1.4.4 Estimating the angular correlation function

For calculating the angular correlation function from real data we apply the estimator by Landy & Szalay (1993),

$$\omega(\theta)\ \delta\theta = \frac{\mathrm{DD} - 2\mathrm{DR} + \mathrm{RR}}{\mathrm{RR}}. \quad (1.63)$$

DD, DR, and RR represent the number-counts of galaxy pairs with a separation between θ and $\theta + \delta\theta$ in catalogues extracted from the data (DD), from a random distribution of galaxies (RR) with the same survey geometry (including masked out regions), and between data and random catalogues (DR). The errors of the angular correlation function are estimated following the Poissonian variance approach of Landy & Szalay (1993), which is justified in the weak clustering regime,

$$\delta\omega(\theta) = \sqrt{\frac{1 + \omega(\theta)}{\mathrm{DD}}}\ . \quad (1.64)$$

In order to get rid of the shot noise from the random mock catalogues one usually creates many of them and averages the counts for DR and RR.

1.4.5 Integral constraint bias

The estimator from Eq. (1.63) is known to be biased low because the galaxy density in the field is estimated from the data itself and no fluctuations on the scale of the field size are accounted for,

$$\omega_{\mathrm{real}}(\theta) = \omega(\theta) + \mathrm{IC}\,, \tag{1.65}$$

with the bias IC usually called "the integral constraint".

It can be shown that the expectation value of this bias equals the variance of galaxy-density fluctuations on the size of the field-of-view. We estimate the integral constraint by the method outlined in Adelberger et al. (2005) from the linear CDM power spectrum. The variance of mass σ^2_{CDM} in our typical survey volumes can be estimated from integrating the power spectrum over the Fourier transform of such a survey volume,

$$\sigma^2_{\mathrm{CDM}} = \frac{1}{(2\pi)^3} \int \mathrm{d}^3k\, P_{\mathrm{lin}}(|\boldsymbol{k}|) \left| W_k^2(\boldsymbol{k}) \right| \,, \tag{1.66}$$

with $W_k(\boldsymbol{k})$ being the Fourier transform of the survey volume window function. For a rectangular field (comoving dimensions l_x and l_y) and a Gaussian distribution in radial direction (comoving width l_z), W_k reads

$$W_k = \exp\left(-\frac{k_z^2 l_z^2}{2}\right) \frac{\sin(k_x l_x/2)}{k_x l_x/2} \frac{\sin(k_y l_y/2)}{k_y l_y/2}\,. \tag{1.67}$$

Assuming a linear relation between the fluctuations of the mass density and the galaxy density, the linear bias factor can be estimated from the correlation function. An iterative approach to estimate the IC first and then the bias factor from the fitted real-space correlation function usually converges quickly. For a detailed description of the method we refer to Adelberger et al. (2005).

2

Imaging data reduction[1]

The data coming from astronomical CCD imaging cameras are influenced by many instrumental effects. The removal of these signatures is one of the main tasks of data reduction. Furthermore, the scientific exploitation of such data requires an accurate mapping of the instrumental quantities like pixel positions, count rates, etc. to physical quantities like sky positions, fluxes, etc. This mapping is achieved by numerous calibration steps during data reduction usually taking place after the removal of the instrumental effects. This thesis deals with multi-colour photometry of very faint sources. Thus, it is necessary to perform a coaddition of many exposures as a last step to reach a deep image finally.

This chapter is a general overview of the tasks carried out by our data reduction pipeline, and most principles laid down here apply to a large variety of imaging data from a variety of different cameras. More specific information about the actual data reduction of the data used for this thesis can be found in Chap. 3.

2.1 THELI

2.1.1 Overview

Over the past years our group has developed a wide-field-imaging reduction pipeline called THELI which is now publicly available.[2] Since data reduction for most optical and near-infrared cameras is very

---

[1] This chapter is partly based on the manuscripts Erben et al. (2005), published in A&A, and Hildebrandt et al. (2007a), to be published in the proceedings of the ESO Calibration Workshop 2007.

[2] ftp://ftp.ing.iac.es/mischa/THELI/

similar, the pipeline was designed in an instrument-independent way from the beginning. It is easily adaptable to new cameras by use of instrument configuration files. Nearly fully automatic processing with little need for interaction quickly leads to scientifically exploitable results making the reduction of one night of wide-field-imaging data possible in a few hours. Large parts of the pipeline are parallelised, increasing speed in presence of a multi-CPU machine. A graphical user interface (GUI) is available which facilitates the handling of reduction and configuration of THELI. A detailed description of the pipeline can be found in Erben et al. (2005). The modules for absolute photometric calibration which are not part of the public release at the moment are described in Sect. 3.2.2 and in Hildebrandt et al. (2006).

2.1.2 Structure

The pipeline is based on a number of excellent existing open source software packages like the *LDAC* tools, different *TERAPIX* packages (*SExtractor*, *SWarp*, etc.), *Eclipse, Astrometrix*, and *IMCAT*, besides others. Bash scripts are wrapped around these packages in order to handle the communication between the different tasks, to control configuration parameters, and to produce some plots for quality control. Due to this modular structure the pipeline is easily extensible and modules can be exchanged if better ones become available. This was done several times in the past; e.g., for resampling and coaddition *drizzle* was replaced by *SWarp*.

Building up a pipeline from many different software packages naturally has some disadvantages when compared to a homogeneous system which is developed from scratch. The data flow is not as transparent and error handling becomes more complicated. For very large projects like the major upcoming imaging surveys it would be desirable to be able to track the history of each of the many thousand reduced images back to the raw images. This can only be done with a sophisticated database system which is at the moment not implemented in THELI.

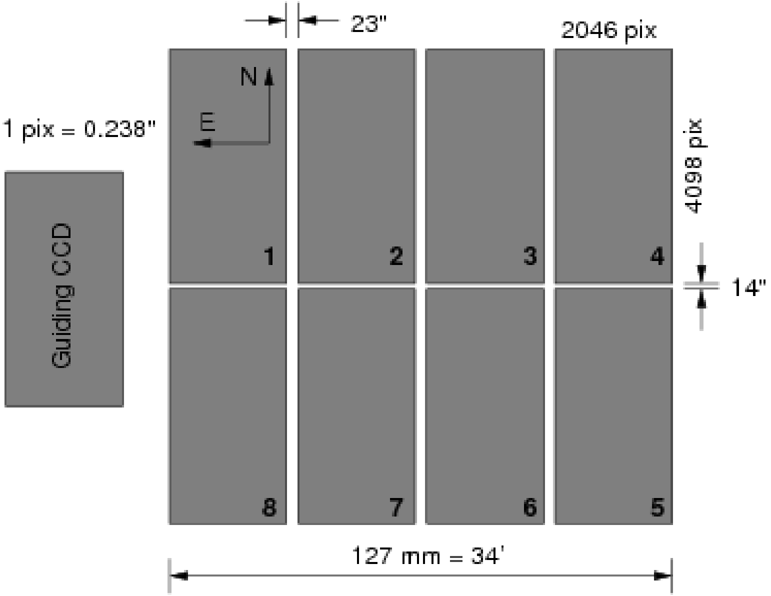

Figure 2.1: Camera layout of WFI@ESO/MPG2.2m (not to scale). Taken from Erben et al. (2005).

2.2 Pre-processing

The reduction of imaging data will be presented with data from the ESO Wide Field Imager (WFI) as an example. The camera layout of WFI is shown in Fig. 2.1.

The pre-processing, i.e. the removal of instrumental signatures from the data, is done on a single-chip basis and does not differ from well-established procedures applied for single-chip cameras. A standard overscan correction, debiasing and flat-fielding is combined with a superflat and, if necessary, a fringe-removal. Furthermore, weight images are created for every chip. See Fig. 2.2 for a visual impression of some WFI data at different reduction steps. In the

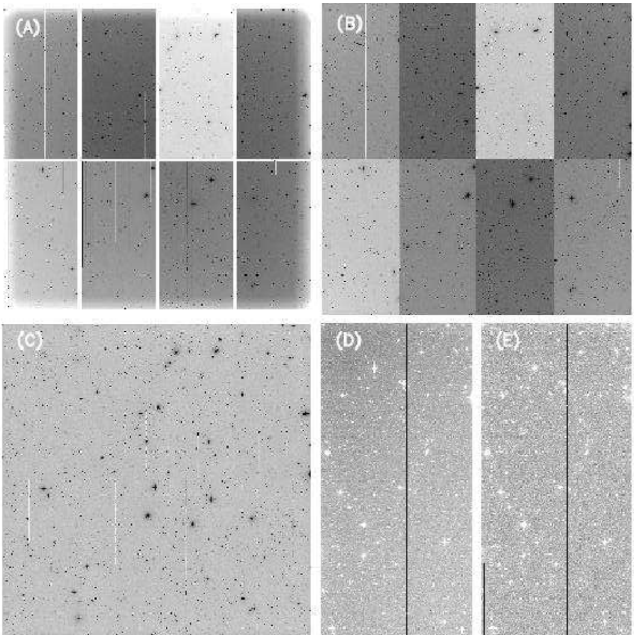

Figure 2.2: Pre-reduction steps for WFI V-band data: (A) raw data, (B)&(D) debiased and flat-fielded data (C)&(E) superflatted data. Taken from Erben et al. (2005).

following, terms that refer to a specific type of images are written in capital letters, like SCIENCE for an exposure of an object on the sky, CALIBRATION for a calibration exposure, BIAS for a bias frame, etc.

2.2.1 Overscan correction

The CCDs of a camera like WFI feature some columns that are not illuminated. These overscan regions are used to estimate the small negative voltage which is given to the whole chip to avoid positive voltages at the A/D converter after the read-out noise is added to the CCD signal. The corresponding counts are subtracted from the chip to bring it to a physical zero voltage level. This overscan correction is applied to each raw-data file coming from the observatory regardless of type (SCIENCE or CALIBRATION).

2.2.2 Debiasing

During the overscan correction the negative potential is assumed to be homogeneous over the chip. Due to inhomogeneities of the CCD chip and the read-out electronics the exact bias level also depends on position. By taking several exposures with an exposure time of 0 s, called BIAS frames, this position-dependent effect can be taken out by subtraction of a master BIAS, a combination of several of these overscan corrected BIAS frames.

2.2.3 Flat-fielding

Quantum efficiency variations over the chip and inhomogeneous illumination from the optical system are always present in modern cameras. Pointing the telescope towards a homogeneously illuminated surface like an illuminated screen or better the twilight sky yields FLAT exposures. Every SCIENCE or STANDARD frame is divided by a combination of 10-15 such FLAT frames called master FLAT. Before division, this master FLAT is normalised to a median count rate of one count. In this way, the small-scale efficiency variations of the telescope-camera system are taken out.

Experience shows that the sky background of astronomical images from a wide-field camera is still not sufficiently flat after applying a standard flat-field division. large-scale gradients on the level of a few percent can still be detected (see Fig. 2.2, panel D). An additional flat-fielding incorporating the sky background information of the SCIENCE frames themselves is used to remove these remaining inhomogeneities (see Fig. 2.2, panel E). This process which requires source detection on the SCIENCE images, masking of these sources, and smoothing of the remaining sky background is called superflatting and it depends on a dither pattern to be larger than the largest object in the field. Alternatively, one can use SCIENCE data from other sky regions in the same filter and from the same epoch, if available, to improve the quality of the SUPERFLAT.

During flat-fielding the different chips of a multi-chip camera are brought to a common photometric zeropoint. This is done by normalising the FLAT or the SUPERFLAT frames of each chip to the median of the count rates of all chips. Consequently, after flat-fielding or after superflatting, respectively, the SCIENCE exposures are brought to the same zeropoint.

2.2.4 Fringe-removal

Most CCD cameras show some amount of fringing in the redder passbands with the fringe amplitude increasing with central wavelength and with filter width. The irregular and timely variable fringe pattern is caused by a thin-layer effect due to interference between incident light-rays and reflected ones, similar to the pattern created by oil on a water puddle. This small-scale pattern can be extracted by subtracting the heavily smoothed SUPERFLAT (also called ILLUMINATION-CORRECTION) from the unsmoothed SUPERFLAT, a procedure similar to “unsharp-masking” in photographic image processing. This FRINGE-MODEL is then subtracted from the SCIENCE images. If a sufficient number of SCIENCE images from a narrow time interval is used for the creation of the SUPERFLAT this procedure usually works fine. The fringes are not visible anymore and the sky-background is flat to the $\sim$1% level. However, this requires an observing strategy in the redder bands that keeps this requirement in mind. If approximately ten exposures per hour are taken during stable sky conditions, fringes can be removed reliably in most cases. For different observing strategies, e.g. only a handful of long exposures spread over the whole night, efficient fringe-removal may be impossible.

2.2.5 WEIGHT images

During flat-fielding the sky-background in the different pixels of a chip is brought to the same level. The sensitivity variations are nevertheless real and photon statistics will be worse in some pixels than in others. Hence, during coaddition more weight should be given to the more sensitive pixels than to the less sensitive ones. A fair representation of the sensitivity of the different pixels of a chip is the normalised FLAT.

Furthermore, bad pixels and columns should be given zero weight to prevent such artifacts in the final coadded image. DARK exposures, i.e. long exposures of 10-15min with the shutter closed, have turned out to be extremely helpful in registering such bad pixels and columns. Though modern cameras cooled with liquid nitrogen do not show a significant amount of dark current (the original reason for taking DARK exposures) the DARK frames should still be taken for bad pixel registration. Some more bad pixels may also be extracted from the SUPERFLAT. Combining these bad pixel maps with the normalised FLAT yields the GLOBAL-WEIGHT image for each chip.

The individual SCIENCE frames are scanned for cosmic ray hits with *SExtractor*. Furthermore, masks for flagging satellite tracks and non-celestial objects like internal reflections, must be created by hand. In Fig. 2.3 some of these image defects are shown exemplarily. For data observed under good conditions with an optimised observation strategy this is the only point where considerable interaction by the user is required. After mask creation, individual WEIGHT images are created for each SCIENCE frame.

With reliable WEIGHT images available a weighted mean coaddition can be performed in the end instead of a median coaddition. The weighted mean coaddition increases the signal to noise considerably and allows for better characterisation of the noise properties of the final image.

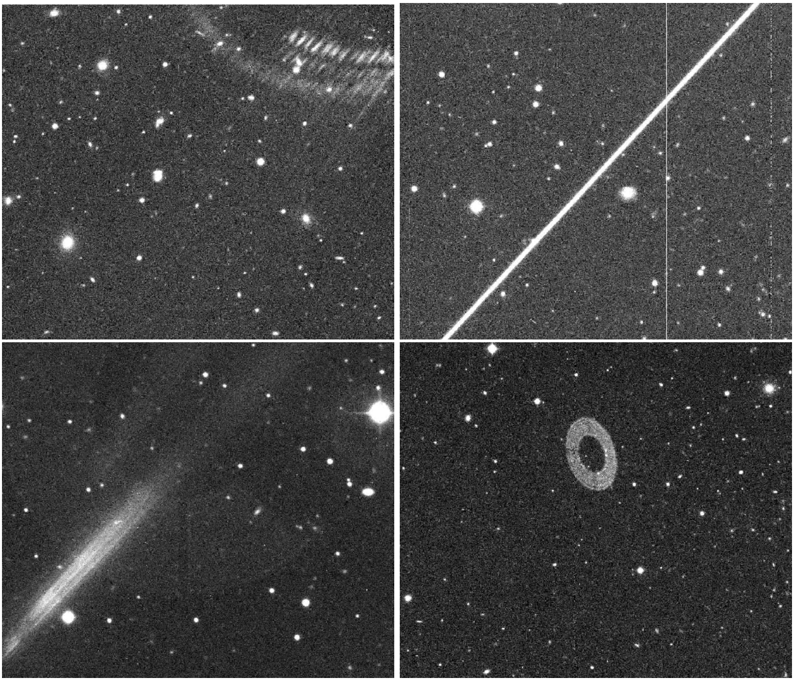

Figure 2.3: Examples of image defects which must be masked out by hand. Taken from Erben et al. (2005).

2.3 Calibration

2.3.1 Calibration requirements

The pre-reduction steps described above require frames which are not too far apart in time because instrument characteristics change. For astrometric calibration and for relative photometric calibration frames from the same region in the sky are required. Therefore, a redistribution of the files is performed after pre-reduction, grouping images for the different pointings together.

2.3.2 Astrometric calibration

THELI was designed with weak-lensing applications in mind. Therefore, a highly accurate internal astrometric calibration was mandatory from the beginning to minimise the impact of coaddition on the shape of the PSF. *Astrometrix* is used on catalogues created with *SExtractor* on each chip relating chip positions to sky positions by a third-order polynomial. With this overlapping astrometry between all chips entering a coaddition, an internal astrometric accuracy of a tenth of a pixel is achieved. The external accuracy is obviously limited by the accuracy of today's astrometric standard star catalogues.

2.3.3 Relative photometric calibration

The equalisation of the sky-background levels with the FLATs or the SUPERFLATs should leave the different chips of a SCIENCE exposure with identical zeropoints. However, different exposures at different airmasses still have different zeropoints. The difference in the zeropoints is checked during the relative photometric calibration comparing the magnitudes of overlap objects from the catalogues created for the astrometric calibration between all chips. A χ^2 minimisation yields the relative zeropoints of the different chips.

2.3.4 Absolute photometric calibration

One of the most critical aspects of imaging data reduction is the absolute photometric calibration relating the counts in the CCD to a physical flux scale. Especially for ground-based data influenced by the Earth's atmosphere this is a complicated task that depends on a sophisticated observing strategy.

Different approaches are possible within THELI:

1. The absolute photometric calibration is done on a night basis. Lacking a dense grid of standard stars across the entire sky SCIENCE frames cannot be calibrated individually. Rather, it is assumed that the atmosphere does not change over a

night and that the photometric calibration derived from additional STANDARD star exposures applies to the SCIENCE frames of the same night as well. A detailed description of this method, also applied to the WFI data used for this thesis and the problems associated with this method, can be found in Sect. 3.2.2.

2. If the observatory features a photometric auxiliary telescope or camera it can be decided already during an observation whether a SCIENCE frame is taken under photometric conditions or not, and a photometric zeropoint can be assigned to it directly at the observatory. This method is applied for MEGACAM@CFHT.

Regardless of whether the zeropoints for some of the images originate from method 1 or 2 they are used as additional constraints for the χ^2 minimisation. See Sect. 3.2.2 for details on how the final zeropoint for the coadded image is estimated finally.

2.4 Coaddition

The *SWarp* code, a derivative of the *drizzle* approach, is used for the coaddition of the astrometrically and photometrically calibrated images. In *drizzle* (see Fig. 2.4) an output grid is defined and the pixels of all input images are projected to this output grid via the astrometric header information. The flux of one input pixel is distributed to different output pixels according to the overlapping area. Thereby, the input pixel is weighted with a kernel function which puts more weight to the central part of a pixel than to the outer parts. A careful choice of this kernel is important to minimise the effects of correlated noise in the output pixels.

SWarp uses a backward-mapping technique projecting the output grid back to all n input grids and averaging over the n estimates for the output pixel flux. Compared to *drizzle* which puts the frames on one another in a serial way this procedure has the advantage that not only linear combinations of the n frames are allowed. For example, one can apply a rejection of outlier frames which is not possible for

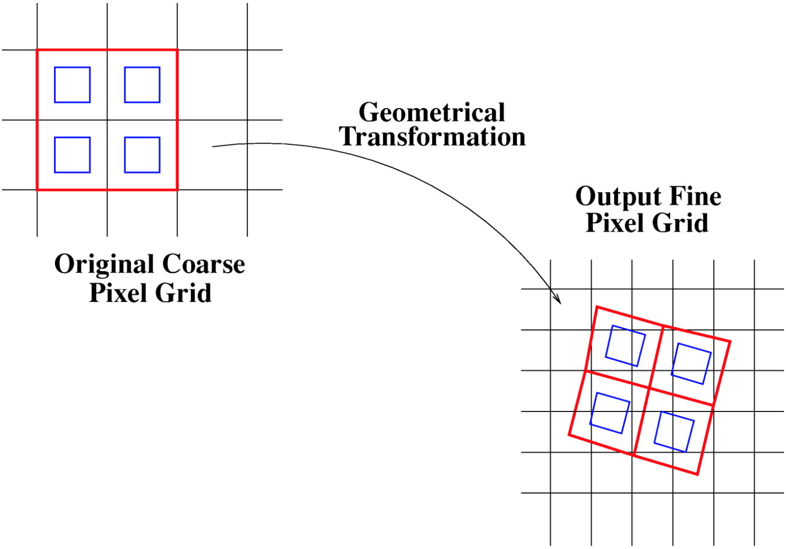

Figure 2.4: Illustration of the *drizzle* approach. Taken from Fruchter & Hook (2002).

drizzle. Furthermore, *SWarp* includes some advanced kernels that improve image quality considerably.

3

The ESO Deep Public Survey[1]

The ESO Deep-Public-Survey (DPS) is a multi-colour imaging survey carried out by the ESO Imaging Survey (EIS) team under the program IDs 164.O-0561 and 169.A-0725. It consists of optical data in the $UBVRI$-bands observed with WFI at the 2.2 m telescope at La Silla and infrared data in the J- and Ks-bands observed with SOFI at the New Technology Telescope. In this chapter we present 64 reduced, photometrically and astrometrically fully calibrated, and stacked images of the optical part which were released to the scientific community.[2] These images were reduced and calibrated with our THELI reduction pipeline (Chap. 2 & Erben et al. 2005).

The main scientific driver for the ESO DPS were searches for high-redshift galaxies, distant clusters, high-redshift QSOs, low surface-brightness galaxies, and gravitational lensing studies. The survey was designed in a way to deliver a unique dataset also for studies on Galactic structure, very low-metallicity stars, white dwarfs, M-dwarfs, and field brown dwarfs. The deep imaging data can be used to pre-select objects by colour for follow-up spectroscopy with VLT instruments.

It was intended to cover an area of three square degrees in three well separated regions of one square degree each at high galactic latitude called Deep1, Deep2, and Deep3. Each region consists of four adjacent WFI pointings ($34' \times 33'$) at the same declination named a, b, c, d in order of decreasing right ascension. Table 3.1 summarises

---

[1] This chapter is largely based on the paper Hildebrandt et al. (2006), published in A&A.

[2] The images can be downloaded at `http://marvin.astro.uni-bonn.de/DPS/` and are available via the ESO archive `http://archive.eso.org/archive/eso_data_products.html`.

Table 3.1: Positions and available colours of the twelve DPS fields and the two mispointings (Deep1e and Deep1f; see text). The region Deep1 overlaps with the ATESP radio survey (Prandoni et al. 2000) and the field Deep2c is centred on the Chandra Deep Field South (CDFS).

field	RA [h m s] J2000.0	DEC [d m s] J2000.0	avail. colours
Deep1a	22:55:00.0	−40:13:00	*UBVRI*
b	22:52:07.1	−40:13:00	*UBVRI*
c	22:49:14.3	−40:13:00	*UBVRI*
d	22:46:21.4	−40:13:00	—
Deep2a	03:37:27.5	−27:48:46	*R*
b	03:34:58.2	−27:48:46	*UBVRI*
c	03:32:29.0	−27:48:46	*UBVRI*
d	03:29:59.8	−27:48:46	*R*
Deep3a	11:24:50.0	−21:42:00	*UBVRI*
b	11:22:27.9	−21:42:00	*UBVRI*
c	11:20:05.9	−21:42:00	*UBVRI*
d	11:17:43.8	−21:42:00	*BVRI*
Deep1e	22:47:47.9	−39:31:06	*URI*
f	22:44:58.4	−39:31:54	*I*

the positions of the twelve DPS fields. The Deep1 region was chosen to overlap with the ATESP radio survey (Prandoni et al. 2000). The Chandra Deep Field South (CDFS) is included in the Deep2 region (centred on the field Deep2c), and Deep3 is a random, empty, high galactic latitude field positioned in such a way that DPS observations are possible over the whole calendar year.

With the wide-field imaging reduction pipeline described in Chap. 2 we began reduction of the DPS data in late 2003 as an ideal test case of a unique, large dataset fitting our scientific goals. We are mainly interested in searches for Lyman-break galaxies (LBGs) and weak-lensing studies supported by photometric redshifts. In November 2004 the EIS team released 40 reduced images of the optical part of the DPS which enabled us to compare the performance of our pipeline to a different one. The released EIS data cover a subset of our data-release and are available via the ESO archive.

The chapter is organised as follows. In Sect. 3.1 the instrument, its photometric system, the observing strategy, and the raw data are described. Sect. 3.2 gives a short summary of the data reduction with the THELI pipeline with emphasis on steps which are important for the end user. The properties of our released images are presented in Sect. 3.3, and in Sect. 3.4 these properties are compared to the EIS data release of the ESO DPS. A summary of the data release and an outlook on projects carried out with the DPS data is given in Sect. 3.5.

3.1 The raw data

All optical data of the DPS were observed with WFI@ESO/MPG2.2m. This multi-chip, focal-reducer CCD camera has a field of view of $34' \times 33'$ with a filling factor of $\sim 96\%$. Eight broad-band filters are used for the DPS some of which are very different from the standard Johnson-Cousins $UBVRI$ filters. Their properties are summarised in Table 3.2 and the total throughput of the main photometric system is shown in Fig. 3.1. The U-band filter ESO # 841, the B-band filter ESO # 842 and the I-band filter ESO # 845 were replaced by new filters for DPS observations after some time. In Table 3.3 the scheduled exposure times and expected limiting magnitudes are shown which were chosen to suffice the scientific goals as described above. For our reduction of the DPS we used all the raw data available until October 2006 including images from the two large programmes 164.O-0561 and 169.A-0725 (P.I. Krautter for both) and, moreover, data from the programmes 67.A-0244 (P.I. Schneider), 75.A-0280 and 77.A-0261 (P.I. Hildebrandt for both), and 77.A-0211 (P.I. Mignano). For the field Deep2c we added also data from the programmes 168.A-0485 (P.I. Cesarsky), 64.H-0390 (P.I. Patat), 66.A-0413 (P.I. Clocchiatti), 68.D-0273 (P.I. Cappellaro), 68.A-0443 (P.I. Clocchiatti), 70.A-0384 (P.I. Vanzi), and 74.A-9001

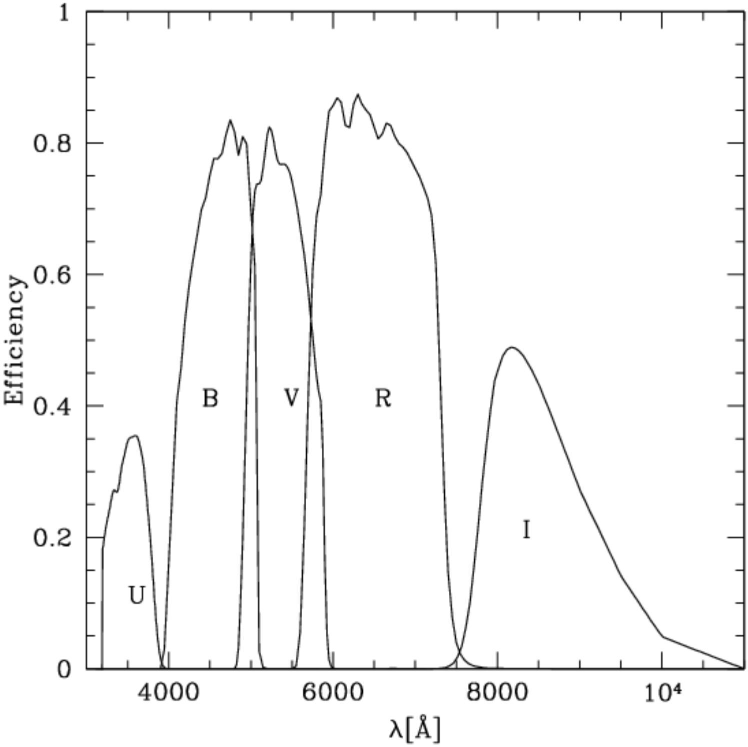

Figure 3.1: Instrumental response of WFI in the different filters. Shown are the new U-band and the old B- and I-band filters which were mainly used for the DPS.

(P.I. Kuijken), and we used the WFI commissioning data and data from the COMBO17 survey (P.I. Meisenheimer) described in Wolf et al. (2004).

Unfortunately, the optical observations of the DPS were not finished. Not all fields were observed to the scheduled depths in all

Table 3.2: Characteristics of the filters used for the DPS. For the computation of the effective wavelength and the filter's width the CCD efficiency is included.

Filter	ESO #	eff. WL [$\AA$]	width [$\AA$]	AB corr. [mag]
U_{35060}	877	3429	705	1.1
U_{38}	841	3647	373	0.8
B_{123}	878	4544	757	−0.1
B	842	4589	871	−0.1
V	843	5377	793	0.0
R	844	6504	1502	0.2
I	845	8635	1425	0.5
I_{EIS}	879	8057	1402	0.5

Table 3.3: Scheduled exposure times and expected limiting magnitudes for the images in the different filters. The limiting magnitudes correspond to 5σ sky level in an aperture of $2''$.

Filter	exp. time [s]	mag lim. [Vega mag]
U_{35060}	43 200	25.7
U_{38}	61 200	26.0
B_{123}	12 600	26.1
B	12 600	26.1
V	9000	26.0
R	9000	26.1
I	27 000	25.5
I_{EIS}	27 000	25.5

filters. Furthermore, some observations were executed under bad sky conditions so that these images are excluded from the data reduction. There are also some mispointings present in the DPS raw data for which the FITS header contains the correct coordinates of one DPS field while the actually observed area is offset by some arcminutes. Some fields (especially the field Deep2c centred on the CDFS) were observed by different programmes to much greater depth, and we also included these data in our reduction. The estimates of the limiting magnitudes in Table 3.3 were quite optimistically calculated at a time when the WFI did not yet exist. The final images are significantly shallower (see Table 3.5). For all these reasons the DPS is rather heterogeneously deep, and this fact should be kept in mind when dealing with data from different DPS fields.

The desired depth usually requires observations in different nights for every field in each colour. For many fields, observations spread over more than a year and images taken under very different photometric conditions have to be combined. Even the instrument setup (baffling, camera rotation etc.) changed over the time of DPS observations. Furthermore, to reach a final coadded image with an exposure time as uniform as possible over the field and in order to get good relative astrometry and photometry between the chips, a wide dither pattern was chosen for the DPS. The whole observing strategy (dither pattern, exposure times of the individual images etc.) was adapted and optimised over the years of DPS observations.

3.2 Data reduction

The data reduction was performed with our THELI pipeline described in detail in Chap. 2 and in Erben et al. (2005). All the raw data were requested from the ESO Science Archive Facility.

3.2.1 Pre-reduction

The data were sorted in so-called observation runs which include science, calibration, and standard star images from some adjacent nights. In the $UBVR$-bands an observation run can easily contain data from one (R) or even two weeks (UBV). The strong fringe patterns in the I-band are variable from night to night and even within one night, so that I-band runs usually contain data from only one or two hours of observation.

On this run basis, the science images were pre-reduced which includes overscan correction, bias subtraction, flat-fielding, superflatting, and in the R- and I-band also fringe-removal. No correction for the inhomogeneous illumination (see Koch et al. 2004) was applied. Weight images were created containing flags for bad pixels, bad columns, and other image defects like satellite tracks which were masked out by hand. The Landolt standard star images which are available for most nights were reduced in the same way as the science images.

3.2.2 Absolute photometric calibration

Here we describe in a bit more detail the absolute photometric calibration which is not covered in Chap. 2 and in Erben et al. (2005). All objects from the standard star frames were extracted and matched to a photometric standard star catalogue, the Landolt catalogue (Landolt 1992) in the U-band and the Stetson catalogue (Stetson 2000)[3] for the other bands, respectively. Per WFI field, this usually led to more than a thousand matched stars per night in the $BVRI$-bands and to around a hundred in the U-band which could be used for calibration. The observed instrumental WFI magnitudes, $m_{\mathrm{inst.}}$, were related to the standard Johnson-Cousins system, m_{JC}, by the following equation:

$$m_{\mathrm{JC}} = \mathrm{ZP} + m_{\mathrm{inst.}} + \mathrm{Colour} \cdot \mathrm{CT} + \mathrm{Airmass} \cdot \mathrm{EXT}\,, \tag{3.1}$$

with ZP the photometric zeropoint, CT the colour term, and EXT the extinction coefficient. It should be stressed that the photometric system of WFI is very different from the standard system and this linear relation fails for some filters (especially the U- and B-filters) and objects with large colour terms. It is in general necessary to work with instrumental magnitudes when doing photometry with WFI. Depending on the number of matched standard stars, different solutions were chosen for the different nights. In nights with standard star fields spanning a wide range of airmasses, the instrumental

---

[3] available at `http://cadcwww.hia.nrc.ca/cadcbin/wdbi.cgi/astrocat/stetson/query`

magnitudes could be fit to the standard system's magnitudes with a three-parameter fit using the zeropoint, colour term, and extinction coefficient as free parameters (for EXT only negative values are permitted). For nights with poorer coverage in airmass, the extinction term was fixed to a default value (taken from the WFI website, `http://www.ls.eso.org/lasilla/sciops/2p2/E2p2M/WFI/`) and a two-parameter fit with the zeropoint and colour term as free parameters was applied. Sometimes, especially in the U-band, also colour coverage was not good enough to fit for the colour term. Then only the zeropoint is estimated in a one-parameter fit with the colour term also set to a default value. The decision was taken with the help of plots showing the differences between instrumental magnitude and standard star magnitude plotted versus airmass and versus colour; see Fig. 3.2 for an example. For this particular night, the three-parameter fit yields an extinction coefficient that is much too large (compared to the default value EXT$= -0.07$) resulting in a bad zeropoint. Obviously this night is not totally photometric. Nevertheless, experience shows that such nights can often still be used to estimate a zeropoint and sometimes even a colour term for the images. The one-parameter fit was chosen here because the non-linearities visible in the colour plots for the one- and the two-parameter fits influence the value for CT2, resulting again in a zeropoint ZP2 that is too large. For ideal photometric nights the values from the three different fits show only a very small scatter. The choice from these plots should be regarded as the first run of our absolute photometric calibration. Experience shows that it is often not possible to entirely judge the photometric quality of a whole night from these plots alone. Sometimes photometric conditions change over the night which cannot be detected from these standard star exposures typically observed two or three times a night only. Thus, we evaluate the calibration by three further means, namely corrected photometric zeropoints (see Sect. 3.2.4), colour-colour-diagrams of stars (see Fig. 3.4), and apparent magnitude number-counts (see Fig. 3.5). The method based on corrected photometric zeropoints fully exploits the long term characteristic of the DPS with observations in one field and colour usually spanning many nights. All photometric solutions were added

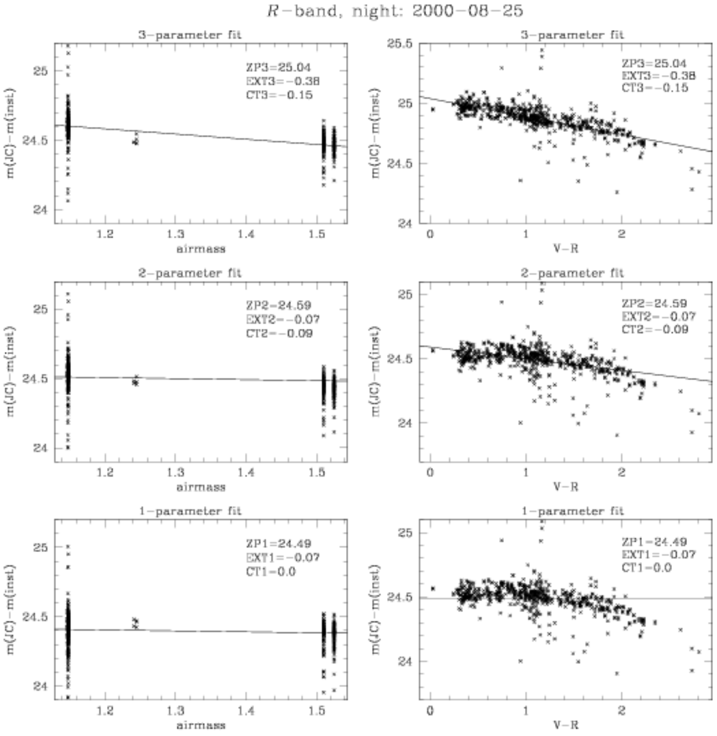

Figure 3.2: Plot for the decision between photometric solutions of one night. The first row shows the results for the three-parameter fit, the difference between instrumental magnitude (data from all eight chips) and standard star magnitude plotted versus airmass (for colour zero) on the *left* and versus colour (for airmass one) on the *right*. In the second and third row the same diagrams are shown for the two- and the one-parameter fit. See the text for a discussion. The instrumental photometric errors are not shown in these plots since at the bright magnitudes of the Stetson/Landolt standard stars they are negligible and would hardly be visible.

to the image header so that later changes could be made (e.g. from 2-parameter fit to 1-parameter fit). Every night was given a unique so-called GaBoDS ID starting with 1 on January, 1st, 1999. The photometric solutions for all the nights of the DPS can be found in the Appendix (see Tables A.1 to A.8). On December, 13, 2004 the EIS team also published their photometric solutions for all nights calibrated by them which can be found on the web.[4]

3.2.3 Transition from run to set

For every run websites were created containing essential image information like exposure time, seeing, and other important quantities.[5] The further steps like relative astrometric and photometric calibration and coaddition require all images from the same pointing. Hence, the images were re-distributed from the runs into so-called sets containing pre-reduced images from different epochs but at the same sky position.

3.2.4 Astrometric and photometric calibration

The astrometric calibration was performed with *ASTROMETRIX* (Radovich 2002) on catalogues created with *SExtractor* (Bertin & Arnouts 1996). The external astrometry was fixed with respect to the USNO-A2.0 (Monet et al. 1998) catalogue. *LDAC* tools were used to do the relative photometric calibration and to bring all images to the same flux scale. First, the relative zeropoints $\mathrm{ZP_{rel}}$ of each single chip was estimated from the flux differences of overlap objects. At this stage it was required that the sum of the relative zeropoints of all images of the set equals zero, $\sum_i \mathrm{ZP}_{\mathrm{rel},i} = 0$. Then the images from calibrated nights (see Sect. 3.2.2) were taken and so-called corrected zeropoints $\mathrm{ZP_{corr}}$ were calculated in the following way:

$$\mathrm{ZP}_{\mathrm{corr},i} = \mathrm{ZP} + \mathrm{Airmass} \cdot \mathrm{EXT} + \mathrm{ZP}_{\mathrm{rel},i}\,, \qquad (3.2)$$

---

[4] `http://www.eso.org/science/eis/surveys/release_70000027_Photometry.html`

[5] available at: `http://marvin.astro.uni-bonn.de/DPS/`

with ZP and EXT being the zeropoint and the extinction coefficient of the particular night, respectively. Theoretically, these corrected zeropoints should then coincide for photometric frames, and usually this assumption holds for most sets to within $\sim 0.05\,\mathrm{mag}$. Deviations were used as hints for bad absolute photometric calibrations or changing photometric conditions over the night. That is to say, if conditions change from the standard star exposure to the scientific exposure over the course of the night, this can be detected from the corrected zeropoints. Furthermore, it can happen that from the plots shown in Fig. 3.2 it is difficult to decide whether solution number two or three should be chosen. Sometimes it is even not entirely clear whether a night should be rejected completely for the absolute photometric calibration. At this point the distribution of the corrected zeropoints can help to identify nights where obviously a bad solution was chosen. In Fig. 3.3 such a situation is illustrated by an extreme example. There, the solutions chosen in the first run of the absolute photometric calibration yielded a large scatter in the corrected zeropoints. After rejecting some nights with small numbers of standard stars or zeropoints far from the default values (hints for non-photometric conditions) in the second run, the scatter in the corrected zeropoints decreased considerably, and the mean of this distribution was taken as the zeropoint for the coadded image. After such a treatment the distribution of the corrected photometric zeropoints shows a HWHM scatter of $\sigma \lesssim 0.05$mag for all images.

With the current calibration plan for WFI, taking standard star frames mostly at the beginning and the end of the night (sometimes in twilight), it is not possible to account for changing photometric conditions during one night. Thus, a night has to be accepted or rejected and no discrimination between good and bad data from one night is feasible since the scientific data usually span the whole night.

After the absolute photometric calibration, the fluxes of the images were multiplied by a factor $10^{-0.4\mathrm{ZP}_{\mathrm{rel},i}}/t_i$, with t_i being the exposure time of the single image. Thus, the counts in our final coadded image are scaled to one second.

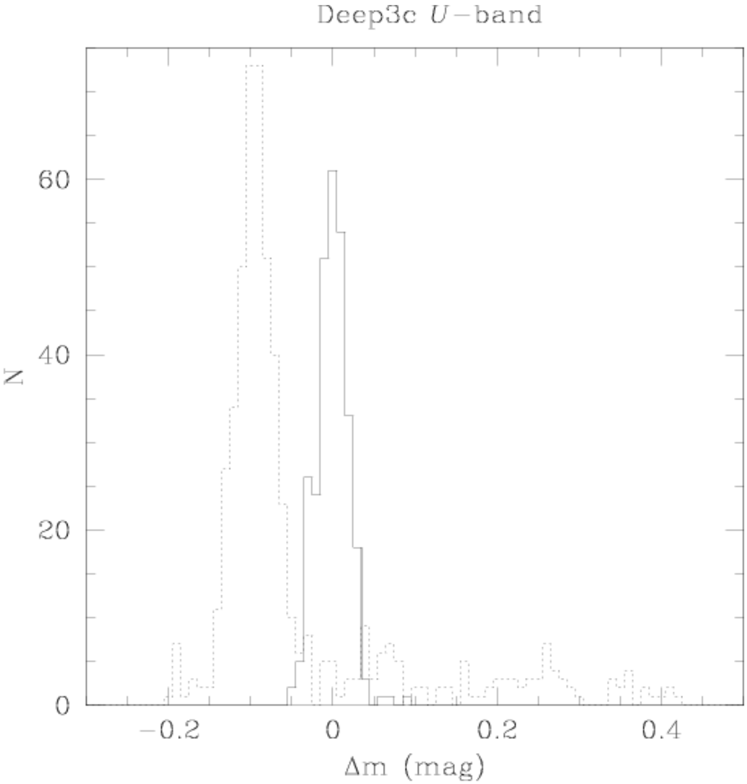

Figure 3.3: Distribution of the differences Δm between the mean of the corrected zeropoints $\frac{1}{N}\sum_i \mathrm{ZP}_{\mathrm{corr},i}$ and the individual corrected zeropoints $\mathrm{ZP}_{\mathrm{corr},i}$ for the field Deep3c (U-band). The dashed line represents the distribution before selecting different photometric solutions by hand, and the solid line represents the distribution after this selection.

3.2.5 Coaddition

The calibrated images then entered the coaddition which was performed with *SWarp* (Bertin 2003). We chose a weighted mean coaddition to maximise the S/N of our final images relying on the efficiency of our weight maps. Some quantities like seeing, sky background, mean stellar ellipticities etc. were estimated for every single frame during the relative astrometric and photometric calibration step. By these values, single frames could be excluded from the coaddition to avoid a degradation of the final image. In fact, different coaddition conditions were often chosen so that for some fields there is more than one image available in some filters. For example, it can be useful for certain scientific purposes to have an image as deep as possible (e.g. for deep multicolour photometry) while seeing is not so important, whereas for other scientific applications a very good seeing is mandatory (e.g. for weak lensing studies). The images were given a unique coaddition ID and the selection criteria applied to the single frames entering the coaddition were written to the FITS header (see Table B.1).

3.3 Released data

3.3.1 Images

During the reduction of the DPS we produced 64 coadded images and their corresponding weight maps. The basic properties of the coadded images are summarised in Table 3.5. Usually the first coaddition aimed at including as many single exposures as possible to maximise the total exposure time. Exposures with a seeing of $> 2''$, unusual high background fluxes, or large relative photometric zeropoints offsets compared to the rest of the set (indicative of twilight, moonlight, or clouds) were excluded from the coaddition. These coadded images were then assigned the letter “A” for “All”. For example, the first coaddition of a Deep1c image then has the coaddition ID D1CA. If the exposures allowed for a second image with better seeing (but certainly with reduced exposure time) a second coaddition was performed and the letter “S” for “Seeing” was assigned to this image (e.g. coaddition ID D1BS). Sometimes, even further conditions were applied to the single frames for another coaddition then denoted by arbitrary letters like “C” or “G”. During 2005 more data from the GOODS program became available for the field Deep2c and in 2006 two of our own WFI observing programs in the field Deep1c were finished. Thus, further images denoted by a letter “B” were created with the same conditions as the images with assigned letter “A” but including more data. Only these “D1CB” and “D2CB” images are shown in Table 3.5.

The mispointings were processed, too. Therefore our data release contains two new fields called Deep1e and Deep1f that are not original DPS fields (see Table 3.1) but that are nevertheless quite deep in some bands and may be useful for some applications.

A description of the GaBoDS image headers can be found in Appendix B.

3.3.2 Photometric accuracy

Since we did not perform an illumination correction in the pre-reduction step, a minimum photometric error of about 0.05 mag (see Koch et al. 2004) is present when using the same zeropoint over the whole field. In order to check the absolute photometric calibration, catalogues of all objects in the coadded images were created and the colours of stars (selected by the *SExtractor*-CLASS_STAR parameter) were plotted in colour-colour diagrams. These measured colours were compared to theoretical stellar isochrones by Girardi et al. (2002). The DPS fields are located at high galactic latitude so that mainly halo stars are observed. Thus, we choose a low mettalicity of $Z = 0.004$ ($\approx 1/5\times$ solar), an age of 10 Gyr, and a mass limit of $m_{\rm ini} < 0.92\,M_\odot$ for the isochrones. It should be stressed that these choices are not critical because the main sequences of isochrones with different parameters differ only slightly. In such a colour-colour diagram similar errors in the different passbands can cancel due to the subtraction. However, for the DPS fields with five filter coverage this is very unlikely to happen in all colour-colour diagrams. An exemplary plot is shown in Fig. 3.4 and all other plots can be found in Appendix A.2. After the thorough inspection of the corrected zeropoints no offsets larger than ~ 0.15mag (perpendicular to the track) were present in these colour-colour diagrams.

Judging from these plots we estimate the error of our absolute photometric calibration to be below ~ 0.1 mag for our released images, in most cases better.

For an alternative check of the absolute photometric calibration, the number-counts of objects were compared to the object number-counts of our old, photometrically well-calibrated CDFS data (ESO Press Photos 02a-d/03) in BVR and to the CDFS data from Arnouts et al. (2001) in U_{35060}, U_{38}, and I, respectively. An example is shown in Fig. 3.5. There is good agreement in these plots for all images within the expected field-to-field variance (especially at brighter magnitudes). For the I_{EIS} filter no such comparison was performed.

The performance of photometric redshifts (photo-z; see Chap. 4) on the data from field Deep2c is analysed in detail in Chap. 5. This is another, yet very indirect, test of photometric quality. Many other parameters influence the performance of the photo-z's so that it is not straightforward to relate a result on the accuracy of photo-z's to the accuracy of the photometric calibration. But the results in Chap. 5 indicate at least that the absolute photometric calibration is not completely off and that the images coming out of the pipeline can be readily used for such analyses.

3.3.3 Astrometric accuracy

While overlap astrometry was used for all single chips entering the coaddition in one band, the final images of different bands were processed astrometrically independently and the catalogues produced for the colour-colour diagrams are also used to check the internal astrometric accuracy between the different colours. The positional differences of associated objects from the B- and V-band images for the field Deep1b are plotted in Fig. 3.6. The dependence of these differences on chip position is shown in Fig 3.7 where no residual effects like chip boundaries etc. are visible. For all fields,

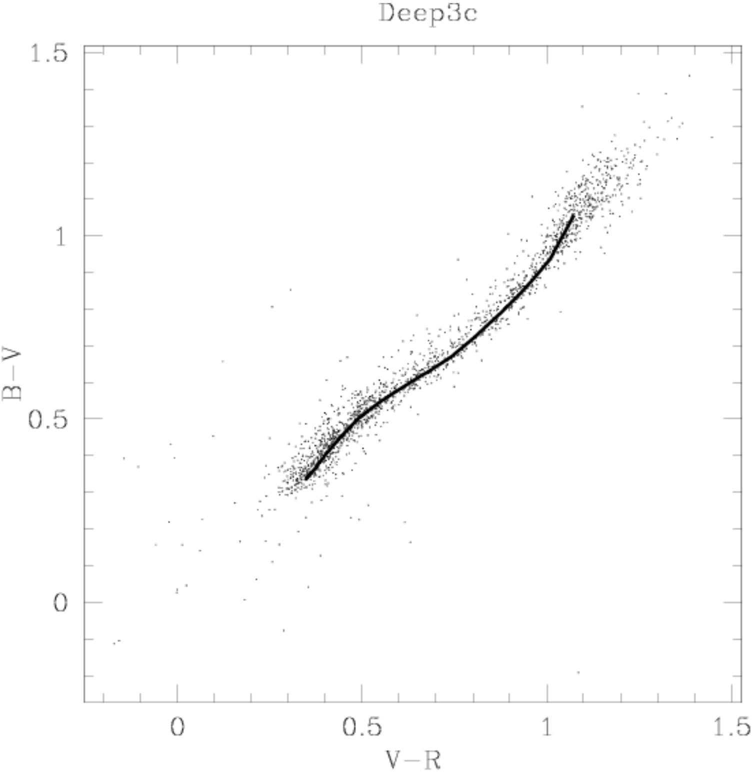

Figure 3.4: Colour-colour diagram for stars in the field Deep3c in comparison to an isochrone from Girardi et al. (2002). The isochrone was calculated for stars with a metallicity of $Z = 0.004$ ($\approx 1/5\times$ solar) and an age of 10 Gyr. Only stars with an initial mass of $m_{\rm ini} < 0.92\,M_\odot$ were included.

the distribution of the position differences between two bands are well described by a Gaussian with a half width at half maximum of $\sigma \lesssim 0\farcs2$ (see Table 3.5). A slightly higher internal astrometric

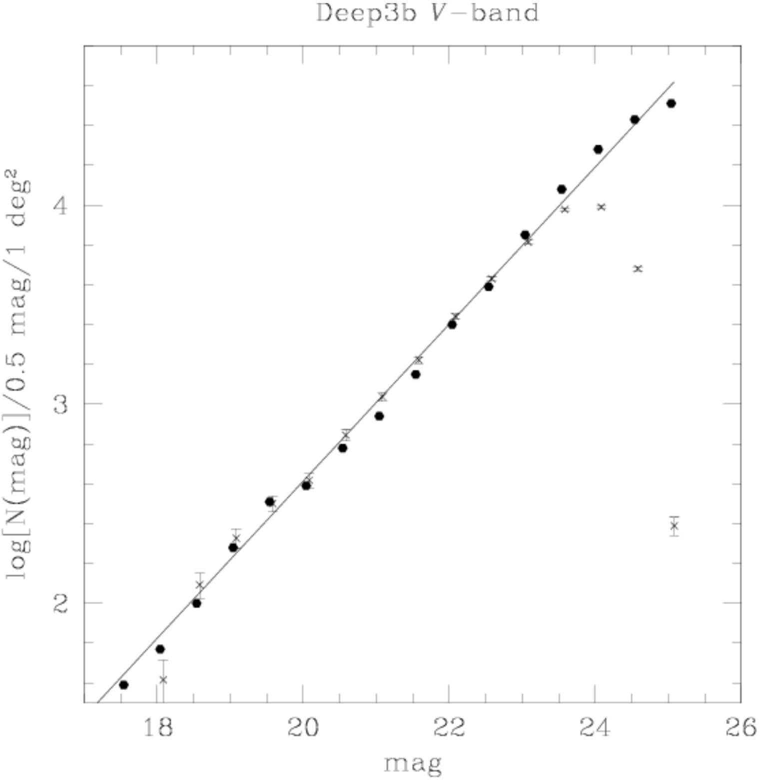

Figure 3.5: Apparent V-band magnitude number-counts for objects in the field Deep3b (crosses) in comparison to the well-calibrated field Deep2c (hexagons) showing no significant offset.

accuracy could have been obtained by using an astrometric catalogue from one band as the reference for the other bands. Given our already high accuracy of $\sigma < 1$pix this step was not performed for flexibility reasons. The quoted accuracy should suffice for almost all scientific goals.

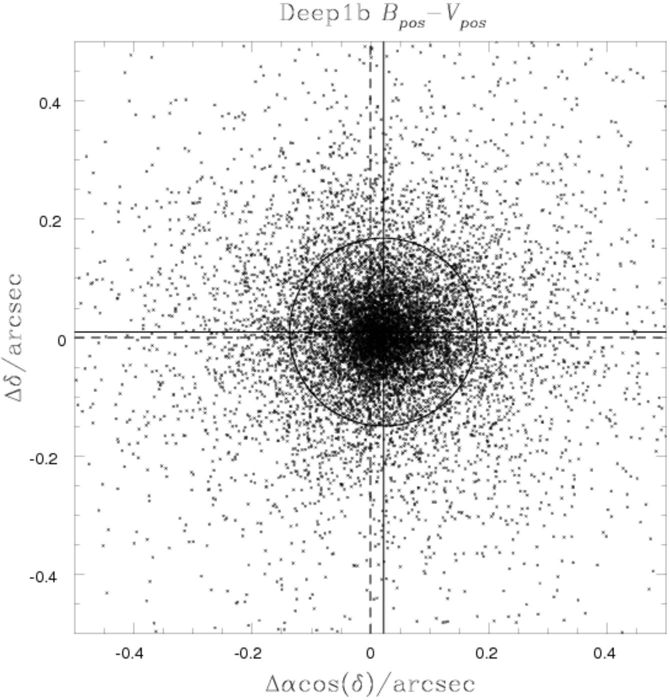

Figure 3.6: Distribution of the positional differences of objects (stars as well as galaxies) in the Deep1b B- and V-band images. The mean and standard deviations here are: $\langle\Delta\alpha\cos(\delta)\rangle = 0\farcs02 \pm 0\farcs16$ and $\langle\Delta\delta\rangle = 0\farcs01 \pm 0\farcs15$. The solid lines mark the centre and the circle represents the one sigma interval (radius: $0\farcs15$).

The external astrometric accuracy is checked with the help of the UCAC-2 catalogue (Zacharias et al. 2004). No offsets larger than the accuracy of the USNO-A2.0 catalogue of $0\farcs3$ are found.

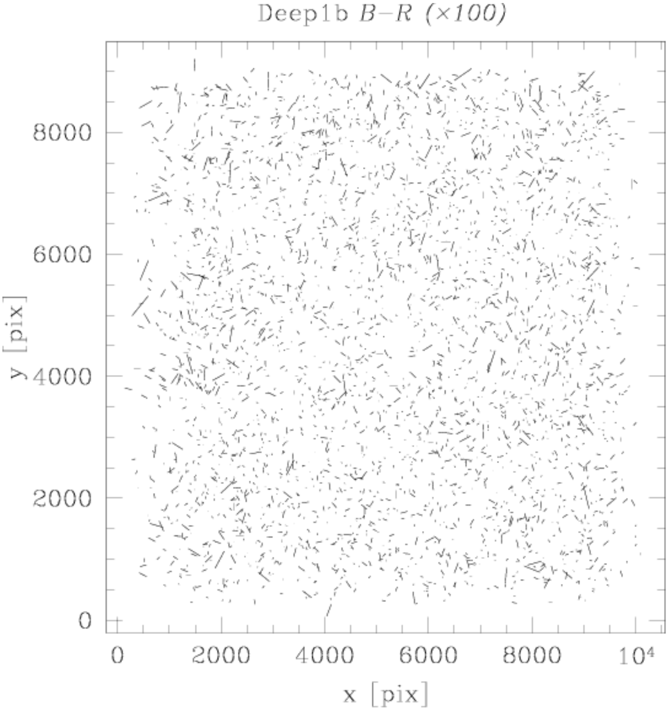

Figure 3.7: The sticks in this plot represent the position differences of objects (stars as well as galaxies) between the Deep1b B- and R-band image. The dependence of the positional differences on chip position is shown. The distribution seems to be random and no chip boundaries etc. are visible. The differences have been multiplied by a factor 100 and the longest sticks correspond to $\sim 1\overset{''}{.}5$.

3.3.4 Websites

The released images are accompanied by websites which give quick access to extensive information about the data reduction. The pages are organised as follows. At the homepage (`http://marvin.astro.`

`uni-bonn.de/DPS/`) links lead to pages for every filter with further links to all observing runs in this band. These run webpages consist of pages for the calibration frames (BIAS, DARK, SKYFLAT) and for the scientific frames. For the calibration images, thumbnails for visual inspection of the master calibration frame and lists for all single frames used for the creation of the master image are presented. For the scientific frames, basic statistics are listed and a thumbnail of every single frame can be inspected together with a plot of the measured PSF. Further links lead to the photometric solutions of the actual night with plots used for the decision between the different photometric solutions (see Fig. 3.2).

For every set, websites listing all coadded images in a particular band and their basic properties are also linked from the homepage. Plots showing the distribution of single exposure quantities like seeing, relative photometric zeropoint, etc. are featured as well as colour-colour-plots for all possible combinations. The website for a coadded image contains crosslinks to the run websites of the single frames included and a number of checkplots created after coaddition. In particular, there are checkplots showing the number-counts (see Fig. 3.5) and angular correlation functions of galaxies in different magnitude bins.

3.4 Comparison to the EIS DPS optical data release

3.4.1 Released data

On November 10, 2004 the EIS team released 40 reduced and stacked optical images of the DPS to the scientific community.[6] Their reduction includes data accumulated until September 28, 2002. Compared to our data release, there are no images for the mispointings, for Deep1a in the V-band, for Deep1c in the U_{35060}, the B_{123}, and the I_{EIS}-bands, for Deep2c in the B- and R-bands, for Deep3a in the

---

[6] available at: `http://www.eso.org/science/eis/surveys/release_60000024_DPS.html`

I-band, and for Deep3d in the B- and R-bands. Furthermore, we used more data for the fields Deep1b (U), Deep2b (I), Deep2c (all bands), and Deep3b (U), reaching greater depth. The EIS release includes an image of the field Deep2c in the U_{38}-band which is not included in our release.

3.4.2 Photometric comparison

After some comparisons between objects' magnitudes in our reduction and in the EIS data release it became clear that the absolute photometric calibration of the EIS images is somewhat different from our calibration. Absolute offsets of some tenths of a magnitude are present for most images. These offsets are also present in colour-colour diagrams when the EIS data are compared to theoretical isochrones by Girardi et al. (2002). Intense cooperation with the EIS team led to the discovery of a mistake in their final application of the absolute photometric calibration to the images (Miralles, private communication) so that no further comparison is meaningful here.

Nevertheless, if both pipelines performed a correct relative photometry between the eight WFI chips there should be only a constant offset between the magnitudes of objects measured in an EIS image and the same objects measured in one of our images. Within photometric errors this is indeed the case for all images from the Deep1 and Deep3 regions. For images from the Deep2 region, the scatter of the magnitude differences is much larger. A careful cooperative analysis with EIS revealed another mistake in the EIS reduction (Miralles, private communication). The chips were not brought to a common zeropoint during the reduction of the Deep2 fields in contrast to Deep1 and Deep3.

3.4.3 Astrometric comparison

The EIS team used the GSC2.2 catalogue for the absolute astrometric calibration which is known to be offset with respect to the USNO-A2.0 catalogue we used. We do not find any offsets between our images and the EIS images of larger than $\sim 0\overset{''}{.}5$. The actual offset depends strongly on the sky position with the best match in the Deep1 fields and the worst in the Deep3 fields. The scatter is comparable to the scatter between the different colour catalogues from our images (see Sect. 3.3.3 and Table 3.5).

Table 3.4 summarises the comparisons between our data release and the EIS data.

Table 3.4: Comparison of the GaBoDS (we always take the deepest image if more than one coaddition is available) and the EIS data. The mean and the standard deviation of the magnitude differences ($\Delta m = m_{\mathrm{GaBoDS}} - m_{\mathrm{EIS}}$) are computed for objects in the magnitude range $17 < m < 21$. Due to the mistakes in the EIS photometric calibration (see text), in column four the mean is always negative and only the numbers printed in bold face are meaningful.

Field	Filter	Δm [mag]	$\cos(\delta) \cdot \Delta\alpha$ [$''$]	$\Delta\delta$ [$''$]
Deep1a	U_{35060}	$-0.19 \pm \mathbf{0.04}$	0.05 ± 0.12	0.03 ± 0.08
Deep1a	U_{38}	$-0.17 \pm \mathbf{0.04}$	0.04 ± 0.12	0.04 ± 0.08
Deep1a	B	$-0.24 \pm \mathbf{0.05}$	0.00 ± 0.11	0.04 ± 0.08
Deep1a	R	$-0.02 \pm \mathbf{0.04}$	0.00 ± 0.12	0.02 ± 0.08
Deep1a	I	$-0.19 \pm \mathbf{0.05}$	0.00 ± 0.12	0.02 ± 0.08
Deep1b	U_{35060}	$-0.33 \pm \mathbf{0.11}$	0.08 ± 0.11	0.00 ± 0.09
Deep1b	B	$-0.32 \pm \mathbf{0.07}$	0.07 ± 0.10	-0.02 ± 0.07
Deep1b	V	$-0.03 \pm \mathbf{0.03}$	0.04 ± 0.09	-0.03 ± 0.07
Deep1b	R	$-0.12 \pm \mathbf{0.04}$	0.04 ± 0.10	-0.02 ± 0.06
Deep1b	I	$-0.14 \pm \mathbf{0.05}$	0.05 ± 0.09	-0.02 ± 0.06
Deep1c	V	$-0.03 \pm \mathbf{0.09}$	-0.02 ± 0.15	-0.04 ± 0.07
Deep1c	R	$-0.11 \pm \mathbf{0.03}$	-0.04 ± 0.15	-0.04 ± 0.07
Deep2a	R	-0.10 ± 0.08	-0.15 ± 0.12	-0.09 ± 0.13
Deep2b	U_{35060}	-0.02 ± 0.11	-0.12 ± 0.13	0.09 ± 0.11
Deep2b	B	-0.31 ± 0.10	-0.18 ± 0.11	0.04 ± 0.09
Deep2b	V	-0.03 ± 0.09	-0.19 ± 0.11	0.03 ± 0.09
Deep2b	R	-0.10 ± 0.09	-0.20 ± 0.12	0.04 ± 0.10
Deep2b	I	-0.21 ± 0.12	-0.20 ± 0.11	0.03 ± 0.10
Deep2c	U_{35060}	-0.49 ± 0.12	-0.12 ± 0.12	0.07 ± 0.10
Deep2c	V	-0.29 ± 0.13	-0.17 ± 0.13	0.02 ± 0.12

Table 3.4: Comparison of the GaBoDS and the EIS data.

Field	Filter	Δm [mag]	$\cos(\delta) \cdot \Delta\alpha$ [$''$]	$\Delta\delta$ [$''$]
Deep2c	I	-0.13 ± 0.08	-0.16 ± 0.11	0.01 ± 0.09
Deep2d	R	-0.32 ± 0.08	-0.19 ± 0.11	0.03 ± 0.08
Deep3a	U_{35060}	$-0.20 \pm \mathbf{0.04}$	-0.21 ± 0.15	-0.39 ± 0.14
Deep3a	U_{38}	$-0.02 \pm \mathbf{0.04}$	-0.21 ± 0.16	-0.42 ± 0.14
Deep3a	B	$-0.20 \pm \mathbf{0.06}$	-0.23 ± 0.13	-0.42 ± 0.13
Deep3a	V	$0.05 \pm \mathbf{0.05}$	-0.22 ± 0.13	-0.43 ± 0.13
Deep3a	R	$-0.08 \pm \mathbf{0.09}$	-0.24 ± 0.11	-0.43 ± 0.13
Deep3b	U_{35060}	$0.09 \pm \mathbf{0.07}$	-0.24 ± 0.07	-0.31 ± 0.08
Deep3b	B	$-0.07 \pm \mathbf{0.03}$	-0.23 ± 0.07	-0.33 ± 0.06
Deep3b	V	$0.01 \pm \mathbf{0.03}$	-0.23 ± 0.07	-0.33 ± 0.06
Deep3b	R	$-0.08 \pm \mathbf{0.03}$	-0.23 ± 0.08	-0.34 ± 0.06
Deep3b	I	$-0.19 \pm \mathbf{0.05}$	-0.24 ± 0.07	-0.34 ± 0.05
Deep3c	U_{35060}	$-0.32 \pm \mathbf{0.08}$	-0.30 ± 0.09	-0.32 ± 0.07
Deep3c	B	$-0.19 \pm \mathbf{0.04}$	-0.32 ± 0.08	-0.33 ± 0.06
Deep3c	V	$-0.02 \pm \mathbf{0.03}$	-0.32 ± 0.08	-0.34 ± 0.06
Deep3c	R	$-0.06 \pm \mathbf{0.03}$	-0.31 ± 0.08	-0.34 ± 0.06
Deep3c	I	$-0.08 \pm \mathbf{0.04}$	-0.33 ± 0.08	-0.35 ± 0.06
Deep3d	V	$-0.08 \pm \mathbf{0.07}$	-0.37 ± 0.07	-0.34 ± 0.07
Deep3d	I_{EIS}	$-0.10 \pm \mathbf{0.04}$	-0.36 ± 0.06	-0.35 ± 0.06

3.5 Summary

We release 64 fully reduced and stacked images of the ESO Deep Public Survey to the scientific community. These images, reduced with our THELI pipeline, are shown to have good internal and external astrometric and photometric calibrations. A comparison to the EIS-DPS release shows significant offsets in the absolute photometric calibration for most images and large scatter in the magnitude differences of objects for a subset. Both effects are shown to originate from the EIS data and this is acknowledged by the EIS team.

Table 3.5: Released optical images from the DPS reduced with the THELI pipeline. The first three characters of the coaddition ID in column three represent the field (D1A for field Deep1a etc.) and the fourth character is an identifier to distinguish the different coadditions of one field. The Vega limiting magnitudes in column six correspond to 5σ sky level measured in a circular aperture of $2''$ radius. The grades in column seven are estimated from visual inspection of the coadded images. Grade "A" is assigned to an image with no special features and an appearance typical for the particular band. Grade "B" is assigned to an image with cosmetic defects. Nevertheless also the grade "B" images are fully usable for most scientific purposes. Astrometric offsets in columns eight and nine are given with respect to the deepest R-band image of that field.

Field	Filter	coadd ID	exp. time [s]	FWHM [$''$]	mag lim	Grade	$\Delta\alpha\cos(\delta)$ [$''$]	$\Delta\delta$ [$''$]
Deep1a	U_{35060}	D1AA	53 095	1.34	25.45	A	-0.01 ± 0.16	-0.01 ± 0.16
Deep1a	U_{35060}	D1AC	71 093	1.44	25.57	A	-0.01 ± 0.15	-0.01 ± 0.15
Deep1a	U_{38}	D1AA	17 398	1.28	24.46	A	-0.00 ± 0.14	-0.01 ± 0.14
Deep1a	B	D1AA	11 696	1.34	26.38	A	-0.00 ± 0.14	-0.01 ± 0.14
Deep1a	V	D1AA	8684	0.98	25.76	A	0.01 ± 0.12	-0.02 ± 0.12
Deep1a	R	D1AA	9597	0.85	25.50	A	-	-
Deep1a	I	D1AA	31 493	0.84	24.33	A	0.00 ± 0.12	-0.00 ± 0.11
Deep1a	I	D1AG	25 194	0.84	24.29	A	0.00 ± 0.12	-0.00 ± 0.11
Deep1b	U_{35060}	D1BA	46 793	1.06	25.27	A	-0.03 ± 0.15	-0.01 ± 0.15
Deep1b	B	D1BA	9597	1.33	26.09	A	-0.02 ± 0.15	-0.01 ± 0.14
Deep1b	B	D1BS	4199	1.09	25.61	A	-0.02 ± 0.13	-0.00 ± 0.13
Deep1b	V	D1BA	9897	1.32	25.60	A	-0.00 ± 0.14	0.00 ± 0.13

Table 3.5: Released optical images from the DPS reduced with the THELI pipeline.

Field	Filter	coadd ID	exp. time [s]	FWHM [$''$]	mag lim	Grade	$\Delta\alpha\cos(\delta)$ [$''$]	$\Delta\delta$ [$''$]
Deep1b	V	D1BS	4499	1.20	25.24	A	0.00 ± 0.13	-0.00 ± 0.12
Deep1b	R	D1BA	19 794	1.28	25.70	B	-	-
Deep1b	R	D1BS	8398	1.10	25.28	A	0.00 ± 0.10	0.00 ± 0.09
Deep1b	I	D1BA	25 493	0.97	24.38	A	-0.01 ± 0.13	-0.00 ± 0.12
Deep1c	U_{35060}	D1CB	45 253	1.09	25.07	A	-0.05 ± 0.16	-0.01 ± 0.14
Deep1c	B_{123}	D1CA	13 796	1.27	26.56	A	0.02 ± 0.12	0.04 ± 0.10
Deep1c	V	D1CB	11 997	1.10	25.78	A	-0.00 ± 0.06	0.02 ± 0.07
Deep1c	R	D1CA	11 696	0.97	25.28	A	-	-
Deep1c	I_{EIS}	D1CB	32 937	1.21	24.83	A	0.02 ± 0.10	0.05 ± 0.10
Deep1e	U_{38}	D1EA	9899	1.78	24.34	A	-0.01 ± 0.09	-0.03 ± 0.21
Deep1e	R	D1EA	8998	0.90	25.37	A	-	-
Deep1e	I	D1EA	11 398	1.26	23.93	A	-0.01 ± 0.07	-0.02 ± 0.15
Deep1f	I	D1FA	14 997	1.16	23.92	A	-	-
Deep2a	R	D2AA	5998	0.83	25.11	A	-	-
Deep2b	U_{35060}	D2BA	53 095	1.15	25.23	A	-0.02 ± 0.14	-0.01 ± 0.14
Deep2b	U_{35060}	D2BC	64 794	1.27	25.30	A	-0.02 ± 0.14	-0.01 ± 0.14
Deep2b	B	D2BA	11 396	0.98	26.23	A	0.00 ± 0.13	0.00 ± 0.13
Deep2b	B	D2BS	9597	0.91	26.13	A	0.00 ± 0.13	0.00 ± 0.13
Deep2b	V	D2BA	9297	0.88	25.59	A	-0.00 ± 0.12	0.01 ± 0.12
Deep2b	R	D2BA	10 497	1.33	25.42	A	-	-

Table 3.5: Released optical images from the DPS reduced with the THELI pipeline.

Field	Filter	coadd ID	exp. time [s]	FWHM [$''$]	mag lim	Grade	$\Delta\alpha\cos(\delta)$ [$''$]	$\Delta\delta$ [$''$]
Deep2b	R	D2BS	4799	1.09	25.01	A	0.00 ± 0.08	0.00 ± 0.08
Deep2b	I	D2BA	19 196	0.75	24.25	A	0.00 ± 0.12	0.01 ± 0.11
Deep2c	U_{35060}	D2CB	78 891	1.01	25.49	A	-0.02 ± 0.13	0.03 ± 0.14
Deep2c	B	D2CA	69 431	0.98	27.28	A	0.01 ± 0.11	-0.00 ± 0.11
Deep2c	V	D2CB	104 603	0.92	26.77	A	0.01 ± 0.10	0.00 ± 0.10
Deep2c	R	D2CB	87 654	0.79	26.54	A	-	-
Deep2c	I	D2CA	34 575	0.93	24.49	A	0.00 ± 0.11	0.00 ± 0.11
Deep2d	R	D2DA	2999	1.06	24.81	A	-	-
Deep3a	U_{35060}	D3AA	26 997	1.09	24.77	A	-0.01 ± 0.13	-0.01 ± 0.13
Deep3a	U_{35060}	D3AC	35 996	1.10	24.91	A	-0.01 ± 0.13	-0.00 ± 0.13
Deep3a	U_{38}	D3AA	26 997	1.27	24.59	A	-0.01 ± 0.12	-0.00 ± 0.12
Deep3a	B	D3AA	11 096	0.92	26.08	A	-0.00 ± 0.12	-0.00 ± 0.11
Deep3a	V	D3AA	8998	1.02	25.44	A	-0.00 ± 0.11	-0.00 ± 0.11
Deep3a	R	D3AA	8997	0.81	25.28	A	-	-
Deep3a	R	D3AS	7798	0.79	25.23	A	-0.00 ± 0.05	0.00 ± 0.05
Deep3a	I	D3AA	23 396	0.95	24.22	B	0.01 ± 0.11	0.00 ± 0.11
Deep3a	I	D3AG	13 198	0.91	23.94	A	0.00 ± 0.11	0.00 ± 0.11
Deep3b	U_{35060}	D3BA	53 995	1.01	25.13	A	0.01 ± 0.13	-0.00 ± 0.13
Deep3b	B	D3BA	9897	0.95	26.06	A	0.00 ± 0.12	-0.00 ± 0.12
Deep3b	V	D3BA	8998	0.88	25.45	A	0.00 ± 0.11	0.00 ± 0.11

Table 3.5: Released optical images from the DPS reduced with the THELI pipeline.

Field	Filter	coadd ID	exp. time [s]	FWHM [$''$]	mag lim	Grade	$\Delta\alpha\cos(\delta)$ [$''$]	$\Delta\delta$ [$''$]
Deep3b	R	D3BA	9297	0.79	25.30	A	-	-
Deep3b	I	D3BA	26 993	0.82	24.27	A	0.00 ± 0.11	0.01 ± 0.10
Deep3c	U_{35060}	D3CA	46 845	0.97	25.22	A	0.01 ± 0.13	-0.00 ± 0.13
Deep3c	B	D3CA	13 496	0.93	26.24	A	0.01 ± 0.12	0.00 ± 0.12
Deep3c	B	D3CS	11 996	0.90	26.17	A	0.01 ± 0.12	-0.00 ± 0.12
Deep3c	V	D3CA	5998	0.79	25.30	A	0.01 ± 0.11	-0.00 ± 0.10
Deep3c	R	D3CA	8998	0.81	25.26	A	-	-
Deep3c	I	D3CA	25 193	1.02	24.04	B	0.01 ± 0.11	0.00 ± 0.11
Deep3d	B	D3DA	11 097	0.88	26.22	A	0.01 ± 0.12	-0.01 ± 0.12
Deep3d	V	D3DA	8998	0.90	25.37	A	0.01 ± 0.11	-0.00 ± 0.10
Deep3d	R	D3DA	8998	0.73	24.76	A	-	-
Deep3d	I_{EIS}	D3DA	22 293	0.76	24.51	A	0.01 ± 0.10	-0.00 ± 0.10

4

Photometric redshifts[1]

In Sect. 1.2 it was described how the observable redshift of a galaxy is related to its distance. Due to this connection the measurement of galaxy redshifts is one of the most important tools in extragalactic astronomy. Usually a spectrum of a galaxy is needed to measure a precise redshift. With today's multi-object spectrographs on 8m-class telescopes this can be done for thousands of galaxies on a reasonable timescale. However, imaging with a wide-field camera in optical broad-band filters is still superior by a factor of 100-1000 in terms of objects per unit telescope time. Therefore, over the past years considerable effort was put into the estimation of approximate galaxy redshifts from these broad-band colours alone without taking spectra. This highly efficient method called *photometric redshift* (photo-z) is described in this chapter.

4.1 Photo-z techniques

There are basically two different approaches to estimate a photometric redshift for a galaxy from its broad-band colours, the "SED-fitting"-method and the "empirical training set"-method. The former relies on a sample of synthetic or observed spectral-energy-distribution (SED) templates and on theoretical knowledge how those SEDs evolve with redshift. The latter relies on a colour catalogue of spectroscopically observed galaxies as large as possible to cover essentially every galaxy type at all redshifts. See e.g. Koo (1999) or Benítez (2000) for detailed reviews of both techniques and their differences.

---

[1] This chapter is partly based on the manuscript Hildebrandt et al. (2007c), submitted to A&A.

The empirical approach (for an example using artificial neural networks see Collister & Lahav 2004) can lead to very precise results if an extensive, complete spectroscopic catalogue with colour information is available. But it is not as flexible as the "SED-fitting"-method because for every new filter set or camera the colour catalogue must be recreated. Moreover, it is essentially limited to the magnitude range where spectra can be taken in large numbers ($I \lesssim 24$).

The "SED-fitting"-method, however, can be applied to every dataset for which the filter transmission curves are known and this is the reason why we concentrate on this approach here. It is described in the following Sect. 4.2.

In practice, a photo-z analysis often involves aspects from both approaches. Empirical colour-redshift relations can certainly be extrapolated in magnitude or redshift. Also, a spectroscopic catalogue can help to optimise parts of an "SED-fitting" approach. Ilbert et al. (2006), e.g., present a method to improve the photo-z estimates in the CFHT Legacy Survey. They adjust the photometric zeropoints of their images and optimise the template SEDs with help of more than 3000 spectroscopically observed galaxies in the range $0 < z < 5$. The optimisation of templates was already used for improving template based photo-z estimates in the SDSS (Csabai et al. 2003). Gabasch et al. (2004) claim to obtain highly accurate photo-z's in the FORS Deep Field (FDF) by constructing semi-empirical template SEDs from 280 spectroscopically observed galaxies in the FDF and the Hubble Deep Field (HDF).

4.2 The SED-fitting method

This method rests on the assumption that all galaxies in the considered redshift range can be represented in their restframe by a number of spectral energy distributions (SEDs). One basically distinguishes between empirical template sets which are derived from spectroscopic observations and synthetic template sets created from stellar population models.

4.2.1 Template sets

The empirical template set by Coleman et al. (1980) is shown in Fig. 4.1. They observed local galaxies, classified them into four spectral types, and reported their mean SEDs. This template set is the benchmark for many studies on the HDF and is therefore still widely used today for comparisons.

Synthetic template sets are created by first choosing an initial mass function with a particular metallicity. A composite stellar population is built by considering different star-formation histories. It can be shown theoretically that under the assumptions of gas mass conservation and instantaneous recycling of stellar ejecta the star formation rate, $\psi(t)$, takes the form of a decaying exponential, $\psi(t) \propto \exp(-t/\tau)$, with τ being the characteristic timescale of the star-formation process. It is clear that the star-formation history of real galaxies is usually much more complicated. But it is justified by the success of SED based photo-z's to assume that the SEDs created in this way give a fair representation of the SEDs of real galaxies. Creating these composite stellar populations for different ages yields a large number of templates for the photo-z estimation. In Fig. 4.2 the synthetic template set created from the spectral evolutionary models of Bruzual & Charlot (1993) and used by the *Hyperz* photo-z code (Bolzonella et al. 2000) is shown. The synthetic approach has the advantage that more information than the redshift alone can be extracted from the maximum likelihood fit, e.g., stellar mass, age, star-formation rate etc. The accuracy of these other physical quantities is, however, hard to determine and depends on a large variety of parameters.

The template set can be further enlarged by applying different internal reddenings to each template according to empirical dust extinction laws; see Fig. 4.3.

It is evident from Figs. 4.2 & 4.3 that the effects of different ages, different star-formation histories, and different reddenings can be degenerate.

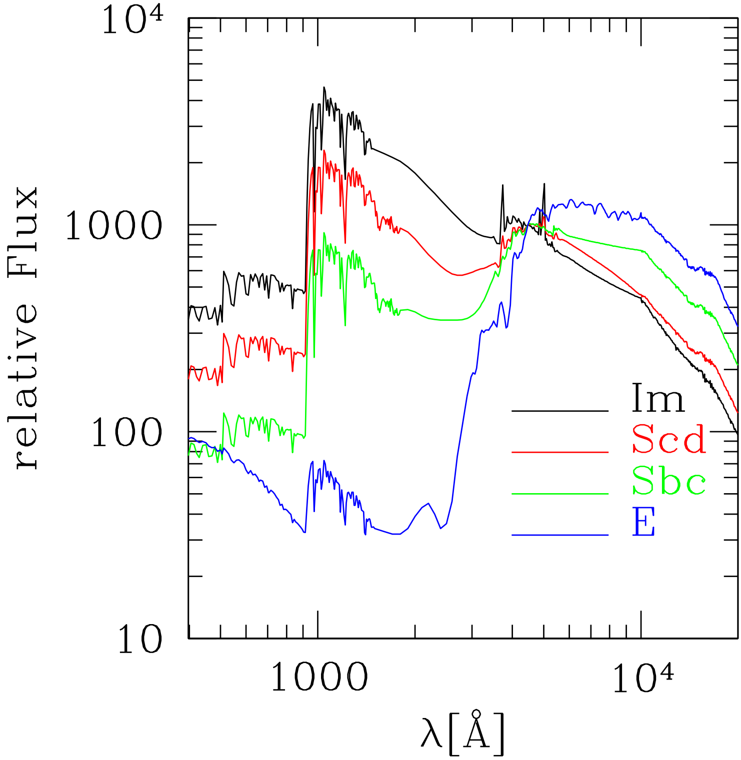

Figure 4.1: Observed SEDs by Coleman et al. (1980). These are the mean SEDs of local galaxies of different types and they are widely used in the community.

4.2.2 Photo-z estimation

Once the template set is defined, the templates are redshifted stepwise with a particular step size (some codes use a variable step size increasing with redshift or apply a refining of the step size around χ^2 minima) and at each redshift step they are convolved with the known

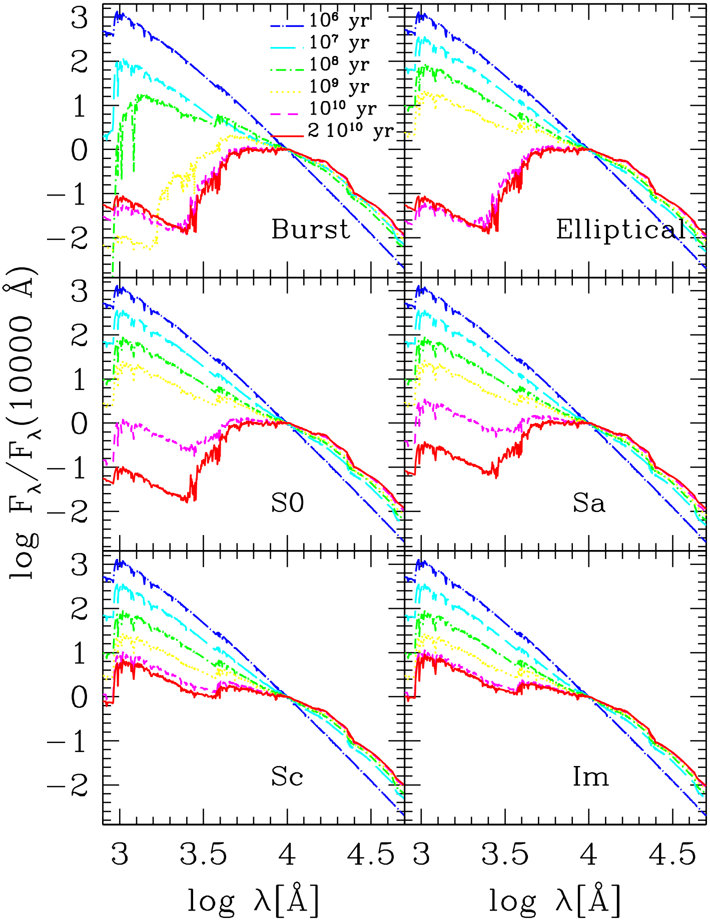

Figure 4.2: Synthetic SEDs of different spectral types and different ages created from the spectral evolutionary models of Bruzual & Charlot (1993) and used by the *Hyperz* photo-z code (Bolzonella et al. 2000). The different spectral types correspond to different star-formation histories ($\tau = 0, 1, 2, 3, 15, \infty$ Gyr for Burst, Elliptical, S0, Sa, Sc, and Im, respectively). Taken from the *Hyperz* manual.

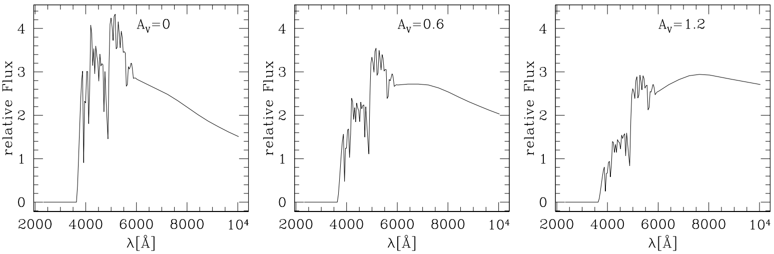

Figure 4.3: Template spectra of $z = 3$ star-forming galaxies with three different amounts of dust reddening calculated with the reddening law by Calzetti et al. (2000).

spectral response of the imager to yield flux estimates. Additionally, some photo-z codes apply a redshift-dependent modification of the templates before convolution according to the known influence of the Lyman-α forest (e.g., with the recipe from Madau 1995). This is especially important for very high-redshift galaxies where most of the flux shortward of the Lyman-break is suppressed (see Sect. 4.5).

At each redshift step and for every template a χ^2 minimisation is carried out by minimising the scaling parameter b in the following equation,

$$\chi^2(z, \mathrm{temp}) = \sum_{i}^{N_\mathrm{filter}} \left[\frac{F_{\mathrm{obs},i} - b \cdot F_{\mathrm{temp},i}(z)}{\sigma_i} \right]^2 , \tag{4.1}$$

with $F_{\mathrm{obs},i}$ being the observed flux in the filter i, $F_{\mathrm{temp},i}(z)$ being the modelled flux at redshift z in the filter i, and σ_i being the error of the observed flux in the filter i. By doing so, a two-dimensional χ^2 matrix is created with one redshift and one template axis. Taking the best-fit template and the probability associated with the χ^2 value of this best-fit template at each redshift step one can produce redshift vs. probability, $p(z)$, plots as shown in Fig. 4.4. Assuming Gaussianity for the photometric errors one can easily calculate the redshift likelihood L from the χ^2 at the different redshift steps (see

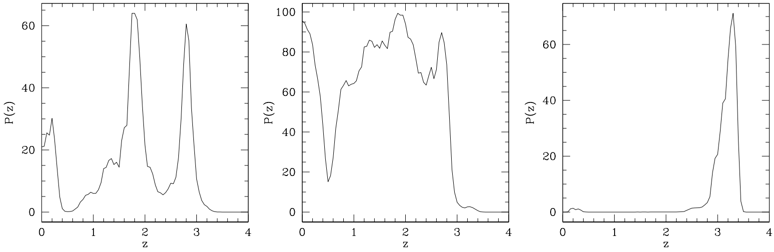

Figure 4.4: Redshift vs. probability for three different U-dropout galaxies (see Sect. 4.5) from the sample presented in Chap. 6. *Left:* An object with assigned redshift phot-z = 1.81. This example illustrates that the assignment of a single number for the photometric redshift can be misleading. The peak at $z = 2.8$ has nearly the same probability. *Middle:* An object with assigned redshift phot-z = 1.85. This example illustrates that sometimes the photometric redshift estimation totally fails but nevertheless a primary solution is put out. *Right:* An object with assigned redshift phot-z = 3.29. The ideal case of an object with a definite single redshift estimate.

e.g. Barlow 1989),

$$\ln L(z, \text{temp.}) = -\frac{1}{2}\chi^2(z, \text{temp.}) - \sum \ln \sigma_i \sqrt{2\pi}\,. \tag{4.2}$$

The standard and most basic way to assign a photo-z to a galaxy at this point is to take the one with the lowest χ^2 value, hence the most likely redshift and template. A different, more sophisticated way to assign a redshift is to locate confidence regions in the χ^2 landscape corresponding to $\Delta\chi^2$ thresholds and to estimate e.g. a weighted mean inside this peak. The details are dependent on the specific photo-z code that is used. Also the χ^2 minimisation may be slightly different for different codes with some codes using magnitudes or colours instead of fluxes during the minimisation.

4.3 Problems of SED-based photo-z's

In general there are two important shortcomings of the SED-fitting technique:

- biased photo-z estimates
- catastrophic outliers.

4.3.1 Bias in photo-z estimates from SED-fitting

It is well-known that photo-z's estimated by SED-fitting in contrast to empirical approaches are prone to be biased so that $\langle z_{\mathrm{spec}} - z_{\mathrm{phot}} \rangle \neq 0$, with $\langle ... \rangle$ denoting an ensemble average. The method relies on the perfect match between the photometric calibration of the templates and the absolute photometric calibration of the imaging dataset. Often this requirement is only poorly fulfilled leading to a systematic over- or underestimation of the redshift. The situation is usually worse for ground-based data in comparison to space-based data because of the Earth's atmosphere complicating photometric calibration and colour measurements. With an appropriate spectroscopic training set this bias can be decreased or even totally removed by means of e.g. photometric re-calibration which will be shown in Sect. 5.4. Certainly, this represents a loss of flexibility which is comparable to the disadvantage of the purely empirical techniques discussed above.

4.3.2 Catastrophic outliers

A catastrophically wrong assignment of an object's photo-z can happen if colour-redshift degeneracies exist, i.e. if for the given filter set there are very different redshift ranges for which the colours of different templates are similar. For example, when using just optical filters the colours of an elliptical galaxy at low to intermediate redshift can resemble the colours of a star-forming galaxy at high redshift. This particular degeneracy is due to the fact that the Balmer-/4000Å-break cannot be clearly distinguished from the Lyman-break without near-IR information. The occurrence of these colour-redshift degeneracies is certainly more serious if fewer filters are used, the total wavelength coverage is decreased, or photometric errors are increased.

The points in multi-dimensional colour-space that correspond to one specific template at the different redshift steps define a curve. For each template there is one such redshift curve. In Fig. 4.5 this is illustrated for a two-dimensional colour-space and a template set made of five SEDs. An observed galaxy is represented by a point in colour-space and an error region around this point. If the galaxy is located very near to two such curves or to an intersection of two curves in colour-space one cannot decide which curve gives the better fit to the object's colours. Such a situation is usually characterised by bimodal or multimodal $p(z)$ distributions; see the left and middle panels of Fig. 4.4. The danger of assigning a catastrophic wrong redshift is large in such cases.

One method to reduce these degeneracies and the resulting catastrophic photo-z failures is to introduce Bayesian priors. There is usually more information available about a galaxy than just its colours. It is well-known that some of these quantities (e.g. the apparent magnitude or the apparent size) are not evenly distributed in redshift. By applying an empirical or theoretical prior one can put more or less weight on different peaks in the redshift likelihood function, and it can be shown that in such a way one reduces the rate of catastrophic outliers considerably (see e.g. Benítez 2000; Ilbert et al. 2006).

For the example of an apparent magnitude prior and if we consider just one template, Bayes' theorem is applied in the following way. It is used to connect the likelihood defined above (see Eq. 4.2) with the prior probability to yield the posterior probability. For a data vector $D = \{C, m\}$, including colours C and the apparent magnitude m in a chosen reference filter, let $p(C|z) \equiv L(z)$ be the likelihood, i.e. the probability of observing the colours C if the galaxy is located at redshift z. Furthermore, a prior probability $p(z|m)$ is needed, which is just the redshift distribution of the galaxies with apparent magnitude m in the reference filter. Multiplying the likelihood with the prior probability yields an estimate for the posterior probability

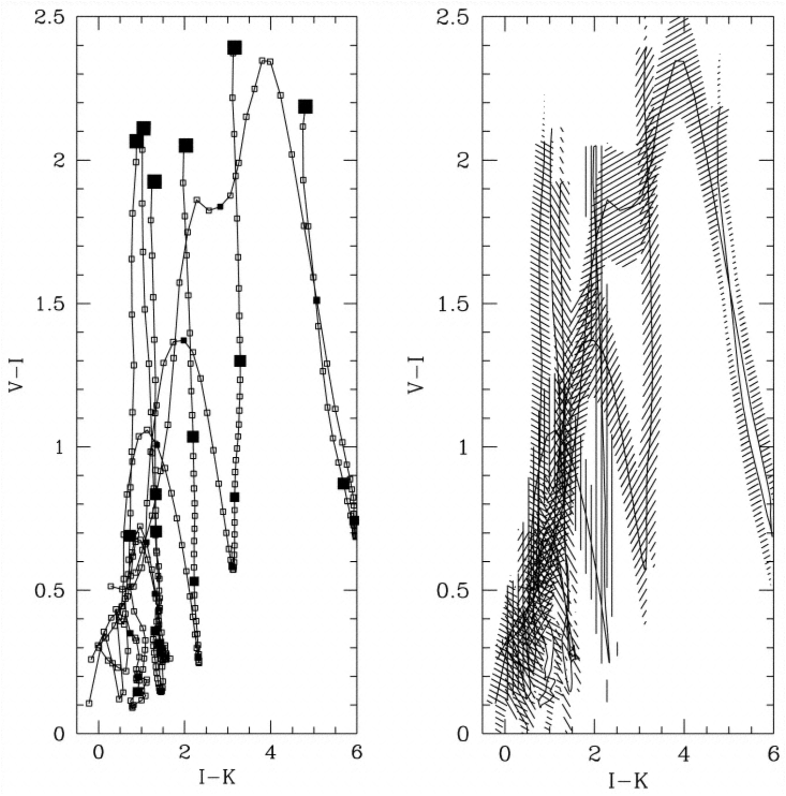

Figure 4.5: *Left:* $V-I$ vs. $I-K$ colours for five different templates in the redshift range $1 < z < 5$. The filled squares correspond to redshift steps of $\Delta z = 1$ and the size of these squares increases with redshift. For this filter set every crossing of two lines corresponds to a colour-redshift degeneracy. *Right:* If photometric errors are added (here: 0.2mag) the regions are broadened and the degeneracies become more serious. In reality, one would not dare to estimate a photo-z from this 3-filter set in general over the range $1 < z < 5$. However, three filters can be sufficient to select galaxies in certain redshift ranges with unique colours (see Sect. 4.5). Taken from Benítez (2000).

$p(z|D)$, i.e. the probability of a galaxy being located at redshift z given the data D:

$$p(z|D) = p(z|C, m) = \frac{p(z|m)p(C|z)}{p(C)} \, . \tag{4.3}$$

The normalisation constant $p(C)$ called evidence is usually not needed for the relative comparison of redshift probabilities. This is Bayes' theorem which is valid in general, relating the conditional and marginal (prior) probabilities of two stochastic events. Here, the theorem, for the particular case of photo-z's with an apparent magnitude prior, relates the conditional probability of observing a galaxy with colours C given a redshift z with the prior probability of observing a galaxy of apparent magnitude m.

For more than one template we need more than one prior and these priors should be scaled with respect to each other to represent the true prior information about the relative observed fraction of different galaxy types. The mode of operation of this technique is further illustrated in Fig. 4.6.

4.4 Photo-z codes

In the following we describe the three different photo-z codes which are used for this study.

4.4.1 Hyperz

The "SED-fitting" photo-z code *Hyperz* (Bolzonella et al. 2000) is publicly available[2], well documented and widely used by the community. For detailed information on the code see the manual at the website or the reference paper mentioned above.

Hyperz comes with two different template SED sets, the mean observed spectra of local galaxies by Coleman et al. (1980), hereafter CWW (see Fig. 4.1), and synthetic spectra created from the spectral

---

[2] `http://webast.ast.obs-mip.fr/hyperz/`

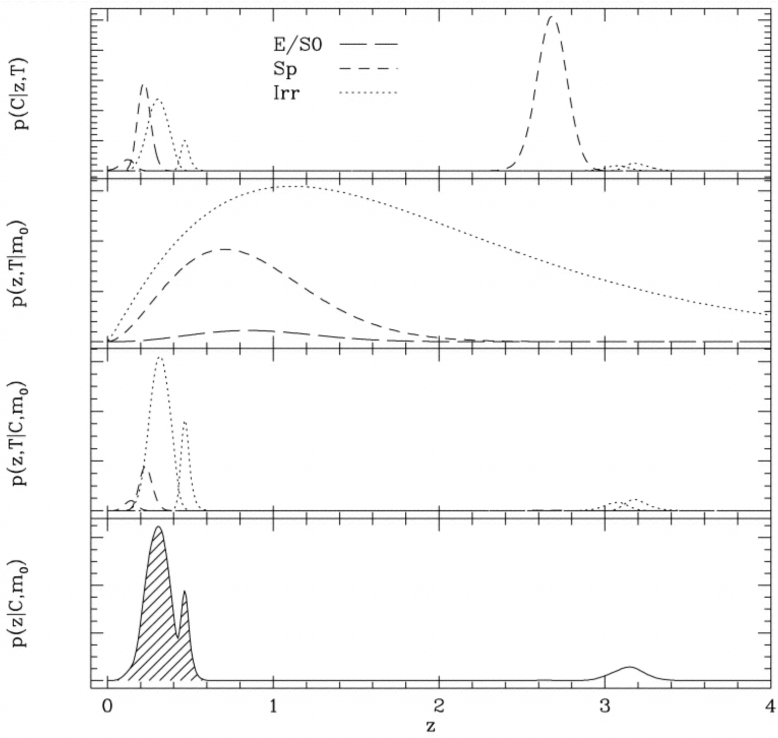

Figure 4.6: From top to bottom: (a) likelihood function for three (for simplicity) different templates, (b) prior probability for the three templates at an apparent magnitude of m_0, (c) posterior probability distributions for the three templates separately, (d) final posterior probability after summation over the three templates. Taken from Benítez (2000).

evolution library of Bruzual & Charlot (1993), hereafter BC (see Fig. 4.2). We use the BC templates for *Hyperz* since for all tested setups the performance with the CWW templates is worse. Different reddening laws are implemented to account for the effect of interstellar dust on the spectral shape. By default we use the reddening law of Calzetti et al. (2000) derived for local star-forming galaxies. The damping of the Lyman-α-forest which increases with redshift is modelled according to Madau (1995).

Another important option influencing performance strongly is the application of a prior on the absolute magnitude. Once the most likely redshift of an object is estimated by the code, the absolute magnitude in a reference filter is calculated from the apparent magnitude in that filter for a chosen cosmology. The user can specify limits to exclude unrealistically luminous or under-luminous objects. In the following we allow galaxies to have an absolute I-band magnitude of $M_* - 2.5 < I_{\rm abs} < M_* + 2.5$ using the local SDSS-value of $M_{*,\rm AB} = -21.26$ from Blanton et al. (2001). M_* is the characteristic absolute magnitude of a galaxy in the Schecter luminosity function (Schechter 1976). An evolution of M_* with redshift due to the K-correction or due to an evolution of the luminosity function is not featured by *Hyperz*.

Besides reporting the most likely redshift estimate as a primary solution, *Hyperz* can also store the redshift probability distribution (see Fig. 4.4). Furthermore, the width of this distribution around the primary solution provides a confidence interval, which allows the user to identify objects with very uncertain estimates.

4.4.2 COMBO-17 code

The photo-z code of COMBO-17 performs two simultaneous tasks: it classifies objects into stars, galaxies, QSOs and white dwarfs based on their colours, and for galaxies and QSOs it also estimates redshifts. Here, we used a setup skipping the classification and assuming a galaxy a priori in order to better compare the results to the other codes which also assume that all objects are galaxies. The code is currently not publicly available.

It uses a set of synthetic templates produced with the PEGASE population synthesis code (Fioc & Rocca-Volmerange 1997) and an SMC reddening law. For all details on how the templates were created we refer the reader to Wolf et al. (2004).

The code determines the redshift probability distribution $p(z)$ and tests the shape of $p(z)$ for bimodality. In case of no bimodality it reports the mean of this distribution as a Minimum-Error-Variance (MEV) redshift and its width as an error estimate. In bimodal cases the redshift axis is split into two intervals delimited at the MEV redshift which is most probably a poor estimate. Then two unimodal solutions are estimated in the two intervals and the one with the larger relative probability computed from the integral of $p(z)$ in the interval is put out (for all details see Wolf et al. 2001).

4.4.3 BPZ

BPZ stands for Bayesian Photometric Redshifts and is a public code[3] which implements the method described in Benítez (2000). It is an SED fitting method combined with a redshift/type prior, $p(z, T|m)$, which depends on the observed magnitude of the galaxies. It originally used a set of 6 templates formed by the 4 CWW set, and two starburst templates from Kinney et al. (1996) which were shown to significantly improve the photo-z estimation. It should be stressed that the extrapolation to the far UV and IR of the optical CWW templates used by *BPZ* is quite different from the one used by *Hyperz*. For example, the depression of the Lyman-α forest shortward of the Lyman-break is to some extent incorporated in the *BPZ* version of the template set unlike the one used by *Hyperz* whose SEDs are later altered by the Madau (1995) formula. The *BPZ* template library has been calibrated using a set of HST and other ground-based observations as described in Benítez et al. (2004). This template set has been shown to remarkably well represent the colours of galaxies in HST observations, to the point of being able to photometrically calibrate the NIC3 Hubble UDF observations with a 0.03 magnitude error as shown in Coe et al. (2006). In the latter paper two additional, very blue templates from the Bruzual & Charlot library were introduced, so the current *BPZ* library contains 8 templates.

---

[3] `http://acs.pha.jhu.edu/~txitxo/`

The redshift likelihood is calculated by *BPZ* in a similar way as by *Hyperz* minimising the χ^2 of observed and predicted colours. However, in contrast to *Hyperz* no reddening is applied to the templates relying on the completeness of the given set and the "automatic" interpolation of the maximum likelihood technique. After the calculation of the likelihood, Bayes theorem is applied incorporating the prior probability (see Sect. 4.3.2). The actual shape of this prior is dependent on template type and I-band magnitude and was derived from the observed redshift distributions of different galaxy types in the Hubble Deep Field. For details on the procedure see also Benítez (2000). The posterior redshift probability distribution is integrated over an interval around the most probable redshift, $z_{\rm phot}$, to get an estimate of the redshift reliability. This probability is called the *ODDS* parameter. *BPZ* has been extensively used in the ACS GTO program, the GOODS survey and others.

4.5 The Lyman-break technique

A special case of the photo-z technique described above is the selection of high-redshift, star-forming galaxies by their pronounced Lyman-break, a discontinuity in the spectral energy density by an order of magnitude at 912Å caused by neutral hydrogen in the stellar atmospheres and the interstellar medium of the galaxy. At higher redshifts the Lyman-break is further enhanced by neutral hydrogen clouds in the intergalactic medium along the sightline. For redshifts $z > 2$ the Lyman-break enters the optical window (see Fig. 4.7) and the U-band flux drops rapidly. Due to this fact, star-forming galaxies at $z \sim 3$ are usually not observable in the U-band anymore and thus are called U-dropouts. For the selection of these Lyman-break galaxies three optical filters are sufficient. The definition of efficient selection criteria in two-colour space and the scientific exploitation of the selected samples is covered in Chap. 6.

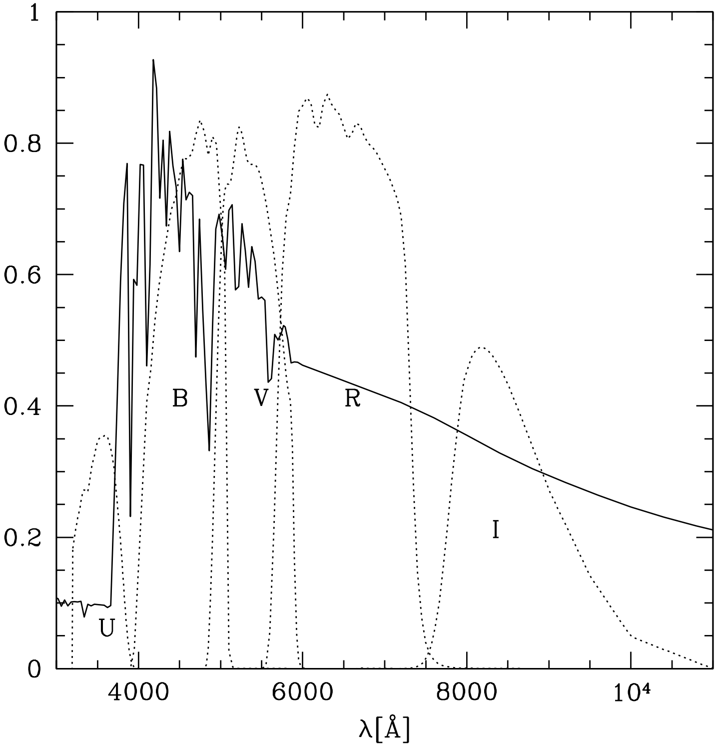

Figure 4.7: Spectrum of a $z = 3$ starburst galaxy and the instrumental response of the WFI@ESO/MPG2.2m. Damping by the Ly-α forest is not taken into account here. Such a galaxy has a very red $U - B$ or $U - V$ colour and a blue $V - R$ colour so that it can be selected easily from three-filter data (see Sect. 6.1.4).

5

Photometric redshifts in comparison[1]

Photometric redshifts have become a standard tool for the observing astronomer in the past years. Not only are large multi-colour imaging surveys planned and executed with the goal of estimating the redshift of as many galaxies as possible from their broad-band photometry, but also many smaller projects benefit from this technique by providing redshifts that are much cheaper in terms of telescope time than spectroscopic ones and may go deeper.

Large survey projects often use abundant manpower and expertise to create a specialised photometric redshift code from scratch or adapt an existing one to optimally exploit the colour information. Supported by large spectroscopic redshift catalogues these projects can improve the match between their photometry and templates, which can lead to very precise results. However, many authors apply the photometric redshift technique to smaller datasets and do not have the resources to improve their redshifts in the sophisticated ways adopted by large survey teams. Usually, a publicly available photometric redshift code is used in combination with filter transmission curves from an instrument webpage and with the available standard photometry.

The typical benchmarks that come with a photometric redshift code often refer to extremely deep, space-based datasets such as the Hubble Deep Field and do not necessarily help much in anticipating the accuracy on shallower ground-based data with a different filter set. In this chapter we give estimates of photometric redshift accuracy for projects that use publicly available tools from the shelf. We also compare them against the performance achieved after a re-

---

[1] This chapter is largely based on the paper Hildebrandt et al. (2007c), submitted to A&A.

calibration using massive spectroscopic comparisons. We analyse the dependence of the accuracy on magnitude, depth of the data, filter coverage, redshift region and code settings, on real spectroscopic data.

The chapter is organised as follows. In Sect. 5.1 the imaging and spectroscopic datasets are presented. The photometric redshift codes used for this study were already described in Sect. 4.4. In Sect. 5.2 we outline our approach for describing photometric redshift accuracy. The results are presented and discussed in Sect. 5.3. The following Sect. 5.4 deals with re-calibrations for enhancing performance in the presence of a spectroscopic catalogue. A final summary and general conclusions are given in Sect. 5.5.

5.1 Datasets

We investigate the performance of photometric redshifts on three different imaging datasets:

1. We use five-colour $UBVRI$ data from the ESO DPS field Deep2c centred on the CDFS which were observed with WFI at the 2.2m telescope at La Silla, Chile, reduced with the THELI reduction pipeline (Erben et al. 2005), and described in detail in Chap. 3 and Hildebrandt et al. (2006) as part of the Garching-Bonn Deep Survey (GaBoDS).

2. On the same field and taken with the same camera there are catalogues available from the COMBO-17 survey covering the same broad-band filters in $BVRI$ to considerably shallower depth, a different U-band filter, and 12 additional medium-band filters in the optical wavelength range. These data are described in detail in Wolf et al. (2004).

3. Furthermore, we use catalogues from the FORS Deep Field (FDF; see Heidt et al. 2003; Gabasch 2004) involving eight broad-band filters, $UBgRIZJK_{\mathrm{s}}$, observed with FORS@VLT in the optical and SOFI@NTT in the near-infrared.

Spectroscopic catalogues are publicly available for both fields. The VVDS team carried out an $I_{AB} < 24$ magnitude limited spectroscopic survey on the CDFS with VIMOS@VLT (Le Fèvre et al. 2004) yielding 1599 redshifts, and the FDF team measured the redshifts of 355 objects pre-selected by photometric redshifts with FORS@VLT (341 of which are published in Noll et al. 2004).

5.1.1 Imaging data and photometric catalogues

GaBoDS imaging data in the CDFS

This dataset is described in Chap. 3 and the photo-z results for these data can be regarded as representative for very deep ground-based wide-field surveys with the typical photometric accuracy achievable for multi-chip camera, multi-epoch data.

In order to measure unbiased object colours the $BVRI$ images were convolved to the seeing of the U-band which has the worst seeing ($\approx 1\farcs0$). The photometric catalogue was created with *SExtractor* (Bertin & Arnouts 1996) in dual-image-mode with the unconvolved R-band image as the detection image.

COMBO-17 imaging data in the CDFS

The $BVRI$ images of the CDFS in the COMBO-17 survey were created from a subset of the raw data used for the reduction of the GaBoDS images while the U-band image uses a different filter than the one used for the GaBoDS reduction. In terms of exposure time the COMBO images are shallower by a factor of 2.5 (R-band) to 12.5 (V-band) corresponding to approximately 0.4-1.7 magnitudes. Additionally, the field was observed in 12 shallower medium-band filters covering the whole wavelength-range of CCD sensitivity.

The broad-band data from COMBO-17 resemble a medium-deep wide-field survey, while the full 17-filter data are presently unique in its kind. However, we can use them to investigate whether additional telescope time should be spent on increasing depth as in GaBoDS or on obtaining additional SED information as in COMBO-17.

In contrast to GaBoDS, the COMBO-17 photometry was measured directly on unconvolved images. The photometry was obtained in Gaussian apertures whose width was adapted to compensate seeing variations between the frames. Provided the convolution of aperture and PSF yields the same result for each frame, this procedure is mathematically identical to convolving all frames to a constant seeing and extracting fluxes with Gaussian apertures.

FDF imaging data

Heidt et al. (2003) released a photometric catalogue from the FDF which goes very deep in $UBgRIJK_\mathrm{s}$. We also included FORS Z-band data described in Gabasch (2004). At least in the optical, these data are representative of very deep pencil-beam surveys achievable with present day large telescopes. The FDF photometric catalogue contains flux measurements in apertures of different sizes obtained after convolving images to the same PSF. In the following, we use fluxes in aperture diameters of $d = 1\farcs5$. Furthermore, we re-calibrate the data with corrections to the zeropoints derived by the FDF team for improved photometric redshifts.

Comparisons of imaging data

In all three datasets, the multi-band fluxes of a given object were effectively measured in identical physical apertures (with the effect of the atmosphere taken into account and with identical spatial weighting) for all filters, assuming that seeing produces a Gaussian-shaped PSF. Colours could still be biased by non-Gaussianity of the PSF and by suboptimal background subtraction.

The properties of these three imaging datasets are summarised in Table 5.1. Since the limiting magnitudes are estimated in completely different ways in the three data release papers, we decided to calculate hypothetical 10-σ limiting magnitudes with the GaBoDS values as a reference. For GaBoDS the limiting magnitudes are calculated by (see e.g. Steidel et al. 2003):

$$mag_{\mathrm{lim}} = \mathrm{ZP} - 2.5\log\left(\sqrt{N_{\mathrm{pix}}}\cdot\sigma\right)\ . \tag{5.1}$$

ZP is the photometric zeropoint of the image, N_{pix} is the number of pixels in the aperture, and σ gives the global RMS pixel-to-pixel fluctuations of the sky background in the image considered. σ is inverse proportional to the square root of the number of photons from an object and thus inverse proportional to the square root of the area of the telescope's primary mirror and to the square root of the exposure time. Furthermore, for point-like objects N_{pix} is proportional to the area of the seeing disk. Thus, $m_{\mathrm{lim,eff}}$ can be calculated by

$$m_{\mathrm{lim,eff,X}} - m_{\mathrm{lim,G}} = -2.5\log\left[\left(\frac{\mathrm{FWHM_X}}{\mathrm{FWHM_G}}\right)\sqrt{\frac{t_{\mathrm{exp,G}}}{t_{\mathrm{exp,X}}}}\left(\frac{2.2\ \mathrm{m}}{D}\right)\right]\ , \tag{5.2}$$

with G denoting GaBoDS quantities and X denoting quantities of the other dataset. FWHM is the measured seeing, t_{exp} is the exposure time, and D is the diameter of the telescope. By doing so we neglect differences between similar filter transmission curves and variations in observing conditions (moon, sky transparency, etc.). Thus, the limiting magnitudes are only rough estimates for approximate comparison.

The dependence of photometric errors on magnitude and redshift in the three datasets is shown in Fig. 5.1. The errors for the COMBO data are derived from multiple measurements of the same sources, where photon shot-noise is assumed to be a lower limit. The GaBoDS and FDF errors are purely derived from shot-noise as no multiple measurements were made.

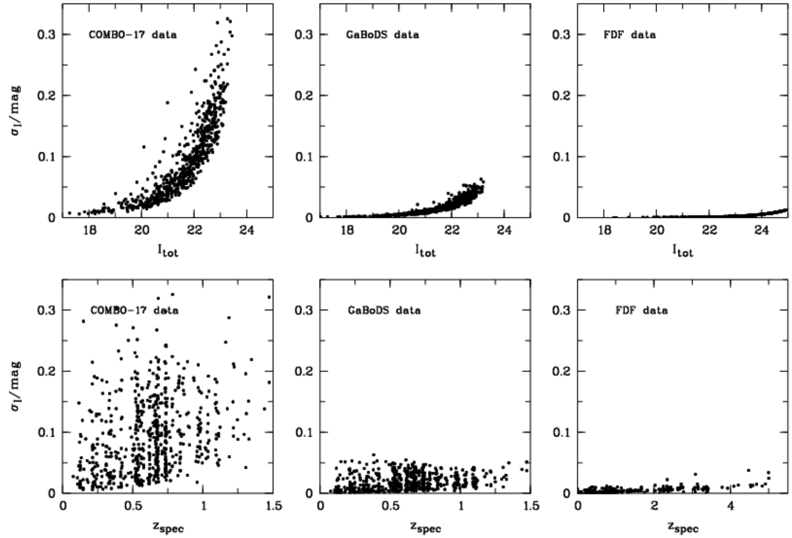

Figure 5.1: Photometric errors in the I-band as a function of I-magnitude (*upper panel*) and as a function of spectroscopic redshift (*lower panel*) for the COMBO-17 data (*left*), the GaBoDS data (*middle*), and the FDF data (*right*); see text for information on how the errors were estimated.

Table 5.1: Properties of the imaging data used in this study. The limiting magnitudes for GaBoDS correspond to 10σ sky noise in an aperture with a diameter of $1 \times$ FWHM. For comparison we have calculated effective limiting magnitudes from the exposure times, seeing values, and telescope apertures neglecting the difference in the filter sets and sky transparency (see text). The FDF limiting magnitudes in the $ZJKs$-bands are the ones given in Heidt et al. (2003) and Gabasch (2004) corresponding to 50% completeness. The total exposure time of the COMBO-17 medium-bands on the CDFS is 108 ksec. Throughout this chapter we use Vega magnitudes if not otherwise mentioned.

	COMBO-17			GaBoDS			FDF		
Band	Exp. time	FWHM	m_{lim}	Exp. time	FWHM	m_{lim}	Exp. time	FWHM	m_{lim}
U	21 600	1.″00	24.6	78 900	1.″01	25.3	44 400	0.″97	26.4
B	11 240	1.″10	26.2	69 400	0.″98	27.3	22 700	0.″60	28.6
V/g	8400	1.″20	25.2	104 600	0.″92	26.9	22 100	0.″87	27.5
R	35 700	0.″75	26.4	87 700	0.″79	26.8	26 400	0.″75	27.6
I	9800	1.″20	23.6	34 600	0.″93	24.6	24 900	0.″53	26.4
Z	-	-	-	-	-	-	18 000	0.″48	25.3
J	-	-	-	-	-	-	4800	1.″20	22.9
K_{s}	-	-	-	-	-	-	4800	1.″24	20.7

5.1.2 Spectroscopic catalogues

VVDS spectroscopic data in the CDFS

The VVDS is a magnitude-limited spectroscopic survey in five fields one of which is centred on the CDFS with a magnitude limit of $I_{AB} = 24$. Le Fèvre et al. (2004) published 1599 redshifts across this field including 1452 galaxies. The redshift measurements have associated reliability flags, and in the following comparisons we use only galaxies with flags 3 or 4 (or secondary targets with flags 23 or 24) indicating 95% and 100% confidence, respectively, to avoid errors introduced by the spectroscopic catalogue. This leaves us with 640 objects with $R_{\mathrm{WFI}} < 24$, whose redshift distribution is shown in Fig. 5.2. Since there are very few objects beyond redshift $z \approx 1.2$ this catalogue is ideally suited to assess the performance of photometric redshifts with optical data alone. The colour-redshift degeneracies (see Sect. 4.3.2) for an optical filter set become increasingly serious for redshifts above $z \approx 1.2$ because the Balmer-/4000Å-break is then outside the optical window and the Lyman-break does not enter before $z \approx 2$. The redshift of a galaxy at these redshifts with such a featureless spectrum (in the optical) is not only hard to determine with photo-z's but also with optical spectrographs, hence the term "redshift desert".

FDF spectroscopic data

Noll et al. (2004) selected objects for follow-up spectroscopy from the FDF imaging data described above. Photometric redshifts were estimated and galaxies were chosen to cover the range $0 < z < 5$. The spectroscopic redshift distribution of all 341 objects plus 14 additional objects observed later is also shown in Fig. 5.2 in comparison to the one of the VVDS data. This deep dataset extends well beyond the region where optical photometric redshifts work well and can illustrate the benefit of near-infrared data on photometric redshift performance for $z > 1$.

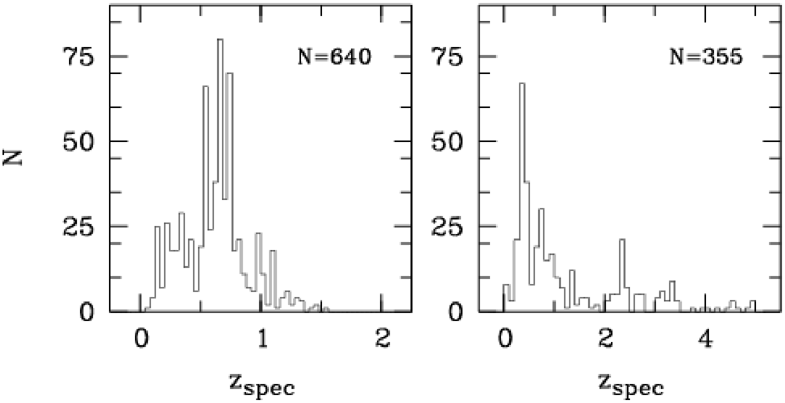

Figure 5.2: Spectroscopic redshift distributions of the comparison samples. *Left:* The VVDS-CDFS spectroscopic data with $R_{\mathrm{WFI}} < 24$ and flags 3,4,23,24. *Right:* The FDF spectroscopic data.

5.2 Description of photo-z quality

The performance of one particular setup is characterised by some basic quantities which are described in the following.

The mean, δ_z, and the standard deviation, σ_z, of the following quantity are calculated:

$$\Delta z = (z_{\mathrm{phot}} - z_{\mathrm{spec}})/(1 + z_{\mathrm{spec}})\,. \tag{5.3}$$

3σ outliers are rejected iteratively and after convergence their fraction is given by $f_{3\sigma}$. By doing so, the outlier fraction $f_{3\sigma}$ is not independent of the scatter σ_z. Therefore, we additionally report the quantity $f_{0.15}$ which is the fraction of objects for which $|\Delta z| > 0.15$.

5.2.1 Rejection of uncertain objects

As described above every photo-z code gives a confidence estimate for each object. *Hyperz* and the COMBO code report confidence intervals on the redshift while *BPZ* uses the *ODDS* parameter. It is obvious that an end-user will reject objects that clearly have uncertain photo-z estimates, although it is a-priori unclear how to define these objects. Since the codes do not estimate confidence measures in identical ways it is not possible to apply a universal threshold. We can get an idea of appropriate thresholds for the different codes by varying the cuts on the confidence intervals or the *ODDS* parameter, respectively. Thus, we see how the quantities δ_z, σ_z, and $f_{3\sigma}$ change with the completeness of the remaining sample.

When using *Hyperz* and the COMBO code all objects with a probability vs. redshift distribution that is too wide are rejected by the following criterion:

$$\sigma > A \times (1 + z_{\mathrm{phot}})\,, \tag{5.4}$$

with σ being the half-width of the 68% confidence interval in *Hyperz* and the standard deviation of $p(z)$ in the COMBO code, respectively, while A is varied from 0 to 1. The fraction of rejected objects is called r_A and the completeness then becomes compl. $= 1 - r_A$.

In *BPZ* we reject all objects with:

$$ODDS < A\,, \tag{5.5}$$

with A varied from 100% to 0%.

In this way diagrams showing δ_z, σ_z, and $f_{3\sigma}$ vs. completeness are created. While δ_z is almost independent of completeness the dependencies of σ_z and $f_{3\sigma}$ on completeness for selected setups are shown in Fig. 5.3.

The COMBO code (dashed line) hardly reacts to a variation of the cut criterion. Only the scatter values decrease slightly. This indicates that the photo-z errors are of limited value for describing the actual photo-z accuracy.

The curves corresponding to *Hyperz* (dotted lines) show a strong dependence of the outlier rate, $f_{3\sigma}$, on a tightening of the cut. At some point around 80% completeness a saturation behaviour sets in and a further tightening does not decrease the outlier rates any-

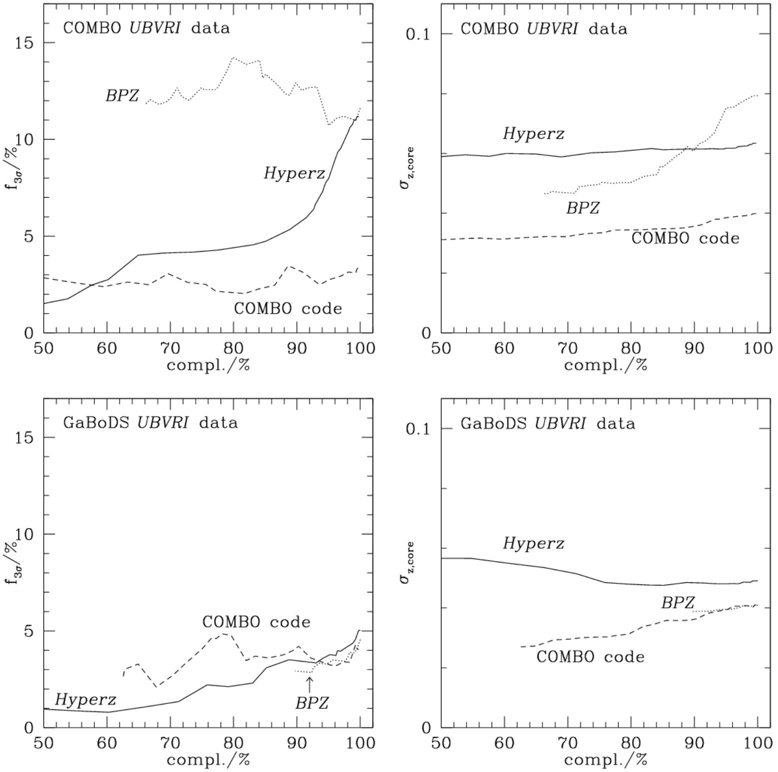

Figure 5.3: Characteristic lines showing completeness vs. 3-σ outlier rate, $f_{3\sigma}$, (*left*) and vs. scatter, σ_z, (*right*) for the COMBO (*top*) and the GaBoDS (*bottom*) *UBVRI* imaging dataset ($17 < R < 23$). The solid line corresponds to *Hyperz*, the dashed line to *BPZ*, and the dotted line to the COMBO code.

more. The scatter values stay fairly constant over the investigated completeness interval.

The *ODDS* parameter put out by *BPZ* does not allow to vary the completeness over a large interval since many objects are assigned an *ODDS* value of 1. *BPZ* combined with the GaBoDS data (grey lines, lower panel) shows a similar behaviour as *Hyperz* with decreasing outlier rates and constant scatter when completeness is decreased. In contrast, *BPZ* with the COMBO data (grey line, upper panel) shows a monotonically decreasing scatter while the outlier rate is nearly constant on a high level.

Investigating these characteristic lines of many more set-ups, we find that the most obvious feature is that σ_z as well as $f_{3\sigma}$ are often insensitive to a tightening of the cut criterion. This immediately tells us that the errors put out by the photo-z codes are not proportional to the real errors on an object-by-object basis and thus of limited use. The real accuracy of the photo-z is not tightly correlated with the error estimate. However, sometimes a saturation behaviour is observed: very large confidence intervals or very low *ODDS* values indicate that the photo-z estimation failed indeed. At some point (often roughly at 80% completeness) a further tightening of the cut criterion does not help in decreasing outlier rates or scatter any more. We assume that at this point the width of the confidence interval is not dominated by the photometric errors but becomes influenced by systematic uncertainties in the photometric calibration, the template set, the filter curves, or the code itself.

From the preceding paragraph it should be clear that the choice of A in Eq. 5.4 and Eq. 5.5 as a criterion for a reliable redshift estimate is somewhat arbitrary. After careful investigation of all characteristic line plots for all setups we decided to fix the cut for the rejection of uncertain objects in *Hyperz* at $\sigma > 0.125$, in the COMBO code at $\sigma > 0.15$, and in *BPZ* at $ODDS < 0.95$. This appears to eliminate the most uncertain objects in the datasets studied here.

There is clearly some amount of degeneracy between the quantities defined in this section. If the photo-z error distribution was purely Gaussian, scatter and bias would be sufficient numbers to characterise the accuracy of one particular setup. As described above, this is not the case for real data (see also Figs. 5.4 & 5.5). Usually, there is a core which might be offset by some bias and there are very extended wings containing catastrophic outliers. This complex error distribution is not easily described by a few numbers, and a specific choice must be a compromise between clarity and degeneracy.

For example, a smaller core scatter will probably produce more 3σ outliers than a larger core scatter. With no alternative at hand to condense the performance of one particular setup into a handful of numbers we can only refer to the $z_{\rm phot}$ vs. $z_{\rm spec}$ plots shown in Figs. 5.4-5.6 which give an uncompressed view of the data.

5.2.2 VVDS setups

The VVDS is complete down to $I_{\rm AB} = 24$ on the CDFS which corresponds to $I_{\rm Vega} \approx 23.5$. Thus, we decided to asses the photometric redshift accuracy for all objects with $17 < R < 23$ to achieve a reasonable level of completeness. We use R-band magnitudes instead of I-band magnitudes because both imaging catalogues are R-band selected. Considering objects with VVDS flags 3/4/23/24 we find that their fraction has dropped to approximately 50% compared to the whole VVDS catalogue at $R = 23$. We also present the results for a fainter magnitude bin with objects in the range $23 < R < 24$ which are then possibly biased in terms of selection. The mean redshifts in the bright and the faint bins are $z = 0.55$ and $z = 0.75$, respectively.

The different setups are given three-letter/digit-acronyms with the first letter denoting the code ("H" for *Hyperz*, "C" for the COMBO code, and "B" for *BPZ*), the second letter denoting the dataset ("B" for GaBoDS and "C" for COMBO), and the digit at the third position denoting the filter set ("5" for $UBVRI$, "4" for "BVRI", and "17" for the full COMBO-17 filter set including medium-band-filters).

The COMBO code considers only redshifts in the range $0 < z < 1.4$ for the VVDS setups since for this depth ($I_{\rm AB} < 24$) only very few objects are expected beyond $z \sim 1.4$ (see also Fig. 5.2).

5.2.3 FDF setups

Since the FDF data are extremely deep, and subtle trends in photometric errors make little difference to the photo-z quality, we do not split the FDF sample into magnitude bins. Furthermore, given the selection choices made for the spectroscopic sample it is not complete at or representative for any particular magnitude limit. Hence, we split the FDF spectroscopic catalogue into two samples at $z = 2$ to show the effects of different filter sets and especially NIR bands on the performance at low and high redshift in comparison. The FDF setups are denoted by a second letter "F" for FDF and the filter set is spelled out.

5.3 Results

In the following we report the results from our blind test of photo-z performance for the different setups. As a complete coverage of all possible data-code-parameter combinations is beyond the scope of this study we concentrate on well-chosen setups to illustrate the effects of key parameters.

5.3.1 VVDS results

The statistics for all photo-z setups that are compared to the VVDS spectroscopic catalogue are shown in Tables 5.2 & 5.3. Selected setups are also illustrated in Figs. 5.4 & 5.5 by plots showing photometric redshift versus spectroscopic redshift.

COMBO-17 imaging data

The best results in both magnitude bins are achieved with the full 17-filter set of COMBO-17 combined with the COMBO code. Especially in the bright magnitude bin the scatter and the outlier fractions are very small compared to all 4- or 5-filter setups. In the fainter bin, however, the difference is not as dramatic due to the lack of depth in many of the medium-bands. *Hyperz* also shows relatively accurate results for the 17-filter set (HC17) but far from being as accurate as the CC17 results.

Clearly, the COMBO code performs best in comparison to the two other codes with the 17-filter set as well as with the 4- and 5-filter sets.

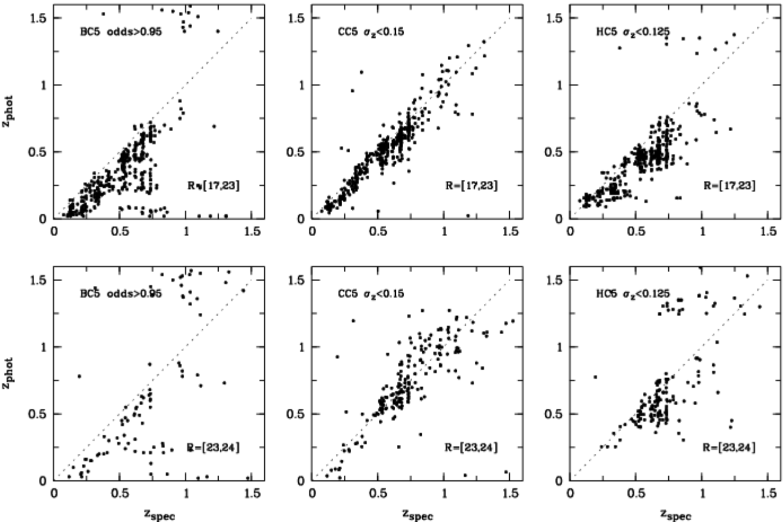

Figure 5.4: Photometric redshifts from the CDFS-COMBO *UBVRI* imaging data vs. spectroscopic redshifts from the VVDS. The *left* panels show results for *BPZ*, the *middle* panels for the COMBO code, and the *right* panels for *Hyperz*, respectively. Bright objects with $17 < R < 23$ are shown in the *top* panels, faint objects with $23 < R < 24$ in the *bottom* panels.

BPZ and *Hyperz* produce large negative biases when the U-band is included in the photometric redshift estimation. An exclusion of the U-band decreases this bias albeit alongside decreasing completeness. The presence of this bias motivated us to investigate a re-calibration of the photo-z's which is discussed in Sect. 5.4.

The COMBO code shows the expected behaviour that the photo-z accuracy decreases when further filters are excluded. Completeness decreases while outlier rate and scatter increase. No significant bias is produced in any setup.

A very interesting fact concerning *Hyperz* is that the exclusion of the U-band *increases* the accuracy in nearly all statistics. Only

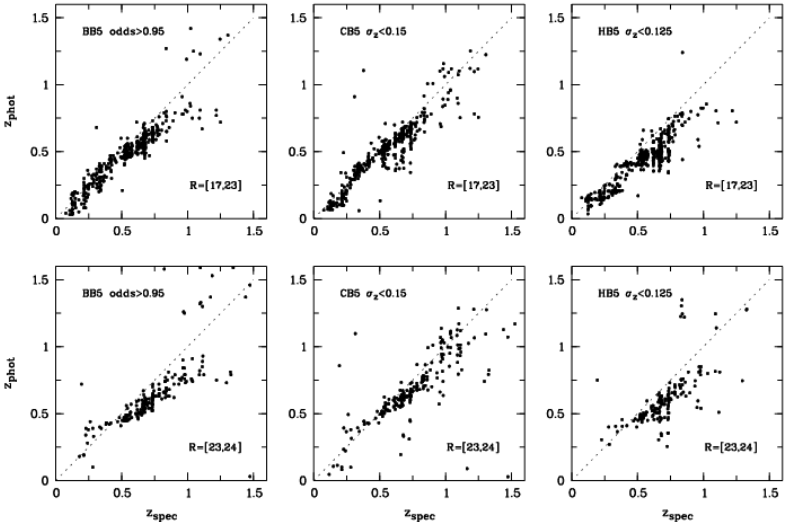

Figure 5.5: Same as Fig. 5.4 but for the GaBoDS $UBVRI$ imaging data.

the completeness decreases slightly. Clearly, the results from the COMBO code show that this behaviour is most likely not caused by a badly calibrated U-band. Such an unexpected behaviour must instead be due to systematics in the code or template set.

BPZ seems to have similar problems with the U-band as *Hyperz*. An exclusion of this band again *increases* accuracy. In particular, for fainter objects in the magnitude bin, $23 < R < 24$, *BPZ* produces very inaccurate results when used with the $UBVRI$ filter set, which is also illustrated by the bottom-left panel of Fig. 5.4. The completeness drops to $\sim 40\%$ but still the outlier and scatter values are quite high. The situation changes for the four filter data for which the results are reasonably accurate though only for a third of the sample. Also, the exclusion of objects with uncertain photo-z estimates by the $ODDS$ parameter works fine in contrast to the five-filter $UBVRI$ data.

GaBoDS imaging data

Owing to their greater depth the GaBoDS data mostly lead to better results than the shallower COMBO data, in $UBVRI$ as well as in $BVRI$. As expected, the effect is much more pronounced in the faint bin, while the depth helps less in the estimation of redshifts for high S/N objects at the bright end of our catalogue. Nevertheless, also bright objects with > 20-σ detections in the R-band benefit from the depth in the other bands.

The negative bias in the photo-z estimation with *BPZ* and *Hyperz* is also present in GaBoDS setups with the U-band included. However, it is clearly not as severe as for the COMBO data. At this point, it is important to mention again that the GaBoDS U-band filter is different from the COMBO U-band filter. The GaBoDS filter is wider and bluer.

For the COMBO code, the 4- and 5-filter results are nearly indistinguishable. Only in the faint bin the outlier rates increase slightly when the U-band is excluded.

Hyperz again shows the unexpected feature that most statistics become more accurate when going from five to four filters.

BPZ shows a similar behaviour as the COMBO code. The statistics are nearly independent on the choice between 4- and 5-filter set. For this code, a bias of ~ 0.06 mag in the faint bin is also present when using just $BVRI$.

Common trends

Typically, the COMBO code rejects fewer objects than *Hyperz*. Nevertheless, the outlier rates, $f_{3\sigma}$, produced by *Hyperz* are in most cases larger than the outlier rates produced by the COMBO code. *BPZ* produces $f_{3\sigma}$ values which are not too different from *Hyperz*.

The outlier-excluded scatter values, σ_z, do not show a clear trend with every code being the most accurate in at least one setup. There is clearly some amount of degeneracy between completeness, $f_{3\sigma}$, δ_z, and σ_z. The plots in Figs. 5.4 & 5.5 provide a more complete view of the performance.

Remarkably, there are only very small or negligible biases for the setups with the COMBO code when compared to the ones reported above for the other codes. Interestingly, similar negative biases as in our *BPZ* and *Hyperz* setups are found by Csabai et al. (2003) in SDSS photo-z estimates that were obtained by using the CWW or BC templates, which are also used by *Hyperz*. This strongly supports the assumption that the absolute photometric calibration of the imaging data is not completely off. Rather the combination of these ground-based photometric datasets with particular codes and template-sets seems to be problematic.

BPZ and *Hyperz* are tested in their release papers (Benítez 2000; Bolzonella et al. 2000) only against real data from the Hubble Deep Field, besides simulations. *BPZ* now incorporates a new template set (see Sect. 4.4.3) that was specially calibrated for HST photometry. With these very deep, high-resolution datasets all the non-trivial problems of ground-based galaxy photometry (e.g. different PSFs in different filters, changing photometric conditions, etc.) do not play a role. The COMBO code, however, was originally designed for the ground-based survey CADIS (Wolf et al. 2001), where colours were measured bias-free from seeing adaptive photometry, and included photometric redshifts for point-source QSOs.

One of the biggest differences between the codes is the template set chosen and one might presume that most of the difference in performance originates from this point. However, we run *Hyperz* with the PEGASE templates used by the COMBO code as well as with the CWW templates plus two Kinney starburst templates originally used by *BPZ* in Benítez (2000) (switching off the reddening because it is already included in the *BPZ* and COMBO-PEGASE templates). The results can neither compete with the best *Hyperz* setups incorporating the BC templates nor with the COMBO code plus PEGASE templates. Hence, the implementation of user-defined templates appears to be not straightforward and results may not be competitive with the template sets that are shipped with the code and were tested and optimised by the author.

Another interesting point is the comparison of the CC17 setup with the CB5 setup. While the total exposure time with WFI is lower for

CC17, the performance of CC17 is better in all statistics described here. It is clear that for the particular application of photometric redshifts for bright objects, the exposure time was well spent on more filters (which is an important result for future surveys). However, the GaBoDS data of the CDFS are completely based on archive data and no specific observing programme was proposed to create these deep images. Furthermore, for applications that depend on deeper data, such as Lyman-break galaxy studies, the GaBoDS data are certainly highly superior to the COMBO data.

Table 5.2: Photometric redshift errors and outlier rates for selected setups on the COMBO-CDFS data (bright sample *left*, faint sample *right*). The different setups are described by configuration acronyms in the first column as described in the text. Each setup is characterised by its completeness after applying the cuts on the confidence intervals (COMBO code and *Hyperz*) or the *ODDS* parameter (*BPZ*), respectively. The fraction of objects with $|\Delta z| = |(z_{\rm phot} - z_{\rm spec})/(1 + z_{\rm phot})| > 0.15$ is represented by $f_{0.15}$. The mean, δ_z, and the standard deviation, σ_z, of the Δz are calculated after rejecting 3σ outliers iteratively, whose fraction is given by $f_{3\sigma}$.

Sample	$R = [17, 23]$				$R = [23, 24]$			
Mean redshift	0.55				0.75			
Configuration	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$
BC17	91.7	30.0	17.1	-0.095 ± 0.053	55.0	49.2	17.5	-0.097 ± 0.154
BC5	87.0	25.3	12.0	-0.096 ± 0.061	40.1	42.5	1.1	-0.066 ± 0.212
BC4	82.5	5.5	5.5	-0.023 ± 0.052	32.7	8.5	9.9	-0.035 ± 0.040
CC17	99.8	1.4	4.3	-0.014 ± 0.021	97.7	7.5	8.9	0.000 ± 0.043
CC5	99.5	2.9	3.3	-0.031 ± 0.040	91.3	10.1	6.0	-0.004 ± 0.069
CC4	93.1	6.4	5.3	-0.017 ± 0.054	61.9	20.7	15.6	-0.014 ± 0.072
HC17	100.0	5.7	6.2	-0.031 ± 0.043	99.1	20.9	20.5	-0.032 ± 0.045
HC5	94.8	18.0	7.8	-0.074 ± 0.062	78.8	30.4	25.1	-0.051 ± 0.065
HC4	78.4	10.0	7.3	-0.034 ± 0.057	71.0	18.2	11.7	-0.049 ± 0.061

Table 5.3: Same as Table 5.2 but for the GaBoDS-CDFS data.

Sample	$R = [17, 23]$				$R = [23, 24]$			
Mean redshift	0.55				0.75			
Configuration	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$
BB5	97.4	2.7	3.4	-0.037 ± 0.040	87.7	6.7	12.4	-0.062 ± 0.041
BB4	88.5	3.1	10.4	-0.010 ± 0.045	85.0	5.3	9.6	-0.060 ± 0.043
CB5	99.5	3.6	4.3	-0.024 ± 0.041	98.6	8.8	8.8	-0.028 ± 0.044
CB4	99.8	4.8	3.8	-0.019 ± 0.049	96.8	11.4	12.3	-0.026 ± 0.042
HB5	95.7	9.0	3.8	-0.065 ± 0.048	79.1	18.4	19.0	-0.060 ± 0.039
HB4	74.2	4.9	5.5	-0.034 ± 0.040	75.5	13.3	13.9	-0.052 ± 0.035

5.3.2 FDF results

Table 5.4 summarises the results on the FDF and Fig. 5.6 shows photometric vs. spectroscopic redshift for selected setups.

In the lower redshift bin again the COMBO code combined with imaging data in 8 filters delivers the smallest outlier rate, bias, and scatter when compared to *BPZ* and *Hyperz* in 8 filters. At least in this redshift interval the results are nearly as good as the results produced by Gabasch et al. (2004) with a template set specifically calibrated for the FDF.

In the high redshift domain, however, the COMBO code does not perform well with an outlier rate and scatter twice as large as the ones produced by *Hyperz* and with a considerable bias. *BPZ* performs not too different from *Hyperz*. Apparently, the COMBO code in combination with the PEGASE templates has problems when the Lyman break enters the filter set: many objects appear at too low redshifts, hence the large negative bias (see also Fig. 5.6). The inferior performance in the high redshift domain can then be attributed to colour-redshift-degeneracies described in detail in Benítez (2000). Simplified, a larger number of templates can lead to better low-z performance with the tradeoff of poorer high-z performance due to increasing degeneracies. Designed for medium-deep surveys the COMBO code was naturally not optimised to work at high redshifts in contrast to *BPZ* and *Hyperz*. There, the application of a Bayesian prior on the apparent magnitude combined with a sparse template set (*BPZ*) or a top-hat prior on the absolute magnitude (*Hyperz*) delivers significantly better results. Edmondson et al. (2006) tested the COMBO code with a Bayesian prior on the COMBO-17 data and slightly improved its performance for fainter $z < 1$ galaxies ($R > 23$); however, it was not tested on the high-z FDF objects where priors should be even more beneficial.

The dependence of photometric redshift performance on the filter set is also shown in Table 5.4. In the lower redshift interval the outlier rates increase as soon as the NIR filters J and K_s are dropped. The scatter, however, remains nearly constant. No significant bias is observed in contrast to the VVDS-*BPZ* and VVDS-*Hyperz* setups

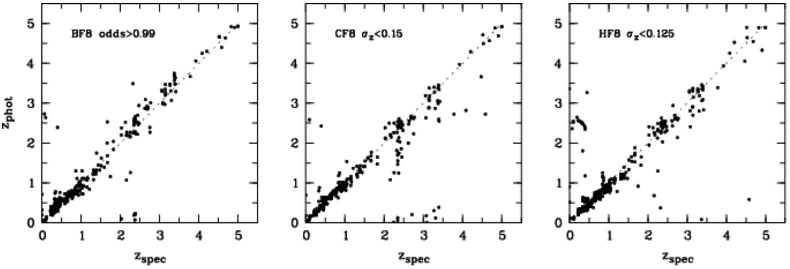

Figure 5.6: Photometric vs. spectroscopic redshifts for the FDF full 8-filter set imaging data. The *left* diagram shows results for *BPZ*, the *middle* diagram for the COMBO code, and the *right* diagram for *Hyperz*, respectively.

(see Table 5.2 and 5.3), but as mentioned in Sect. 5.1.1 the FDF photometric catalogue was already calibrated before by the FDF team for better photo-z performance.

Moreover, we especially checked whether it is better to drop UV or near IR data. In *BPZ* we noted virtually no difference between the $BgRIZJK$s and $UBgRIZ$ setups keeping the degeneracy between $f_{3\sigma}$ and σ_z in mind. However, with the COMBO code dropping the U-band is preferred at high-z while with *Hyperz* dropping J and K_s decreases performance less than dropping U.

Table 5.4: Same as Tables 5.2 and 5.3 but for the FDF data (low-z sample *left*, high-z sample *right*).

Sample		$z = [0, 2]$				$z = [2, 5]$		
Mean redshift		0.65				2.94		
Configuration	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$
$\mathrm{BF}_{-UBgRIZJK\mathrm{s}}$	98.1	5.4	5.4	0.005 ± 0.053	93.3	12.0	13.3	0.026 ± 0.046
$\mathrm{CF}_{-UBgRIZJK\mathrm{s}}$	99.6	3.4	4.9	0.001 ± 0.034	100.0	29.2	13.5	-0.046 ± 0.093
$\mathrm{HF}_{-UBgRIZJK\mathrm{s}}$	99.6	10.2	9.8	-0.019 ± 0.051	100.0	8.0	5.7	-0.004 ± 0.056
$\mathrm{BF}_{-BgRIZJK\mathrm{s}}$	98.1	9.2	2.3	0.019 ± 0.073	92.1	20.7	20.7	0.033 ± 0.047
$\mathrm{BF}_{-UBgRIJK\mathrm{s}}$	98.5	8.4	9.2	0.011 ± 0.058	89.9	18.8	18.8	0.024 ± 0.050
$\mathrm{BF}_{-UBgRIZ}$	97.7	7.7	9.6	-0.010 ± 0.041	91.0	17.3	19.8	0.032 ± 0.047
BF_{-UBgRI}	98.1	8.4	9.2	0.000 ± 0.042	74.2	27.3	27.3	0.025 ± 0.057
BF_{-BgRI}	90.2	12.5	8.3	0.020 ± 0.060	53.9	27.1	27.1	0.017 ± 0.051
$\mathrm{CF}_{-BgRIZJK\mathrm{s}}$	99.6	6.8	10.6	0.002 ± 0.033	100.0	25.8	16.9	-0.033 ± 0.077
$\mathrm{CF}_{-UBgRIJK\mathrm{s}}$	99.6	4.5	4.5	0.006 ± 0.044	95.5	30.6	14.1	-0.058 ± 0.102
$\mathrm{CF}_{-UBgRIZ}$	99.6	6.0	7.9	-0.006 ± 0.039	98.9	54.5	47.7	-0.061 ± 0.077
CF_{-UBgRI}	96.6	5.8	7.0	-0.001 ± 0.043	94.4	57.1	38.1	-0.100 ± 0.082
CF_{-BgRI}	92.5	16.3	13.0	0.006 ± 0.066	93.3	62.7	50.6	-0.095 ± 0.099
$\mathrm{HF}_{-BgRIZJK\mathrm{s}}$	100.0	12.8	15.4	-0.019 ± 0.048	100.0	14.8	10.2	-0.009 ± 0.066
$\mathrm{HF}_{-UBgRIJK\mathrm{s}}$	99.6	10.2	10.6	-0.022 ± 0.049	100.0	9.1	8.0	-0.012 ± 0.053
$\mathrm{HF}_{-UBgRIZ}$	100.0	12.8	11.7	-0.020 ± 0.048	100.0	12.5	12.5	0.009 ± 0.052
HF_{-UBgRI}	99.2	14.4	17.8	-0.028 ± 0.040	97.7	20.9	14.0	-0.012 ± 0.072
HF_{-BgRI}	99.6	19.6	22.6	-0.021 ± 0.038	96.6	24.7	16.4	-0.010 ± 0.086

5.4 Re-calibrating the data-template match

In the following, we show how the redshift estimates can suffer from sub-optimal match between data and templates, and how they can be improved by a re-calibration of the data. As an example, we show procedures for *BPZ* and *Hyperz*, which can easily decrease the large negative biases observed with the CDFS data here. Such a procedure is only feasible when a spectroscopic catalogue of considerable size is available. Otherwise, the reader should take the results from Sect. 5.3 as final estimates of photo-z performance. Interestingly, the FDF data have already been re-calibrated by the FDF team and do not benefit from further modifications here.

The methods applied here rely on the re-calibration of the photometry to match the templates. Certainly, it is also possible to optimise the templates to better match the observed colours which was done by e.g. Csabai et al. (2003) for the SDSS LRG sample. See Ilbert et al. (2006) for a combination of both re-calibration methods on CFHTLS-Deep data.

As described above the relative photometric calibration between the different bands of the COMBO-17 and GaBoDS datasets is most likely better than 0.1 mag at least for stars. If the galaxy photometry in one particular filter was affected by some systematic offset (in comparison to the template set or space-based photometry) adjusting this offset should lead to more accurate photometric redshifts.

5.4.1 Method for Hyperz

Hyperz enables the user to set the redshift range for the template fit not only globally for all objects but also individually for every single object. Confining these individual redshift ranges to the spectroscopic redshifts of the VVDS objects with $17 < R < 23$ we find the best fit templates at the spectroscopic redshifts. The modelled magnitudes of these best-fit templates, which are also put out by the code, deviate slightly from the observed magnitudes. By averaging these differences over all objects we derive constant corrections for the photometry in all filters. Iterating this procedure the offsets drop below a reasonable

Table 5.5: Magnitude offsets derived with *Hyperz* in the different broad-band filters. It should be noted that there is not a filter for which the offset vanishes because the modelled fluxes are not scaled to the observed ones in just one filter but the scaling is fitted by the parameter b over all filters in Eq. (4.1).

Filter	COMBO	GaBoDS
U	-0.160	-0.185
B	0.132	0.062
V	0.099	0.038
R	0.042	0.030
I	-0.060	-0.021

threshold of 0.01 mag quite quickly (after < 10 iterations) and the cumulative offsets are applied to the photometric catalogue. We use the BC templates again for consistency but it should be noted here that after re-calibration the performance differences between the BC and the CWW templates mentioned in Sect. 4.4.1 vanish. Table 5.5 presents the constant offsets applied to the COMBO and GaBoDS CDFS data to improve the *Hyperz* photometric redshift estimates.

5.4.2 Method for BPZ

BPZ allows the user to assume that these offsets may not be constant over the whole magnitude range. Thus, we vary the photometry in the different filters by a quadratic function,

$$\Delta m_f = a_{1,f}(I - I_0) + a_{2,f}(I - I_0)^2 \,, \tag{5.6}$$

with I_0 being a reference magnitude in the I-band for which the correction vanishes ($I_0 = 16; 19$ for GaBoDS; COMBO; fixed to these values in all filters) and $a_{1,f}$ and $a_{2,f}$ being two variable parameters for the filter f. In principle one could also let I_0 vary in the different filters but that did not improve the results for the five-filter sets so that we keep it fixed here. By varying the $a_{i,f}$ as free fit parameters and keeping the redshift fixed to the spectroscopic redshift we find the

Table 5.6: Parameters for the magnitude corrections derived with *BPZ* in the different broad-band filters (see Eq. 5.6). Fig. 5.7 shows these corrections as a function of magnitude.

	COMBO		GaBoDS	
Filter	$a_{1,f}$	$a_{2,f}$	$a_{1,f}$	$a_{2,f}$
U	0.3117	−0.0571	0.1880	−0.0175
B	−0.0854	0.0127	0.0004	0.0010
V	−0.0343	0.0047	−0.0017	0.0007
R	−0.0173	0.0041	−0.0144	0.0008
I	0.0164	0.0187	−0.0003	0.0014

magnitude corrections in the different bands for which the differences between the modelled fluxes of the best-fit *BPZ* template at redshift z_{spec} and the corrected observed fluxes become minimal. In Table 5.6 the parameters for the *BPZ* correction scheme (see Eq. 5.6) are shown and in Fig. 5.7 these quadratic corrections are displayed graphically. For the full COMBO-17 filter set we additionally allow for I_0 to vary for the different filters because there are considerable offsets from the *BPZ* templates even at bright magnitudes.

5.4.3 Results

The magnitude corrections, derived in the ways described above, indeed lead to a decrease in the bias for the *Hyperz* and *BPZ* setups. Additionally, completeness increases as well as outlier rates and scatter decrease. The results are presented in Tables 5.7 & 5.8 and Figs. 5.8 & 5.9.

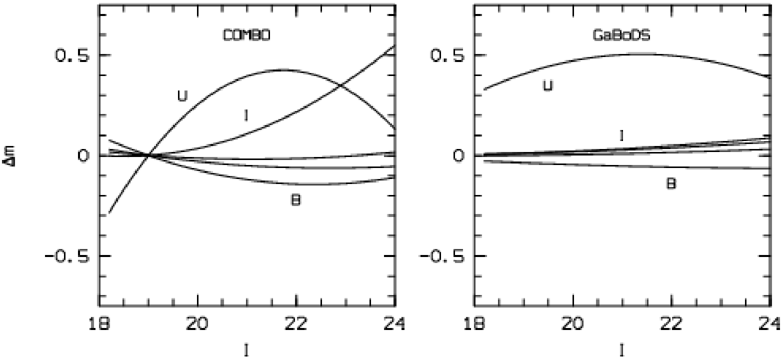

Figure 5.7: Magnitude corrections applied to the COMBO (*left*) and GaBoDS (*right*) photometry in order to minimise flux offsets between the best-fit *BPZ* templates at z_{spec} and the observed fluxes.

Table 5.7: Photometric redshift errors and outlier rates for selected setups on the re-calibrated COMBO-CDFS data (bright sample *left*, faint sample *right*). The columns are identical to Table 5.2. Here, the re-calibrated setups are represented by the same acronyms as above with an additional "c".

Sample	$R = [17, 23]$				$R = [23, 24]$			
Mean redshift	0.55				0.75			
Configuration	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$
BC17c	99.1	1.0	4.8	-0.005 ± 0.024	82.6	2.8	6.7	-0.004 ± 0.031
BC5c	93.4	1.3	3.6	-0.009 ± 0.036	42.4	5.4	4.3	0.005 ± 0.059
BC4c	79.1	3.3	7.2	-0.005 ± 0.037	28.6	4.8	3.2	0.020 ± 0.055
HC17c	100.0	5.9	6.4	-0.024 ± 0.037	100.0	15.1	15.6	-0.027 ± 0.042
HC5c	97.2	4.6	2.7	-0.034 ± 0.052	87.8	11.6	6.8	-0.034 ± 0.072
HC4c	76.5	8.4	6.2	-0.027 ± 0.054	72.9	13.8	8.8	-0.036 ± 0.065

Table 5.8: Same as Table 5.7 but for the re-calibrated GaBoDS-CDFS data. This values should be compared to the uncalibrated results in Table 5.3.

Sample	$R = [17, 23]$				$R = [23, 24]$			
Mean redshift	0.55				0.75			
Configuration	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$	compl. [%]	$f_{0.15}$ [%]	$f_{3\sigma}$ [%]	$\langle\delta_z\rangle \pm \sigma_z$
BB5c	96.9	2.0	5.9	-0.009 ± 0.029	78.6	4.6	7.5	-0.015 ± 0.037
BB4c	87.6	1.9	4.6	-0.013 ± 0.032	70.9	5.1	7.7	-0.015 ± 0.037
HB5c	96.9	4.2	4.0	-0.038 ± 0.042	86.8	11.5	9.9	-0.045 ± 0.050
HB4c	73.0	4.9	6.2	-0.031 ± 0.038	75.9	12.0	13.8	-0.045 ± 0.037

The re-calibration of the 17-filter set with *Hyperz* could probably be further improved in terms of scatter but this may involve modifications in the source code, such as actually fitting the magnitude offsets as free parameters, which are beyond the scope of this study.

As described above the corrections should not be misunderstood as corrections for bad absolute photometric calibrations. It is also not likely that these corrections represent aperture corrections of ground-based galaxy photometry alone. The COMBO as well as the GaBoDS photometric catalogues were created in a way to take PSF variations in the different filters into account and as the results with the COMBO code show, these photometric measurements can lead to accurate photo-z estimates. Rather an intrinsic mismatch of ground-based colours and modelled template colours can create the systematic deviations presented here. Even if PSF effects are taken properly into account, the outer parts of a faint galaxy image will most probably be missed by flux measurements from the ground. This can lead to precise colours but for a much smaller aperture, and thus for a different part of a galaxy than for spectroscopically observed bright galaxies or for space-based photometry (the test case for *Hyperz* and *BPZ*). There, essentially the total fluxes are measured. In this way the SEDs used here may be a bad fit to ground-based photometry. The same is certainly true for synthetic templates which, by construction, represent total fluxes of a stellar population.

There are several other reasons that could contribute to the biases. Templates are usually not as accurate in the UV as in the optical. This plays a major role for intermediate redshift galaxies ($0.5 < z < 1$) where the restframe UV dominates the optical window. Furthermore, the different CCDs in a multi-chip camera show their largest sensitivity variations in the U-band. Taking these issues properly into account is likely to improve the overall redshift accuracy irrespective of the reliance on large spectroscopic calibration catalogues.

Interestingly, the COMBO code does not benefit from an application of the corrections to the photometry suggested by either *Hyperz* or *BPZ* and indeed produces much worse results when corrections

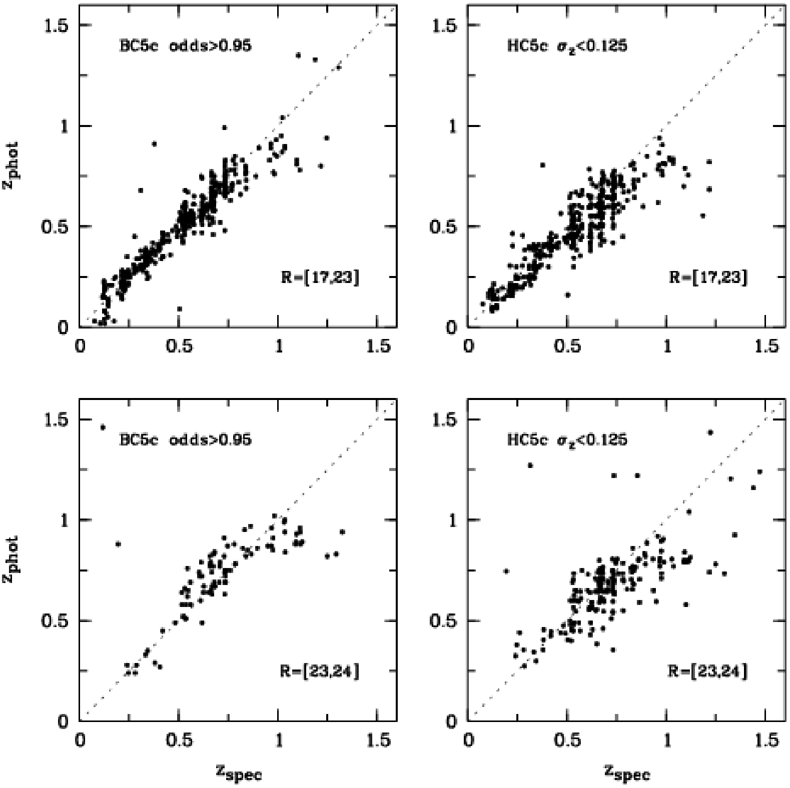

Figure 5.8: Photometric vs. spectroscopic redshifts for the re-calibrated COMBO-CDFS data in *UBVRI*. Results for *BPZ* are presented in the *left* panels and results for *Hyperz* in the *right* panels, respectively. The *top* panels show the bright subsample and the *bottom* panels the faint subsample, respectively. Note the improvement over the version presented in the left and right panels of Fig. 5.4.

are applied. In contrast, it already produced no major bias with the original datasets, suggesting that its templates are already well matched to the ground-based photometric data.

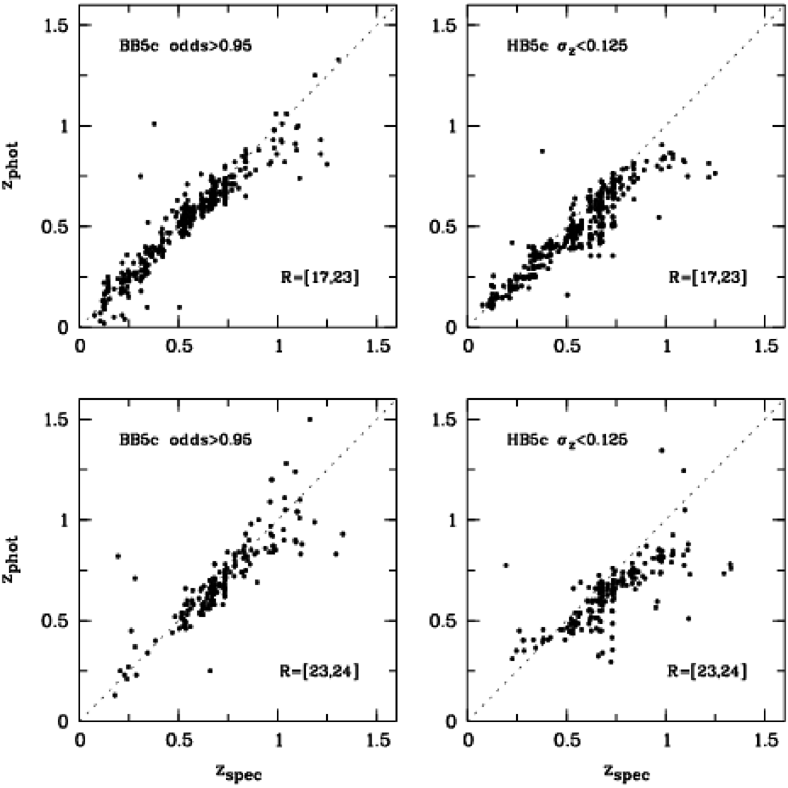

Figure 5.9: Same as Fig. 5.8 but for the GaBoDS-CDFS imaging data. To be compared to the left and right panels of Fig. 5.5.

5.5 Summary and Conclusions

We have shown that photometric redshifts estimated with publicly available tools can produce a reasonable accuracy “from the shelf”. The performance of a particular photometric redshift code, however, cannot easily be characterised by a mere two numbers such as scatter and global outlier rate. The benchmarks are rather sensitive functions of filter set, depth, redshift range and code settings. Moreover, there is at least a factor of two possible difference in bias, scatter, and outlier rates between different publicly available codes which is again not stable for all setups but can vary considerably from one setup to the other. There are e.g. redshift ranges where one code clearly beats the other one in terms of accuracy only to loose at other redshifts. We give estimates of the performance for a number of codes in some practically relevant cases.

With a large spectroscopic catalogue at hand, remaining systematic biases in the photo-z estimation can be removed by some re-calibration of the match between data and templates. The performance can benefit considerably from such a procedure, but again this benefit is not stable for all reasonable setups. There are code-data combinations which need more re-calibration to work properly and others which are more robust against the difficulties associated with ground-based galaxy photometry.

We have shown that the estimation of accurate photometric redshifts from different ground-based datasets is not straightforward at all and results should not be expected to be identical to simulated photometric redshift estimates. Also photo-z simulations often seem to neglect critical steps in ground-based photo-z estimation. For example, the match between observed colours and some template sets commonly used may be suboptimal (see Sect. 5.4). Moreover, PSF variations across the field can pose a considerable problem that is usually not accounted for in simulations.

In the preceding sections we have identified several aspects which are relevant to future optimisations of photo-z codes. The photo-z error estimation is one of the most unsatisfying aspects to date with error values often only very weakly correlated with real uncertainties. This is probably due to the insufficient inclusion of systematics since very low-S/N objects, for which the errors should be dominated by photon shot-noise, show a tighter correlation. Chip-to-chip sensitivity variations, especially in the UV, could either be taken into account more accurately within the photo-z codes or could be tackled by improved instrument design, survey strategy, and data reduction. The optimisation of template sets can be expected to be successfully done with larger and better spectroscopic catalogues becoming available (see e.g. Ilbert et al. 2006).

The "state-of-the-art" approaches on photometric redshifts now undertaken by large survey teams like the CFHTLS- and FDF-teams can produce results that are more accurate by up to a factor of two when compared to some setups presented here. Complex work on template sets and photometric calibration and extensive support by large spectroscopic campaigns are, however, necessary to reach such a level of accuracy. Another route to better photometric redshifts is improving the spectral resolution provided by the data as demonstrated successfully by the COMBO-17 survey, a route that is also taken by the new ALHAMBRA survey (Moles et al. 2005, Benítez et al. in prep.).

6

Lyman-break galaxies in the DPS[1]

For more than a decade now, the high-redshift universe has become reachable by observations mainly due to the development of efficient colour selection techniques. The group around C. C. Steidel (Steidel & Hamilton 1993; Steidel et al. 1996, 1999) has introduced the Lyman-break technique selecting high-redshift, star-forming galaxies from optical multicolour data by their pronounced Lyman-break (see also Sect. 4.5). Many groups have used this technique and analysed various properties of these galaxy populations from redshifts $z \approx 3$ up to $z \approx 7$.

A particular emphasis in these studies was given to the clustering properties of the Lyman-break galaxy (LBG) samples (Steidel et al. 1998; Giavalisco et al. 1998; Adelberger et al. 1998, 2005; Giavalisco & Dickinson 2001; Ouchi et al. 2001, 2004, 2005; Porciani & Giavalisco 2002; Bouché & Lowenthal 2004; Foucaud et al. 2003; Allen et al. 2005; Lee et al. 2006; Kashikawa et al. 2006; Gawiser et al. 2006). Going back in cosmic time the correlation strength of these high-z galaxies can be directly compared to N-body simulations or semi-analytical predictions yielding estimates of, e.g., the galaxy bias for early epochs. These results can then be used as an input for models of galaxy formation, and help to constrain the large number of free adjustable parameters. The more precise our knowledge of the clustering evolution becomes the more accurate our understanding of galaxy evolution will be.

In this context it is of crucial importance to reach a similar precision for the same galaxy populations at different redshifts. While in the beginning most LBG studies concentrated on relatively bright $z = 3$

---

[1] This chapter is largely based on the papers Hildebrandt et al. (2005, 2007b), published in A&A.

U-dropouts, in the last years more and more groups have focused on the investigation of LBGs at redshifts around $z = 4$ selected as B-dropouts. This is obviously due to the fact that deep and wide U-band images are still very telescope-time consuming, and some recent wide-field cameras like *Suprimecam* or space-based surveys like GOODS even lack a U-filter entirely.

By estimating the angular correlation function of nearly 17 000 B-dropouts selected from the Subaru/XMM-Newton Deep Field, Ouchi et al. (2005) find evidence for a departure from a pure power law on small scales. The same trend is reported by Lee et al. (2006) for B- and V-dropouts from the very deep GOODS ACS data. Neither the number statistics in the former study mentioned nor the depth and angular resolution of the GOODS data have a comparable counterpart at slightly lower redshifts. The most precise estimates of U-dropout clustering to date come from Adelberger et al. (2005) estimating the angular correlation function but not reporting any obvious excess on small scales. The question whether this is an evolutionary effect or whether this feature is only visible in the $z = 4$ data because of their superior quality can only be answered by a U-dropout survey comparable in size to the B-dropout surveys mentioned above.

In this chapter we describe our investigations of $z = 3$ LBGs in the ESO DPS. The methods presented here are based on our investigations in the CDFS (the DPS field Deep2c) presented in Hildebrandt et al. (2005). In Sect. 6.1 the data of the DPS and the selection of the LBGs are described. Simulations to assess the performance of our LBG selection are presented in Sect. 6.2. The clustering analysis is covered in Sect. 6.3. A summary and conclusions are given in Sect. 6.4.

Throughout this chapter we adopt a standard ΛCDM cosmology like before with a normalisation of the power spectrum of $\sigma_8 = 0.9$. We use Vega magnitudes if not stated otherwise.

6.1 The data and the samples

6.1.1 DPS images

For our studies on LBGs we use the eight DPS fields (2.5 sq. deg) with complete coverage in the *UBVRI*-filters, in particular the fields Deep1a, 1b, 1c, 2b, 2c, 3a, 3b, and 3c, respectively. Their properties are summarised in Table 6.1. Details on the data reduction and calibration can be found in Chap. 3 and in Hildebrandt et al. (2006).

Note that the images of the field Deep2c used in the current study are different from the images used in Hildebrandt et al. (2005). More data have become available so that the current images in this field are considerably deeper.

6.1.2 Catalogue extraction

First, the different colour images of one field are trimmed to the same size. The seeing is measured for these images and *SExtractor* (Bertin & Arnouts 1996) is used to create RMS maps representing the local pixel-to-pixel sky background fluctuations. From these RMS images the local limiting magnitude (1σ sky background in a circular aperture of 2×FWHM diameter) in each pixel is calculated and limiting magnitude maps are created. Then every image is convolved with an appropriate Gaussian filter to match the seeing of the image with the worst seeing value.

For the catalogue extraction *SExtractor* is run in dual-image mode with the unconvolved R-band image for source detection and the convolved images in the five bands for photometric flux measurements. In this way it is assured that the same part of a galaxy is measured in every band. In the absence of a strong spatial colour gradient in an object this method should lead to unbiased colours especially for high-redshift objects of small apparent size. In the following, we use isophotal magnitudes when estimating colours or photometric redshifts of objects. The total R-band magnitudes always refer to the *SExtractor* parameter MAG_AUTO measured on the unconvolved R-band image.

Table 6.1: The DPS fields with five colour coverage. The limiting magnitudes in columns two to six are measured in a circular aperture of 2×FWHM diameter from the 1σ sky background fluctuations of the images before convolution with a Gaussian filter (see text). In column seven the seeing FWHM values after convolution are given. Column eight contains the number of objects satisfying Eq. (6.1). In the fields Deep1c, Deep2b, and Deep3a fewer U-dropouts are selected due to inferior quality of the imaging data (see text). The completeness limits for LBG selection in column nine are estimated from Fig. 6.4. These visual estimates are very rough with an accuracy of ~ 0.5 mag. The effective area used for LBG selection is given in the last column.

field	1σ mag lim. [Vega mags]					conv.	N_{LBG}	LBG compl. limit	eff. area
	U	B	V	R	I	seeing		R	[arcmin^2]
Deep1a	27.0	27.8	27.5	27.4	26.3	1.″3	1420	25.0	1045
Deep1b	27.0	27.5	27.1	27.2	26.2	1.″3	1114	24.5	1036
Deep1c	26.7	28.0	27.4	27.1	26.4	1.″3	685	24.0	1031
Deep2b	26.8	28.0	27.5	26.9	26.3	1.″3	492	24.0	1025
Deep2c	27.1	29.0	28.6	28.5	26.3	1.″0	2181	25.5	1064
Deep3a	26.6	27.9	27.2	27.3	25.8	1.″1	456	24.0	1033
Deep3b	26.9	27.9	27.3	27.3	26.2	1.″0	1484	25.0	1072
Deep3c	27.0	28.1	27.3	27.2	25.8	1.″0	1679	25.0	998
							$\Sigma = 9511$		$\Sigma = 8304$

6.1.3 Photometric redshift estimation

Photometric redshifts are estimated for all objects in the catalogue from their *UBVRI* photometry using the publicly available code *Hyperz* (see Chap. 5 Bolzonella et al. 2000). The technique applied is essentially the same as described in Hildebrandt et al. (2005) with the only difference that in the current study we use isophotal magnitudes extracted from images with matched seeing instead of seeing adapted aperture magnitudes. In Chap. 5 and in Hildebrandt et al. (2007c) we compare our photometric redshift estimates to several hundred spectroscopic redshifts from the VIMOS VVDS in the field Deep2c and find that the combination of isophotal magnitudes with the templates created from the library of Bruzual & Charlot (1993) yields the smallest scatter and outlier rates at least for the redshifts probed by the VVDS ($z < 1.4$). For statistics on photo-z on this dataset we refer to Table 5.3 and Fig. 5.5. With increasing redshift and decreasing angular size of the objects the isophotal magnitudes on images with matched seeing should approach the seeing adapted aperture magnitudes. Investigating the redshift distributions of our U-dropout sample (see below) we see no significant difference between the two approaches so that the results of the clustering measurements are not influenced by this choice.

6.1.4 Sample selection

In Hildebrandt et al. (2005) we chose quite conservative criteria for our U-dropout selection. In particular, we tried to define colour cuts in such a way to avoid regions in colour space with a considerable amount of contamination. Supported by our simulations (see Sect. 6.2), and after refining our photometric measurements (see above), which should yield better colours, especially for the possible contaminants, we decided to relax our selection criteria. The performance of our selection in terms of contamination and efficiency is analysed in Sect. 6.2 by means of simulated colour catalogues. In Fig. 6.2 the colour-colour diagram for one field is shown with the old and the new selection criteria represented by the boxes and Fig. 6.1 shows the selection criteria in combination with theoretical colour tracks. We select candidates for $z \sim 3$ LBGs in the following way:

$$\begin{aligned} 0.3 &< (U-V)\ , \\ -0.5 &< (V-R) < 1.5\ , \\ 3(V-R) &< (U-V) + 0.2\ , \end{aligned} \tag{6.1}$$

Furthermore, we require every candidate to be detected in V and R and to be located in a region on our images where the local limiting magnitudes in UVR are not considerably decreased. This is necessary to avoid complex selection effects in colour space due to varying depths over a single field. In our catalogues we find 8826 objects satisfying these selection criteria.

Every candidate is inspected visually in all five filters; see Fig. 6.3 for an example of the visual appearance of one such candidate in the five optical filters. Moreover, the redshift-probability distribution (see Fig. 4.4 for some examples), the location in the field, and the location in the colour-colour diagram is checked for every object. Approximately one third of the candidates is rejected in this way. Most of these rejected objects are influenced by the straylight from bright neighbouring objects (indicated by their extraction flags or visible in the $10'' \times 10''$ thumbnail images) so that their colours cannot be trusted. Merely 72 objects ($< 1\%$) are rejected due to their redshift-probability distribution in combination with a suspicious location in the colour-colour diagram near the stellar locus. Thus, the photometric redshift distribution is almost not affected by the exclusion of these objects.

The magnitude dependent angular number-densities of the accepted candidates for the eight fields and for the whole DPS are displayed in Fig. 6.4 in comparison to the values found by Steidel et al. (1999). For $R < 24$ all eight fields show approximately the same LBG source density within a reasonable field-to-field variance. The field Deep2b can be regarded as fairly complete down to $R = 24.5$ and the fields Deep1a, 3b, and 3c, respectively, down to $R = 25$. But only the field Deep2c shows the same density as the survey by Steidel et al. (1999) down to $R = 25.5$. The underdensity of Deep1b and Deep2b can be

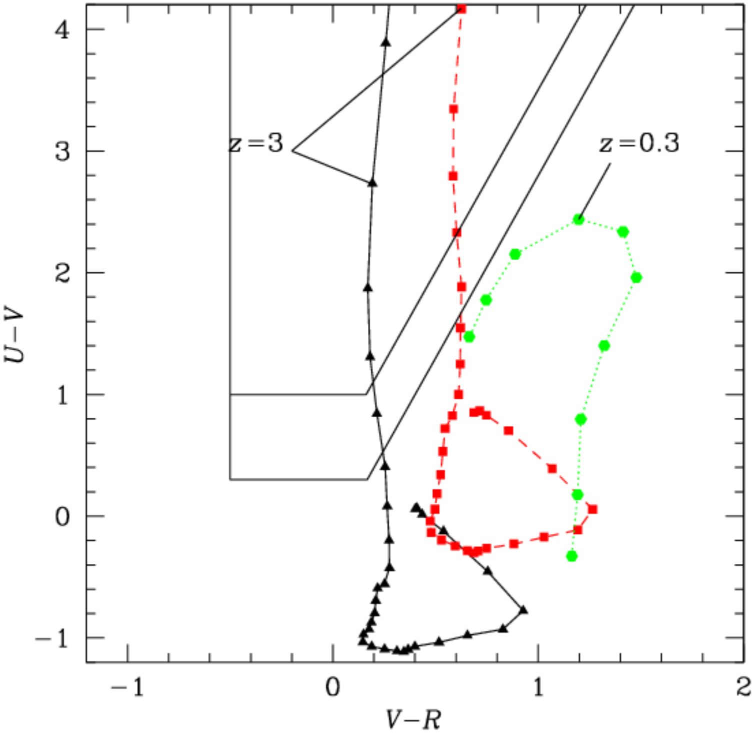

Figure 6.1: Colours of model galaxies in the $(U-V)$ vs. $(V-R)$ colour-colour diagram used for U-dropout selection. The solid line represents galaxies (spectral type Im) with no dust reddening, the dashed line represents galaxies (spectral type Im) with an extinction in the visual of $A_V = 1.5$ mag, and the dotted line represents elliptical galaxies (spectral type E) at low redshift. The points correspond to intervals of $\Delta z = 0.1$. For the Im templates the $(U-V)$ colour increases strongly for redshifts approaching $z = 3$. The upper box represents the old selection criteria adopted in Hildebrandt et al. (2005) while the lower extension represents the selection criteria of the current study.

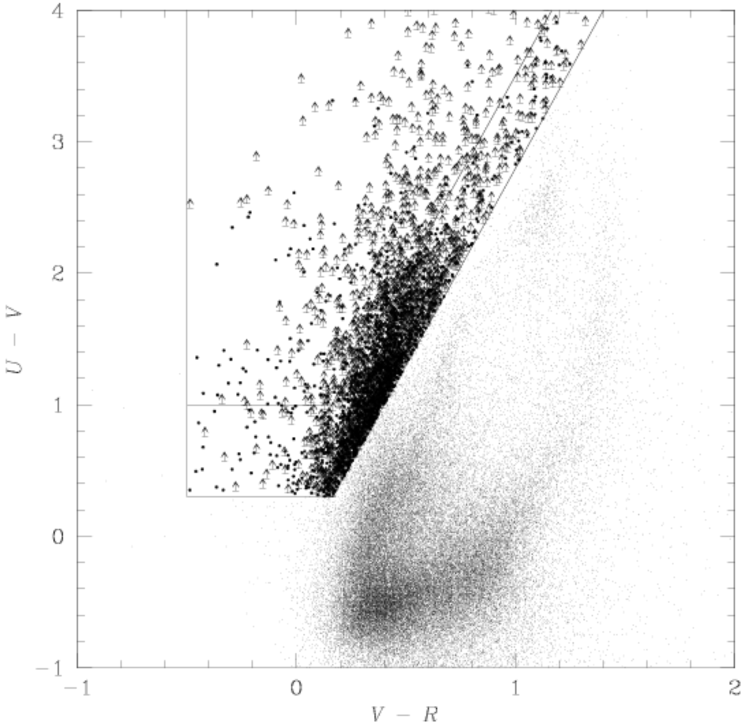

Figure 6.2: $(U-V)$ vs. $(V-R)$ colour-colour diagram of galaxies in the field Deep2c. The upper box represents the old selection criteria adopted in Hildebrandt et al. (2005) while the lower extension represents the selection criteria of the current study. The $(U-V)$ colours of objects that are satisfying the LBG selection criteria but are not detected in the U-band are lower limits (*arrows*). In contrast the detected objects satisfying the selection criteria are plotted as *filled circles*.

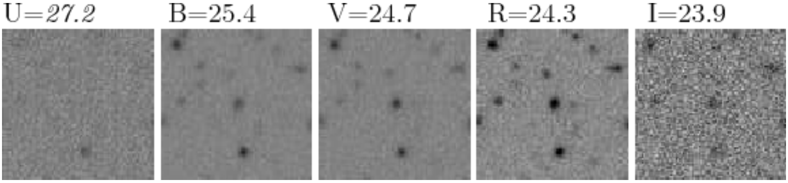

Figure 6.3: Thumbnail pictures in the WFI-$UBVRI$ bands of size $10'' \times 10''$ and measured isophotal magnitudes of one U-dropout candidate. Since the object is not detected in U, the U-band magnitude corresponds to the local 1σ limiting magnitude (indicated by *italics*).

explained by the inferior seeing in the detection images (R) while the underdensity of Deep1c and Deep3a is due to a shallower U-band image.

As long as one restricts investigations in a given magnitude bin to fields that can be regarded as uniform in terms of LBG selection, these investigations are not subject to systematic effects due to varying selection efficiency between the fields. This is in particular important in the clustering analysis to avoid artificial correlations originating from non-uniform depths.

The photometric redshift distribution of all accepted objects is shown in Fig. 6.5. The mean redshift of $\langle z \rangle = 2.96$ (4σ outliers are rejected) is essentially the same as the spectroscopic mean found by Adelberger et al. (2005) for their LBG sample, but our distribution seems to be slightly narrower ($\sigma = 0.24$). However, without spectroscopic information we are not able to judge whether this is due to the different selection method with a different camera and filter set or due to imperfect photometric redshift estimation. In the following we will use the distribution inferred from our photometric redshifts indicating whenever the width has a large influence on our results.

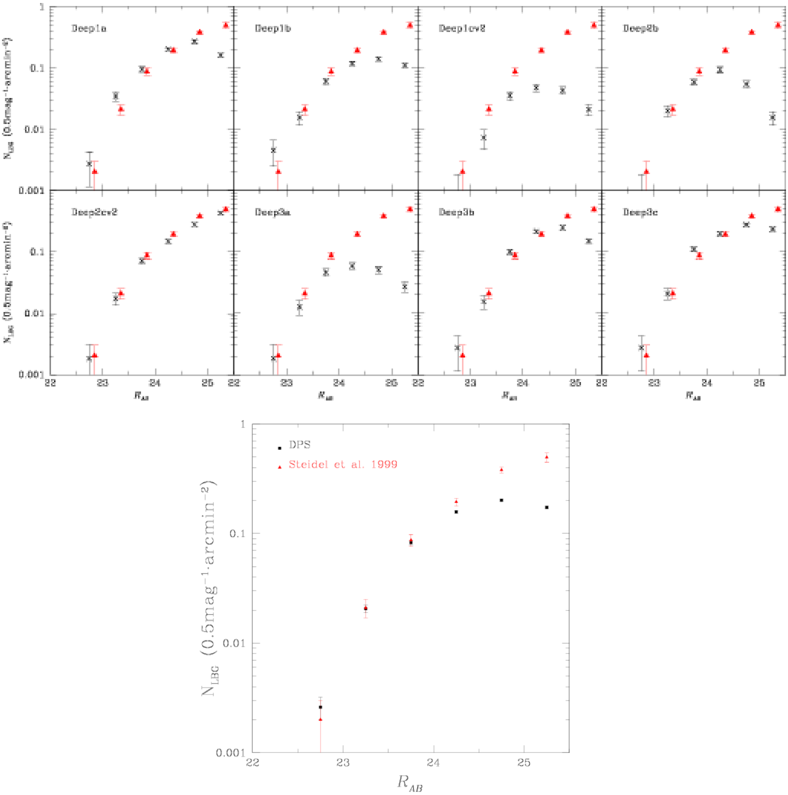

Figure 6.4: Source counts of LBGs as a function of R-band magnitude in the eight DPS fields and for the whole survey. The DPS densities are represented by crosses while the Steidel et al. (1999) densities are represented by triangles which are offset by +0.1 mag just for clarity.

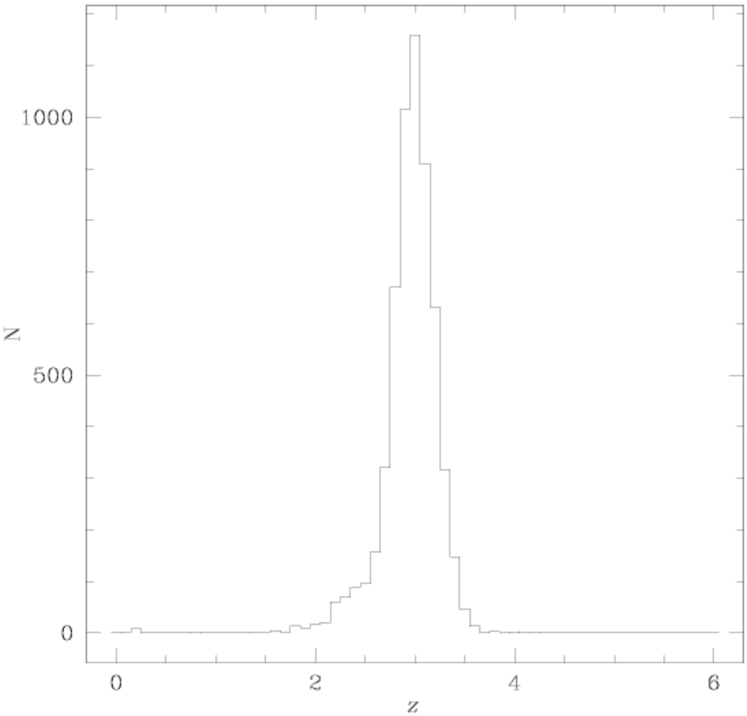

Figure 6.5: Photometric redshift distribution of all accepted LBG candidates. The distribution has a mean of $\langle z \rangle = 2.96$ and an RMS of $\sigma = 0.24$.

6.2 Simulations of objects' colours in the DPS

Without a large spectroscopic survey of our LBGs at hand the only possible way to estimate properties of our LBG samples like, e.g., the contamination rate is to create a simulated colour catalogue.

6.2.1 The role of stars in our sample

The TRILEGAL galactic model by Girardi et al. (2005) is used to simulate the number of stars in all eight fields and their colours in the WFI filter set. In this way we obtain accurate $U - V$ and $V - R$ colours and are able to quantify the amount of stellar contamination in the LBG selection box. In Fig. 6.6 the colours of stars in the field Deep1a are shown, representative for the whole survey. The selection criteria were chosen in such a way that stellar contamination is very low, which is confirmed by Fig. 6.6 (see also the bottom panel of Fig. 6.12).

6.2.2 Colours of galaxies

The code *Hyperz* cannot only be used to estimate photometric redshifts but also to create colour catalogues of galaxies at different redshifts and of different spectral types. We simulate huge random mock catalogues in magnitude bins of width 0.5 mag of 500 000 galaxies each, evenly distributed over the redshift interval $0 < z < 7$ and over all spectral types from the library of Bruzual & Charlot (1993) provided by *Hyperz*. Magnitude errors are simulated by *Hyperz* according to the 1σ limits in the five bands. This is done separately for every field taking into account the different depths in the five bands. From the photometric redshift code *BPZ* (Benítez 2000) we extract magnitude- and spectral type-dependent redshift distributions derived from the Hubble Deep Field. We assign the two reddest *Hyperz* spectral types from the library of Bruzual & Charlot (1993) ("Burst" and "E") to the *BPZ* z-distribution for elliptical galaxies, two intermediate types ("Sb" and "Sc") to the z-distribution for

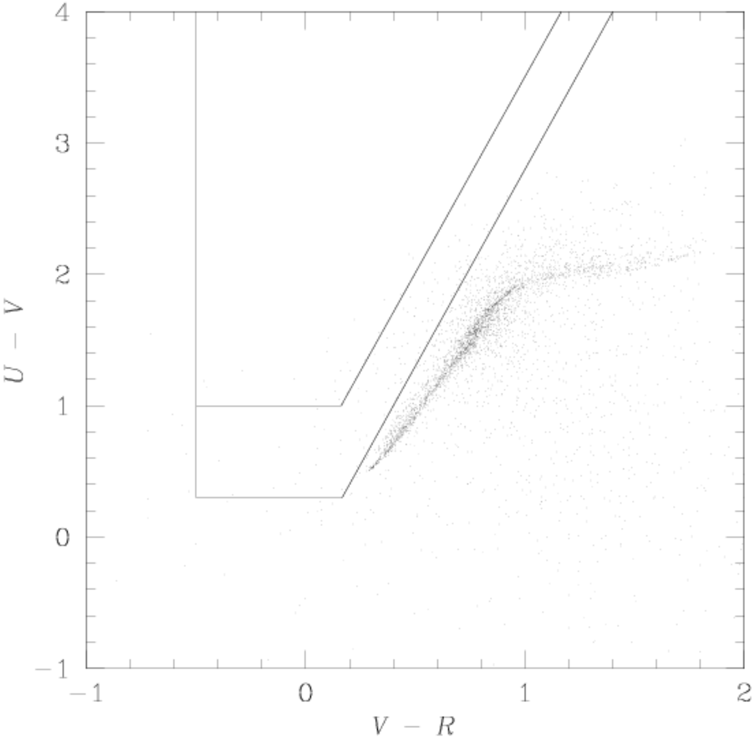

Figure 6.6: $(U - V)$ vs. $(V - R)$ colour-colour diagram of simulated stars in the field Deep1a. The boxes are the same as in Fig. 6.2.

spirals, and the two bluest spectral types ("Sd" and "Im") to the z-distribution for star-forming galaxies. Galaxies are taken from the evenly distributed catalogues with numbers and spectral types according to these redshift distributions to create realistic catalogues for 0.5 mag intervals. Finally, from these catalogues the galaxies are taken with numbers scaled to the I-band number-counts in the eight fields. In this way, for every field a catalogue is created which gives a fair representation of our data in 0.5 mag wide intervals. The $U-V$ vs. $V-R$ colour-colour diagram of the simulated galaxies is shown in Fig. 6.7.

Certainly, it is a strong assumption that the chosen template set represents the galaxy population in our data at all redshifts. Furthermore, the redshift distributions were extracted from the Hubble Deep Field which is subject to strong cosmic variance. These shortcomings are most probably responsible for the slight differences between Fig. 6.2 and Fig. 6.7 and the simulations should only be regarded as rough estimates.

Applying the colour cuts from Eq. (6.1) to the simulated star and galaxy catalogues we obtain estimates for the contamination rate in our LBG sample at different magnitudes which are plotted in the bottom panel of Fig. 6.12 and included in Table 6.2. For the considered magnitude range of $22.5 < R < 26$ the total contamination is below 25%.

We define the completeness in a particular magnitude and redshift bin as the ratio of the number of objects selected by our criteria to the number of objects in the whole catalogue. Since the objects in our mock catalogue are generated in such a way that their magnitude errors are derived from the typical limiting magnitudes in the DPS, with these simulations the completeness of our selection can only be quantified with respect to this catalogue. In Fig. 6.8 the completeness in dependence of R-band magnitude and redshift is shown. The values should be regarded as an upper bound for the total completeness with respect to the whole galaxy population since some of them may be entirely undetectable in our images due to low surface brightness. Certainly, this becomes more serious for $R > 25$ where our LBG number-counts start to drop (see Fig. 6.4).

The redshift distribution of the selected simulated galaxies in the magnitude interval $23 < R < 25$ is shown in Fig. 6.9 with good agreement to the photometric redshift distribution in Fig. 6.5.

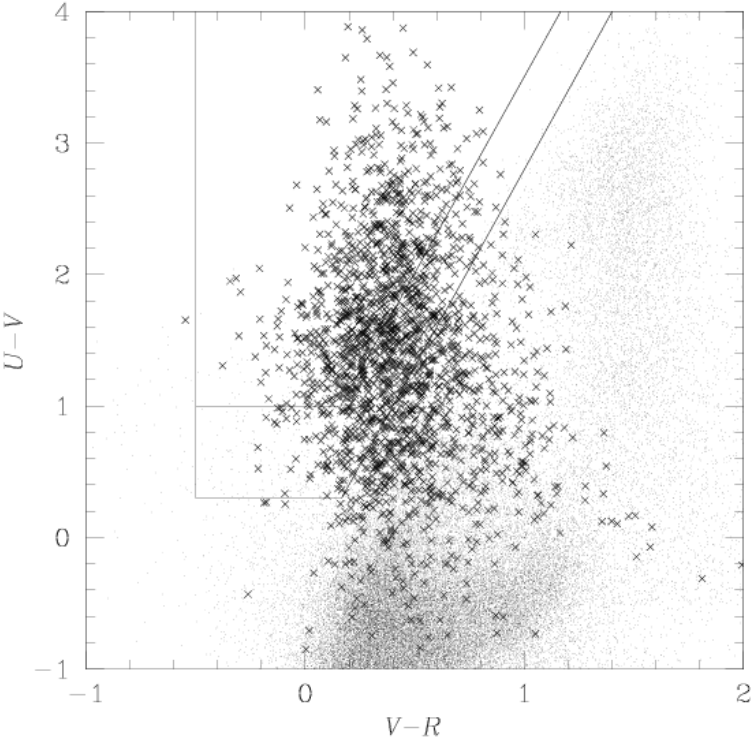

Figure 6.7: $(U - V)$ vs. $(V - R)$ colour-colour diagram of simulated galaxies. The boxes are the same as in Fig. 6.2. Objects with a redshift of $z > 2.5$ are plotted as crosses. We find good agreement in the overall shape of the colour distribution of the simulated galaxies to the real data (see Fig. 6.2). The slight differences, however, may be attributed to an imperfect template set and to a redshift distribution which is subject to cosmic variance.

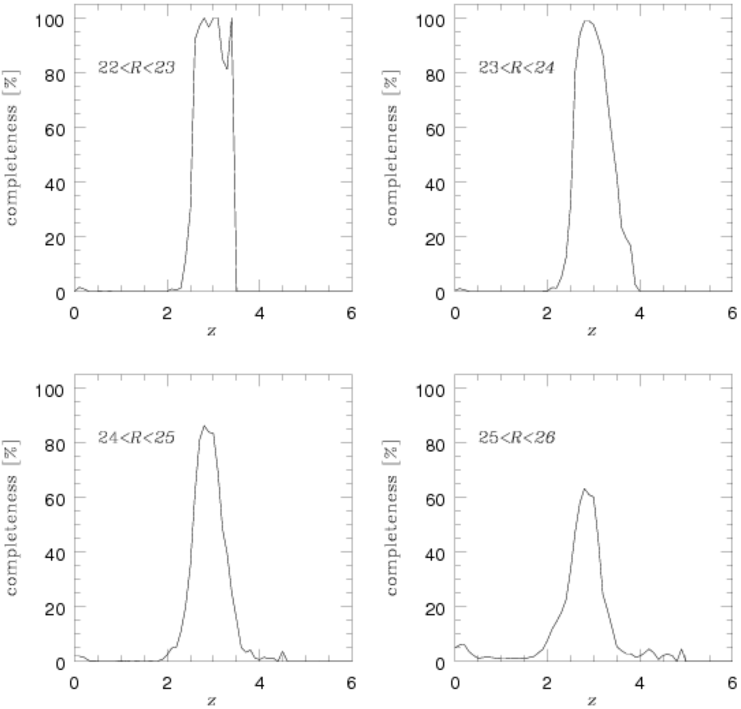

Figure 6.8: Completeness as defined in the text of our LBG selection in dependence of magnitude and redshift averaged over the eight fields.

6.3 Clustering properties

6.3.1 Method

The angular correlation function is estimated as described in 1.4.4 and we apply Poissonian errors for the angular correlation function.

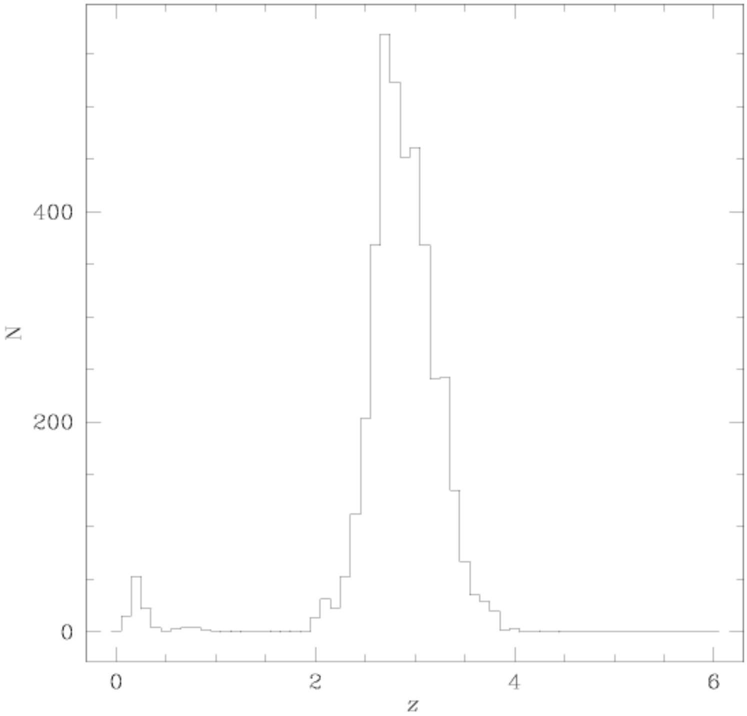

Figure 6.9: Redshift distribution of $23 < R < 25$ (to avoid magnitude regions with higher contamination) galaxies selected from the mock catalogues of Deep1a, 2c, 3b, and 3c by Eq. (6.1) showing good agreement to Fig. 6.5 with only a slightly wider peak. The distribution of galaxies with $2 < z < 4$ has a mean of $\langle z \rangle = 2.89$ and a standard deviation of $\sigma = 0.30$.

Although the area of the DPS is rather large in comparison to previous U-dropout surveys the results may nevertheless be subject to cosmic variance. We estimate the amplitude of the cosmic variance from the field-to-field variance which includes cosmic variance as well as shot-noise from the limited number of galaxies used for the estimate of $\omega(\theta)$. We find values which are comparable in size to the Poissonian errors. This means that the errors are dominated by shot-noise while cosmic variance is negligible. Thus, the application of Poissonian errors is justified.

In Fig. 6.10 the angular correlation function is shown for all U-dropouts with $22.5 < R < 23.5$ of all eight fields as well as with $22.5 < R < 26$ of the fields Deep1a, Deep2c, Deep3b, and Deep3c.

A power law,

$$\omega(\theta) = A_\omega \theta^{-\delta} \,, \tag{6.2}$$

is fitted to the angular correlation function and the Limber equation (see Hildebrandt et al. 2005) is used to estimate the real-space correlation function, ξ.[2] In this step we apply our photometric redshift distribution presented in Sect. 6.1.4. We parametrise the power law approximation of the real-space correlation function in the following form:

$$\xi\left(r\right) = \left(\frac{r}{r_0}\right)^{-\gamma} \,, \tag{6.3}$$

with r being the comoving distance, r_0 being the comoving correlation length, and $\gamma = 1 + \delta$.

We estimate the integral constraint by the method outlined in Sect. 1.4.5 from the linear cold dark matter (CDM) power spectrum. The survey volume of $z = 3$ LBGs in a single DPS field can be reasonably approximated by a square on the sky (comoving dimensions of $42 \times 42 (h^{-1}\mathrm{Mpc})^2$ at redshift $z = 3$) and a Gaussian in radial direction ($\sigma = 88\, h^{-1}\mathrm{Mpc}$) resulting in $\sigma^2_{\mathrm{CDM}} = 0.0017$.

---

[2] Note that the Limber equation is inaccurate to some degree due to the relatively narrow distribution in comoving distance of our LBGs. See Simon (2006) for details. This is certainly also the case for other LBG studies applying the Limber equation so that the relative comparisons presented here are not affected seriously.

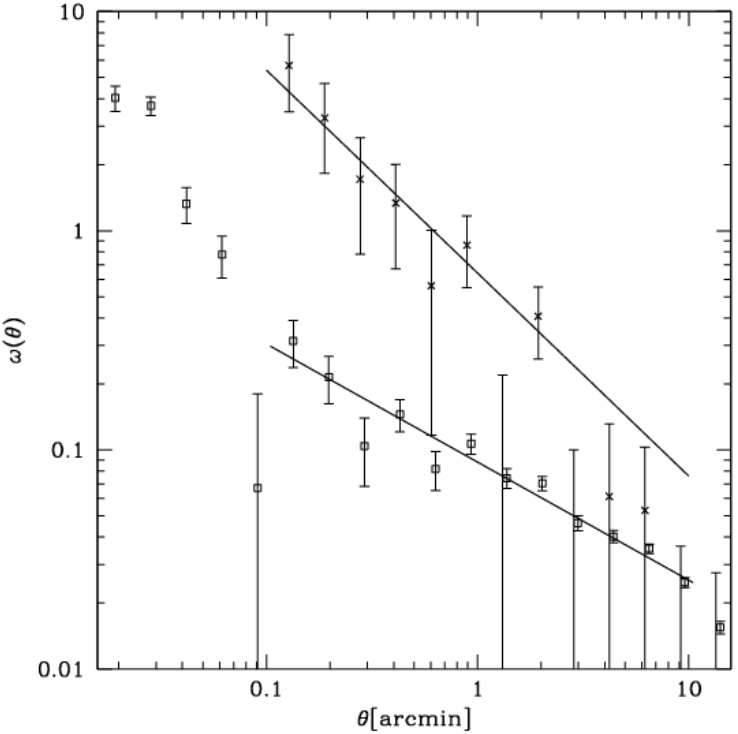

Figure 6.10: Angular correlation function for U-dropouts with $22.5 < R < 23.5$ (crosses) of all eight fields and with $22.5 < R < 26$ (open squares, slightly offset for clarity) of the fields Deep1a, Deep2c, Deep3b, and Deep3c. The solid lines represent power law fits to the data in the range $0.1' < \theta < 10'$. The angular correlation function of the faint sample shows an excess on small scales with respect to the power law fitted to the data on intermediate scales.

6.3.2 Results

We investigate the influence of the angular binning of the data by estimating the correlation length for 10 to 25 bins and find that the standard deviation of r_0 over these 16 binnings is comparable or smaller than the error introduced by the fitting for $R_{\rm lim} \geq 24.5$. We choose a common binning for these magnitude intervals which samples the correlation function well. The errors on the correlation lengths are then derived from Monte-Carlo simulations taking into account the fitting errors for A_ω and δ and assuming that these are Gaussian and uncorrelated. In Fig. 6.11 the confidence regions for A_ω and δ for the $22.5 < R < 26$ subsample are plotted as 2-dimensional contours.

For the $22.5 < R < 23.5$ and $22.5 < R < 24$ subsamples the error introduced by the binning is larger than the fitting error. Therefore, we decided to average over the different binnings and report the standard deviation as the error of r_0 for these two magnitude ranges. By doing so we avoid an underestimation of the error on r_0.

In Table 6.2 the results for the various subsamples of the LBGs are presented in comparison to results from previous studies.

There are, however, still systematic uncertainties in our correlation analysis, the most serious being our selection function. We must rely on the validity of our photometric redshift distribution shown in Fig. 6.5. The derived correlation lengths certainly depend on the width of this distribution, with wider distributions resulting in larger correlation lengths. Moreover, the results from the simulations in Sect. 6.2 may be subject to cosmic variance since the underlying redshift distributions were derived from the small HDF. Thus, our contamination and completeness estimates must be regarded as approximate values.

Furthermore, we cannot correct our clustering measurements for contamination directly as there are no spectroscopic observations of our WFI-selected LBG samples available yet. In general, a contamination rate f of uncorrelated sources in our catalogues will lead to an angular correlation function with a measured amplitude $A = (1 - f)^2 \, A_{\rm real}$ implying a corrected correlation length

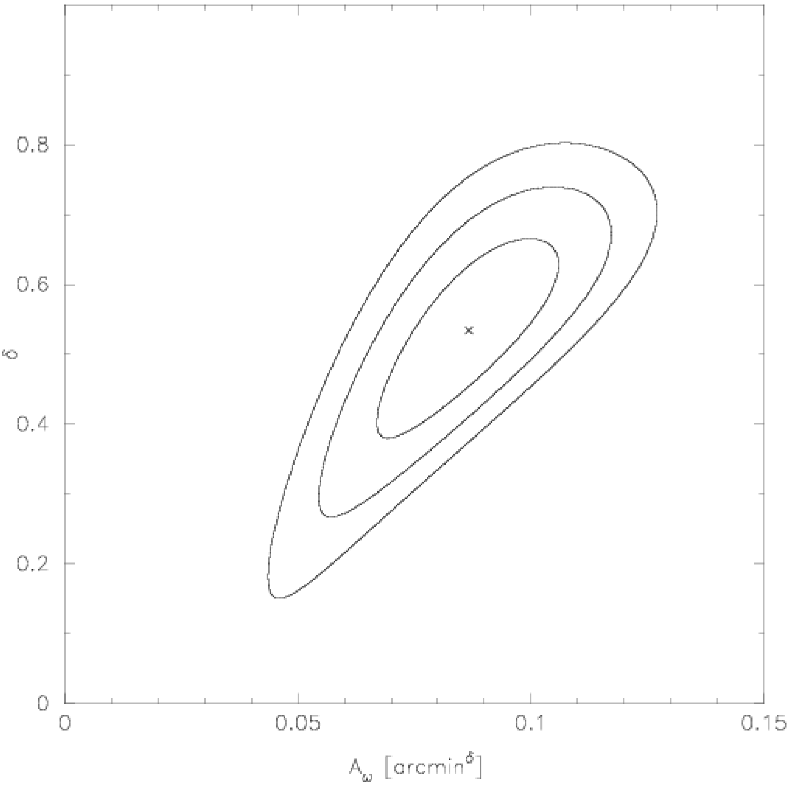

Figure 6.11: Joint 68.3%-, 95.4%-, and 99.8%-confidence regions (corresponding to $\Delta\chi^2 = [2.3; 6.2; 11.8]$) for the power law parameters A_ω and δ (see Eq. 6.2) in the fit to the angular correlation function of the $22.5 < R < 26$ subsample.

$r_{0,\mathrm{corr}} = (1-f)^{-2/\gamma} r_0$. However, as we do not know from our simulations the exact clustering behaviour of the contaminants we do not apply such a correction.

We see clustering segregation with rest-frame UV luminosity in our data. In Fig. 6.12 the dependence of the correlation lengths and the slope of the correlation function are plotted against limiting magnitude along with the contamination estimates from Sect. 6.2.2. The correlation lengths for the different subsamples decrease monotonically with limiting magnitude down to $R_{\mathrm{lim}} = 25$ and then stay constant whereas the slope decreases down to $R_{\mathrm{lim}} = 24.5$.

The observation that more luminous LBGs show larger correlation lengths was reported quite some time ago (see e.g. Giavalisco & Dickinson 2001; Ouchi et al. 2001). In the CDM framework more massive halos are more strongly biased than less massive halos with respect to the whole mass distribution. Thus, brighter LBGs are supposed to be hosted by more massive halos than fainter ones. For a quantitative analysis of halo properties see Sect. 6.3.3.

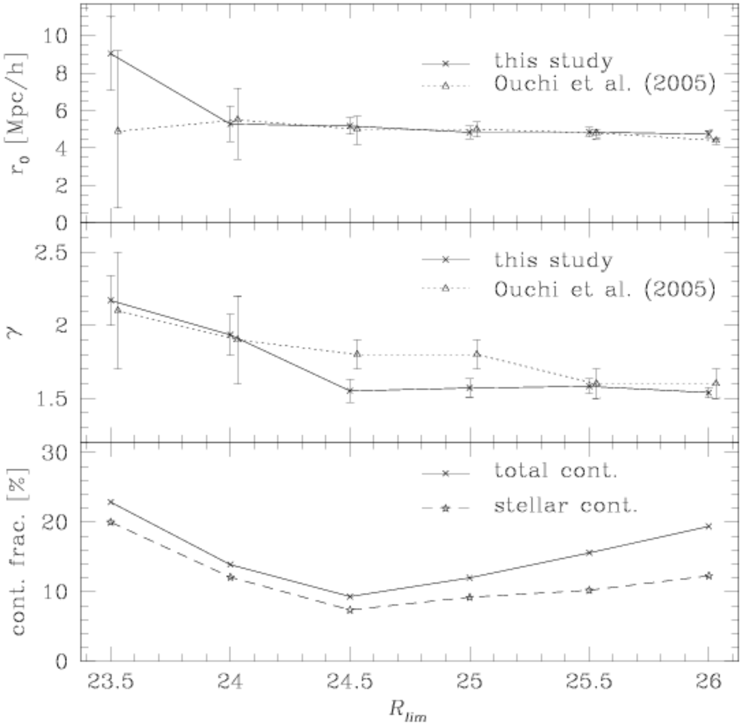

Figure 6.12: Dependence of the correlation length and the slope of the correlation function on R-band limiting magnitude (*upper* and *middle* panels, respectively). The crosses and the solid lines represent our data. The triangles (slightly offset for clarity) and the dotted lines show the values from Ouchi et al. (2005) at $z = 4$ in comparison. We relate their i'_{AB} limiting magnitudes to our R_{Vega} limiting magnitudes as described in the text. The lower panel shows the contamination fraction for the different magnitude-limited subsamples with $22.5 < R < R_{\mathrm{lim}}$ with the solid line representing the total contamination and the dashed line representing the stellar contamination.

Table 6.2: Clustering measurements of LBGs in the DPS fields for different limiting magnitudes. LBGs are cut at the bright end at $R \geq 22.5$. Results from the survey of Ouchi et al. (2005) are shown below for comparison. The power law fits to the angular correlation function are performed in the range $0\overset{\prime}{.}1 \leq \theta \leq 10'$. The errors on the correlation lengths are estimated from the fitting errors of the slope and the amplitude of the angular correlation function for $R_{\rm lim} \geq 24.5$ and from the scatter of different binnings for $R_{\rm lim} \leq 24$, respectively. No possible systematical uncertainties introduced by e.g. the photometric redshift distribution are included. Notice that the $22.5 < R < 25.5$ and $22.5 < R < 26$ samples suffer from some incompleteness at the faint end which may be the reason for non-evolution from $R_{\rm lim} = 25$ to $R_{\rm lim} = 26$. In the second column the number of fields are listed with “8” corresponding to all eight DPS fields with *UBVRI* coverage, “5” corresponding to the fields Deep1a, 1b, 2c, 3b, and 3c, and “4” corresponding to Deep1a, 2c, 3b, and 3c, respectively. The values for the integral constraint, IC, and the linear bias factor, b, in the sixth and seventh column are estimated as described in Sect. 6.3.1 and the contamination fractions, f, in column eight are derived from the simulations presented in Sect. 6.2. The average number of LBGs per halo, $\langle N_{\rm g} \rangle$, and the average mass of an LBG hosting halo, $\langle M_{\rm halo} \rangle$ are estimated as detailed in Sect. 6.3.3. Note that the $i'_{{\rm AB},z=4}$ limiting magnitudes in the study by Ouchi et al. (2005) can be related to our $R_{\rm Vega}$ limiting magnitudes by $i'_{{\rm AB},z=4,{\rm lim}} \hat{=} R_{\rm Vega,lim} + 1$ as described in the text.

Sample	No. fields	N	γ	r_0 $[h^{-1}\mathrm{Mpc}]$	IC	b	f [%]	$\langle N_{\mathrm{g}} \rangle$	$\log \langle M_{\mathrm{halo}} \rangle$ $[h^{-1}M_{\odot}]$
$R \leq 23.5$	8	280	2.17 ± 0.17	9.0 ± 2.0	0.071	6.6	22.9	0.44 ± 0.85	$12.79^{+0.04}_{-0.04}$
$R \leq 24.0$	8	1064	1.93 ± 0.14	6.3 ± 0.6	0.019	3.4	13.9	0.41 ± 0.29	$12.46^{+0.02}_{-0.02}$
$R \leq 24.5$	5	1864	1.55 ± 0.08	5.2 ± 0.4	0.015	3.0	9.3	0.42 ± 0.24	$12.29^{+0.09}_{-0.10}$
$R \leq 25.0$	4	2950	1.57 ± 0.06	4.8 ± 0.3	0.013	2.8	12.0	0.61 ± 0.51	$12.12^{+0.10}_{-0.12}$
$R \leq 25.5$	4	3913	1.58 ± 0.05	4.8 ± 0.3	0.014	2.9	15.6	0.44 ± 0.23	$12.15^{+0.15}_{-0.24}$
$R \leq 26.0$	4	4363	1.54 ± 0.04	4.8 ± 0.2	0.014	2.9	19.4	0.50 ± 0.24	$12.15^{+0.14}_{-0.20}$
Ouchi et al. (2005), $z = 4$									
$i' < 24.5$	–	239	2.1 ± 0.4	$4.9^{+4.3}_{-4.1}$	–	–	–	$0.2^{+0.2}_{-0.2}$	$12.3^{+0.1}_{-0.6}$
$i' < 25.0$	–	808	1.9 ± 0.3	$5.5^{+1.7}_{-2.1}$	–	–	–	$0.3^{+0.4}_{-0.3}$	$12.3^{+0.1}_{-0.2}$
$i' < 25.5$	–	2231	1.8 ± 0.1	$5.0^{+0.7}_{-0.8}$	–	–	–	$0.6^{+0.1}_{-0.5}$	$12.1^{+0.1}_{-0.1}$
$i' < 26.0$	–	4891	1.8 ± 0.1	$5.0^{+0.4}_{-0.4}$	–	–	–	$0.6^{+0.1}_{-0.1}$	$12.0^{+0.1}_{-0.1}$
$i' < 26.5$	–	8639	1.6 ± 0.1	$4.8^{+0.2}_{-0.3}$	–	–	–	$0.6^{+0.1}_{-0.1}$	$11.9^{+0.05}_{-0.05}$
$i' < 27.0$	–	12 921	1.6 ± 0.1	$4.4^{+0.1}_{-0.2}$	–	–	–	$0.6^{+0.1}_{-0.2}$	$11.8^{+0.07}_{-0.04}$

Table 6.4: Clustering measurement in comparison to Adelberger et al. (2005). at $z = 3$.

	Sample	γ	r_0 $[h^{-1}\mathrm{Mpc}]$
this study	$23.3 \leq R \leq 25.3$	1.60 ± 0.03	5.0 ± 0.2
A2005	$23.5 < R_{\mathrm{AB}} < 25.5$	1.57 ± 0.14	4.0 ± 0.6

We compare our results to precise recent measurements of $z = 4$ LBG clustering by Ouchi et al. (2005). For these comparisons our R_{Vega} limits must be converted to the AB system (+0.2 mag) and the distance modulus between $z = 3$ and $z = 4$ must be added (+0.8 mag for ΛCDM). Since the R-band at $z = 3$ closely resembles the I-band at $z = 4$ in terms of restframe wavelength coverage we do not apply a k-correction. The agreement of both studies shown in Table 6.2 and Fig. 6.12 is excellent with most corresponding measurements lying within the 1σ intervals. Considering the systematic differences introduced by different filter-sets, different depths, different selection criteria, etc., the agreement is rather impressive. However, considering the cosmic time and the structure formation that took place between $z = 4$ and $z = 3$ this means that an LBG at $z = 3$ is hosted by a significantly more massive halo than an LBG of the same luminosity at $z = 4$ (see Sect. 6.3.3).

Adelberger et al. (2005) confine their samples of optically selected, star-forming galaxies to the magnitude range $23.5 \leq R_{\mathrm{AB}} \leq 25.5$, which corresponds to $23.3 \leq R_{\mathrm{Vega}} \leq 25.3$. In this magnitude range the correlation length derived from our survey is slightly larger than the value found by Adelberger et al. (2005) while the slopes found in both studies agree very well within the uncertainties (see Table 6.2). To compare the depth of our images with the ones used in Adelberger et al. (2005) we calculate 1σ AB limiting magnitudes in apertures with an area that is three times as large as the seeing disk like in Steidel et al. (2003) where the imaging data used by Adelberger et al. (2005) are described. We find that our images are slightly shallower in all three bands used for the selection of LBGs and thus, our larger correlation length may be due to incompleteness at the faint end of the $23.3 < R < 25.3$ magnitude interval with the Adelberger et al. (2005) LBG sample probing slightly deeper into the luminosity function. The same problem certainly applies for the $22.5 < R < 25.5$ and $22.5 < R < 26$ subsamples; this might be the reason for the non-evolution of the correlation length for limiting magnitudes $R_{\mathrm{lim}} > 25$.

6.3.3 Small-scale clustering

With the unprecedented statistical accuracy of our survey it is for the first time possible to clearly detect an excess of the angular correlation function of faint U-dropouts on small scales with respect to a power law fit. In Fig. 6.13 the deviation of the angular correlation function with respect to the power law fitted at large to intermediate scales is shown. Ouchi et al. (2005) and Lee et al. (2006) report such a small-scale excess for $z = 4$ LBG samples from the Subaru/XMM-Newton Deep Field and the GOODS fields, respectively.

This excess on small scales is interpreted in both studies as being due to the contribution from a 1-halo term of galaxy pairs residing in the same halos. We apply the halo model by Hamana et al. (2004) described in Sect. 1.4 to our data to have a direct comparison with the $z = 4$ results from Ouchi et al. (2005) who use the same model.

Applying a combined maximum likelihood fit to the angular correlation functions and the number densities we find the best-fitting model parameters for the different magnitude limited subsamples. It should be noted that there is quite some degeneracy in the parameters $\mathrm{M}_{\mathrm{min}}$, M_1, and α which is illustrated in the two-dimensional contour plots of Fig. 6.14. From these best-fit parameters we calculate the average mass of an LBG hosting halo, $\langle M_{\mathrm{halo}} \rangle$, and the average number of galaxies inside this halo, $\langle N_{\mathrm{g}} \rangle$ which are also tabulated in Table 6.2.

Given the good agreement between our correlation functions at $z = 3$ and the corresponding ones from Ouchi et al. (2005) at $z = 4$, and given the structure growth of the dark matter density field between $z = 3$ and $z = 4$ it is not surprising that we get slightly larger halo masses. This would imply that star formation, which

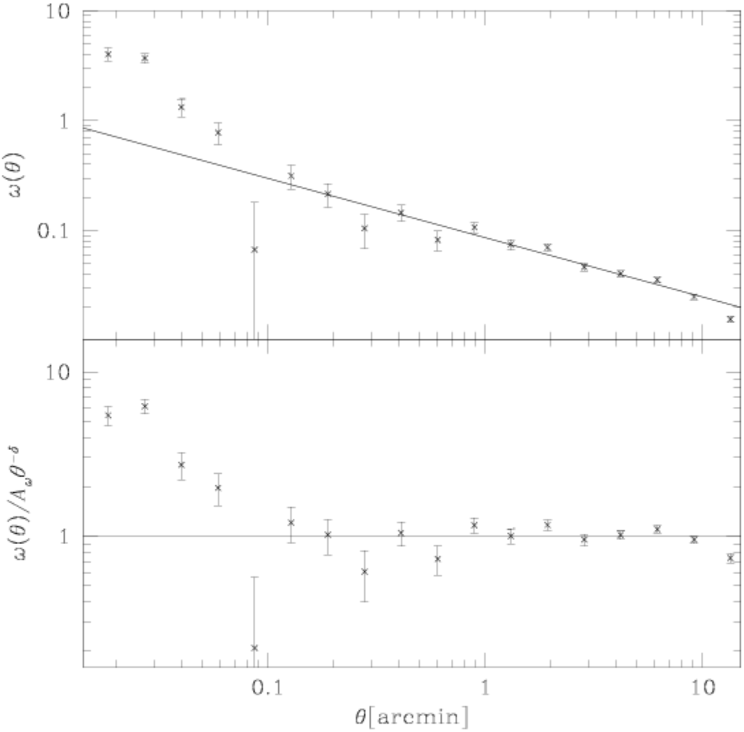

Figure 6.13: *Upper* panel: Angular correlation function for U-dropouts with $22.5 < R < 26$. The solid line represents a power law fit to the data in the range $0\farcm1 < \theta < 10'$. *Lower* panel: Ratios of the angular correlation function to the best-fit power law with a significant excess on small scales.

is mostly responsible for the restframe UV flux, was slightly more efficient at higher redshift. However, the evolution in halo mass is rather small and not very significant. Judging from the residual χ^2

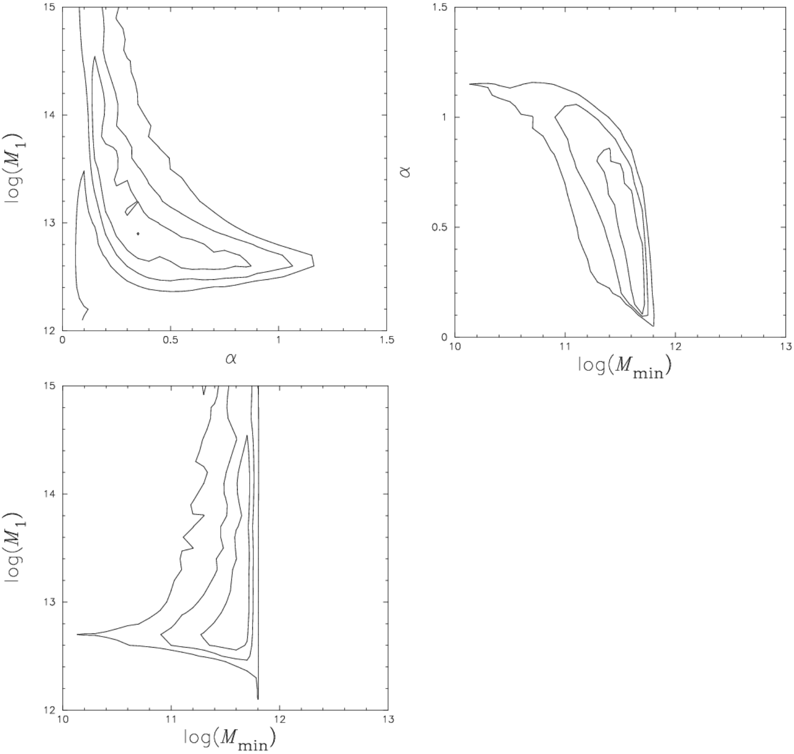

Figure 6.14: Marginalised likelihood contours of the halo-model parameters $M_{\rm min}$, M_1, and α for the $R_{\rm lim} = 25$ subsample. The lines represent to the joint 68.3%-, 95.4%-, and 99.8%-confidence regions (corresponding to $\Delta\chi^2 = [2.3; 6.2; 11.8]$).

values for the best-fit parameters this model is still too simple to account for the shape of the angular correlation functions and the number densities simultaneously.

The mean number of LBGs per halo is well below one. This means that there are a lot of halos which are not occupied by LBGs down to the particular flux-limit. Nevertheless, this does not mean that these are dark matter halos that do not host a galaxy. Massive galaxies that are not actively forming stars may be very faint in the restframe UV and have such red colours that they can easily escape our Lyman-break selection technique. Other techniques incorporating near-IR data must be used to select these populations (see Franx et al. 2003; van Dokkum et al. 2003; Daddi et al. 2004).

6.4 Conclusions and outlook

We measure the clustering properties of a large sample of U-dropouts from the ESO Deep Public Survey with unprecedented statistical accuracy at this redshift.

Candidates are selected via the well-known Lyman-break technique and the selection efficiency is investigated and optimised by means of simulated colour catalogues. The angular correlation function of LBGs is estimated over an area of two square degrees, depending on depth, and a deprojection with the help of the photometric redshift distribution yields estimates for the correlation lengths of different subsamples.

We find clustering segregation with restframe UV-luminosity indicated by a decreasing correlation length and a decreasing slope of the correlation function with increasing limiting magnitude. The latter result was reported at redshift $z = 4$ and is now confirmed at redshift $z = 3$ for the first time. Furthermore, the unprecedented statistical accuracy of our survey at $z = 3$ allows us to study the small-scale clustering signal in detail. We find an excess of the angular correlation function on small angular scales similar to that found previously at $z = 4$.

Applying a halo model we find average masses for LBG-hosting halos at $z = 3$ which are slightly larger than literature values for $z = 4$ implying decreasing star-formation efficiency with decreasing redshift.

The U-dropout sample studied here will be investigated further by means of a spectroscopic campaign which was carried out with VIMOS@VLT (ESO ID: 077.A-0249, P.I. Hildebrandt). Several hundred spectra of LBGs in the field Deep1a will be available after data reduction allowing for the better characterisation of the redshift distribution, the calibration of the photometric redshifts, and the detailed study of the astrophysical properties of these galaxies. Furthermore, we plan to apply this technique to the Canada France Hawaii Telescope Legacy Survey (CFHTLS) Deep fields which are deeper than the DPS and have a larger area. Number statistics of LBGs in the CFHTLS may be sufficient to study the cross-correlation of background LBGs and foreground low-z galaxies induced by cosmic magnification, a weak gravitational lensing effect. See e.g. Scranton et al. (2005) for a similar study with high-z quasars in the SDSS.

7

Summary

Chapter 1: Cosmological framework

Starting from Einstein's field equation the theoretical framework of cosmology is briefly laid down. The field equation can be solved by the Robertson-Walker metric. Additionally, the Friedmann equations are needed to describe the dynamics of the homogeneous and isotropic universe. Introducing the various constituents of the energy density of the universe together with their equations of state, yields a description of the temporal evolution of the cosmic expansion. I define the cosmological parameters like the density parameters and the Hubble constant.

The second section of this chapter deals with the cosmological redshift and distances. After the definition of redshift it is shown how this observable is related to distance measures in cosmology. The differences between the different distance measures are pointed out and it is explained for which applications the different distance measures are used.

Turning from the homogeneous and isotropic universe to the structured universe, it is first presented how structure formation proceeds in the linear regime under Newtonian gravity and how the structures that are formed can be described statistically by the correlation function and the power spectrum. I introduce the primordial power spectrum, setting the boundary conditions of structure formation, as well as the transfer function accounting for deviations from the pressure-less Newtonian case and for the behaviour of different hypothetical dark matter particles. Applying the spherical collapse model it is shown how an analytic expression for the mass function of dark matter halos can be found.

The last section covers the transition from the dark matter structures to the structures observed in the galaxy population which can be understood in the framework of the halo model. The ingredients and assumptions that go into this model are presented together with the model predictions for the galaxy number density, the galaxy power spectrum, and the galaxy two-point correlation function. The chapter concludes with the estimation of correlation functions from real data and their projection onto the sky, both being important theoretical tools for the observing cosmologist.

Chapter 2: Imaging data reduction

The general concepts of imaging data reduction for wide-field CCD camera data are reviewed in this chapter. An emphasis is given to the methods carried out by our data reduction pipeline THELI, the structure of which is briefly discussed.

The different pre-processing steps, including overscan correction, debiasing, flat-fielding, superflatting, fringe-removal, and weight-image creation, are described which are necessary to take out the instrumental signatures of the telescope and the camera.

For the scientific exploitation of these pre-processed data several calibration steps mapping instrumental quantities to physical quantities must be carried out. I present the details of the astrometric as well as the relative and absolute photometric calibrations.

Deep data of the kind that are analysed in the following chapters require a coaddition of many single exposures at the end of data reduction. This complex procedure is detailed in the last section of this chapter including different practical approaches.

Chapter 3: The ESO Deep Public Survey

In this chapter that is largely based on a publication (Hildebrandt et al. 2006), the optical data of the ESO Deep-Public-Survey observed with the Wide Field Imager and reduced with the THELI pipeline are described. These data are the basis of the scientific analyses in later chapters.

After a short introduction to the instrument and the available raw data, we describe the data reduction with emphasis on issues that are special to this dataset. A large part is dedicated to the absolute photometric calibration, arguably one of the most difficult parts in imaging data reduction.

We present 64 fully reduced and stacked images which are publicly released to the community. The astrometric and photometric calibrations are discussed and the properties of the images are compared to images released by the ESO Imaging Survey (EIS) team covering a subset of our data. Thereby, discrepancies in the photometric calibration were found and a collaboration with the EIS team revealed the error to originate from the EIS images.

Chapter 4: Photometric redshifts

The photometric redshift (photo-z) technique is widely used to estimate approximate redshifts for large numbers of galaxies from their colours without taking spectra. In this chapter I introduce the principles of the two different approaches to estimate photo-z's, and I discuss the advantages and disadvantages of both methods.

The SED-fitting method is covered in more detail since it is extensively used in the following chapters.Empirical and synthetic template sets can be used to estimate photo-z's and I describe the steps to create a synthetic template set from a stellar spectral evolution library. Then, the maximum likelihood criteria in the framework of photo-z's from SED-fitting are formulated. Closely connected to this maximum likelihood method is the definition of meaningful redshift confidence intervals.

There are some general problems associated with photo-z's coming from SED-fitting. I discuss these problems and their possible solutions, the one being related to spectroscopic re-calibration and the other one being related to the introduction of priors in the framework of Bayesian statistics. Three different photo-z codes are used in the subsequent chapter of this thesis and therefore their characteristics are summarised.

Finally, I give an overview of a special photo-z technique, the

Lyman-break technique, which is used for the efficient selection of large samples of galaxies at high redshift from just three colours.

Chapter 5: Photometric redshifts in comparison

A blind test of the three different photo-z codes described in the previous chapter on imaging datasets with different depths and filter coverage is performed, and the results are compared to large spectroscopic catalogues. The content of this chapter is part of a manuscript that was submitted to the journal Astronomy & Astrophysics (Hildebrandt et al. 2007c).

By analysing the performance on several subsamples of the complete object catalogue we find criteria to identify objects with uncertain photo-z estimates for the three codes. It is moreover found that the redshift confidence intervals put out by today's codes are only very weakly correlated with the real uncertainty and thus of limited use.

For any given dataset, we find very significant differences in redshift accuracy and outlier rates between the different codes. However, different codes excel in different regimes. We find significant biases in many of the tested setups and we show how a considerable performance increase can be achieved with two of the three codes after applying a complex spectroscopic re-calibration.

Chapter 6: Lyman-break Galaxies in the DPS

A population of ~ 8800 candidates for star-forming galaxies at $z = 3$ is selected via the well-known Lyman-break technique from the ESO Deep Public Survey data presented in the third chapter. This work is published in Hildebrandt et al. (2005, 2007b).

The selection efficiency, contamination rate, and redshift distribution of this population are investigated by means of extensive simulations. Photometric redshifts are estimated for every Lyman-break galaxy (LBG) candidate from its $UBVRI$ photometry yielding an empirical redshift distribution. The measured angular correlation function is deprojected and the resulting spatial correlation lengths and slopes of the correlation function of different subsamples are compared to previous studies.

By fitting a simple power law to the correlation function we do not see an evolution in the correlation length and the slope from other studies at $z = 4$ to our study at $z = 3$. In particular, the dependence of the slope on UV-luminosity similar to that recently detected for a sample of B-dropouts is confirmed also for our U-dropouts. For the first time number statistics for U-dropouts are sufficient to clearly detect a departure from a pure power law on small scales down to $\sim 2''$ reported by other groups for B-dropouts. Applying a halo model fit to our correlation functions we derive mean halo masses which are slightly in excess of the ones of similar bright LBGs at $z = 4$. This result implies a decreasing star-formation efficiency over cosmic time.

A

Photometric calibrations

A.1 Tables with photometric solutions

In Tables A.1 to A.8 the photometric solutions for all calibrated nights of the DPS are shown. The solution chosen for a particular coadded image can be found in the FITS image header.

Table A.1: Photometric solutions in the U_{35060}-band. The colour term corresponds to the Johnson-Cousins colour $U-B$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = −0.48 and the default value for the colour term in the one-parameter fit is CT1 = 0.05.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2000-10-26	665	20.48	−0.00	−10.04	21.15	−10.21	22.06
2000-10-27	666	22.11	−0.47	−0.11	22.12	−0.11	22.12
2000-11-27	697	23.81	−1.93	0.02	22.13	0.02	22.11
2000-11-28	698	27.06	−4.64	0.05	22.12	0.05	22.12
2001-02-21	783	21.21	0.00	0.02	21.79	0.02	21.77
2001-02-22	784	21.35	0.00	0.09	21.95	0.10	21.97
2001-02-23	785	21.78	−0.26	0.10	22.07	0.10	22.09
2001-02-26	788	21.85	−0.37	0.08	21.99	0.08	22.00
2001-03-24	814	22.19	−0.62	0.07	22.02	0.07	22.03
2001-03-25	815	22.55	−0.94	0.09	21.99	0.09	22.00
2001-04-20	841	23.09	−1.38	0.10	21.78	0.10	21.80
2001-04-21	842	21.88	−0.36	0.07	22.04	0.06	22.04
2001-07-20	932	25.73	−3.23	0.23	22.04	0.22	22.17
2001-07-21	933	23.36	−1.51	0.37	21.92	0.44	22.11
2001-07-22	934	22.82	−0.87	0.14	22.33	0.13	22.38
2001-07-23	935	21.58	−0.00	0.11	22.21	0.11	22.26
2001-07-24	936	22.24	−0.40	0.19	22.34	0.19	22.41

Table A.1: Photometric solutions in the U_{35060}-band.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2001-07-25	937	-	-	-	22.05	0.41	22.31
2001-07-26	938	-	-	-	21.92	0.52	22.21
2001-08-19	962	22.46	−0.54	0.08	22.37	0.08	22.38
2001-08-20	963	22.40	−0.51	0.09	22.37	0.09	22.39
2001-08-21	964	22.90	−0.91	0.04	22.37	0.05	22.37
2001-08-22	965	26.52	−3.95	0.02	22.35	0.05	22.35
2001-08-23	966	22.35	−0.48	0.06	22.35	0.06	22.35
2001-11-13	1048	21.73	0.00	0.01	22.29	0.01	22.27
2001-11-14	1049	22.97	−1.06	0.08	22.29	0.09	22.31
2001-11-15	1050	22.78	−0.90	0.05	22.28	0.05	22.28
2002-03-09	1164	22.10	−0.47	0.04	22.12	0.04	22.12
2002-03-10	1165	22.01	−0.45	0.16	22.04	0.16	22.10
2002-03-11	1166	21.95	−0.28	0.11	22.20	0.12	22.23
2002-03-12	1167	22.12	−0.37	0.05	22.26	0.06	22.26
2002-03-13	1168	21.73	0.00	0.03	22.28	0.03	22.27
2002-06-07	1254	21.49	0.00	0.57	22.07	0.56	22.33
2002-06-09	1256	22.88	−1.00	−0.11	22.24	−0.08	22.23
2002-06-10	1257	22.18	−0.39	0.08	22.28	0.08	22.29
2002-06-12	1259	-	-	-	21.82	2.50	22.06
2002-08-11	1319	22.02	−0.35	−0.69	22.18	−0.71	21.99
2002-09-28	1367	-	-	-	21.86	0.70	22.20
2002-12-11	1441	22.28	−0.54	0.12	22.20	0.12	22.24
2004-05-12	1959	21.58	0.00	−0.14	22.15	−0.14	22.12
2004-10-31	2131	23.72	−1.67	0.02	22.35	0.03	22.35
2004-11-02	2133	-	-	-	-	-	22.32
2005-10-30	2495	21.81	−0.09	0.07	22.40	0.00	22.36
2006-09-20	2820	21.79	0.00	0.09	22.37	0.09	22.39
2006-09-21	2821	23.29	−1.23	0.08	22.38	0.08	22.40
2006-09-22	2822	22.26	−0.38	0.10	22.41	0.09	22.43
2006-09-23	2823	21.76	0.00	0.10	22.38	0.10	22.41
2006-09-25	2825	22.26	−0.36	0.08	22.41	0.07	22.42
2006-09-27	2827	22.32	−0.43	0.10	22.38	0.10	22.41
2006-09-28	2828	22.35	−0.47	0.08	22.37	0.08	22.38
2006-09-29	2829	21.79	0.00	0.09	22.39	0.07	22.40

Table A.2: Photometric solutions in the U_{38}-band. The colour term corresponds to the Johnson-Cousins colour $U-B$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = -0.73 and the default value for the colour term in the one-parameter fit is CT1 = -0.01.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2000-03-30	455	21.24	0.00	0.03	22.09	0.03	22.11
2000-03-31	456	21.21	0.00	0.05	22.07	0.05	22.10
2000-04-01	457	25.19	-3.37	0.05	22.09	0.05	22.12
2000-04-05	461	21.17	0.00	0.04	22.03	0.04	22.05
2000-07-29	576	-	-	-	23.76	16.57	22.15
2000-08-01	579	-	-	-	23.09	9.37	22.18
2000-08-26	604	21.34	-0.06	0.09	22.17	0.08	22.24
2000-08-27	605	20.97	0.00	0.55	21.84	0.55	22.17

Table A.3: Photometric solutions in the B_{123}-band. The colour term corresponds to the Johnson-Cousins colour $B-V$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = −0.22 and the default value for the colour term in the one-parameter fit is CT1 = 0.25.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2006-09-12	2812	24.66	0.00	0.21	24.91	0.21	24.83
2006-09-19	2819	24.94	−0.25	0.21	24.88	0.22	24.86
2006-09-20	2820	-	-	-	24.89	0.22	24.83
2006-09-29	2829	24.94	−0.22	0.21	24.94	0.21	24.90

Table A.4: Photometric solutions in the B-band. The colour term corresponds to the Johnson-Cousins colour $B-V$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = -0.22 and the default value for the colour term in the one-parameter fit is CT1 = 0.25.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
1999-12-02	336	-	-	-	24.74	0.27	24.76
2000-03-29	454	25.61	−0.99	0.27	24.59	0.26	24.60
2000-03-30	455	24.72	−0.29	0.23	24.63	0.23	24.61
2000-03-31	456	24.83	−0.38	0.23	24.63	0.23	24.62
2000-04-01	457	24.40	0.00	0.21	24.66	0.21	24.63
2000-04-06	462	24.43	−0.00	0.20	24.69	0.20	24.65
2000-07-02	549	26.29	−1.53	0.17	24.77	0.18	24.69
2000-11-27	697	24.42	0.00	0.22	24.70	0.22	24.68
2000-11-28	698	24.71	−0.23	0.22	24.69	0.22	24.67
2001-02-01	763	24.66	−0.23	0.23	24.65	0.23	24.63
2001-02-02	764	24.45	−0.16	0.23	24.53	0.23	24.51
2001-02-25	787	24.66	−0.27	0.23	24.58	0.23	24.57
2001-02-26	788	24.51	−0.14	0.22	24.61	0.22	24.59
2001-11-13	1048	24.49	0.00	0.21	24.75	0.21	24.71
2001-11-14	1049	24.48	0.00	0.22	24.74	0.22	24.72
2001-11-15	1050	24.69	−0.18	0.23	24.74	0.23	24.72
2001-12-12	1077	24.80	−0.25	0.22	24.76	0.22	24.74
2002-02-01	1128	22.54	−0.00	1.29	22.80	1.29	23.80
2002-02-02	1129	24.72	−0.25	0.26	24.67	0.27	24.69
2002-02-03	1130	24.70	−0.21	0.22	24.71	0.22	24.69
2002-02-04	1131	24.63	−0.18	0.21	24.68	0.21	24.65
2002-02-05	1132	25.01	−0.50	0.20	24.66	0.21	24.63
2002-06-07	1254	24.85	−0.30	0.19	24.75	0.19	24.69
2002-10-12	1381	24.56	−0.14	0.25	24.66	0.25	24.66
2004-01-31	1857	24.41	0.00	0.22	24.67	0.22	24.64
2004-02-01	1858	24.39	0.00	0.23	24.65	0.23	24.63
2004-03-19	1905	-	-	-	24.61	0.29	24.63

Table A.5: Photometric solutions in the V-band. The colour term corresponds to the Johnson-Cousins colour $V - R$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = −0.11 and the default value for the colour term in the one-parameter fit is CT1 = −0.13.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
1999-11-08	312	-	-	-	24.37	−0.16	24.35
1999-12-02	336	24.49	−0.33	−0.12	24.27	−0.12	24.27
1999-12-03	337	24.34	−0.17	−0.18	24.27	−0.18	24.24
1999-12-04	338	24.23	−0.07	−0.17	24.27	−0.17	24.25
2000-03-29	454	24.24	−0.13	−0.17	24.22	−0.17	24.20
2000-03-30	455	24.40	−0.30	−0.12	24.17	−0.11	24.18
2000-03-31	456	24.51	−0.42	−0.12	24.13	−0.11	24.14
2000-04-01	457	23.89	0.00	−0.07	24.02	−0.07	24.05
2000-04-06	462	24.09	0.00	−0.13	24.22	−0.14	24.22
2000-07-03	550	22.81	0.00	−0.93	22.94	−0.93	22.12
2000-11-28	698	24.09	0.00	−0.16	24.23	−0.16	24.21
2000-11-29	699	24.30	−0.17	−0.16	24.22	−0.16	24.20
2001-02-02	764	24.08	−0.10	−0.15	24.10	−0.16	24.09
2001-02-20	782	24.39	−0.34	−0.15	24.08	−0.13	24.08
2001-02-23	785	24.07	0.00	−0.16	24.21	−0.16	24.19
2001-02-25	787	24.25	−0.19	−0.16	24.14	−0.16	24.13
2001-02-26	788	24.15	−0.09	−0.15	24.17	−0.15	24.16
2001-03-26	816	24.53	−0.45	−0.18	23.99	−0.04	24.04
2001-03-27	817	24.04	0.00	−0.17	24.17	−0.17	24.15
2001-06-25	907	24.12	0.00	−0.15	24.26	−0.16	24.24
2001-06-27	909	24.42	−0.22	−0.18	24.25	−0.15	24.23
2001-06-29	911	24.28	−0.11	−0.16	24.28	−0.16	24.25
2001-07-26	938	24.36	−0.13	−0.19	24.34	−0.19	24.27
2001-08-21	964	24.31	−0.13	−0.14	24.28	−0.14	24.27
2001-08-22	965	24.30	−0.15	−0.13	24.25	−0.14	24.24
2001-08-23	966	25.26	−0.94	−0.17	24.25	−0.14	24.24
2001-11-12	1047	24.07	−0.00	−0.23	24.20	−0.23	24.15
2001-11-19	1054	24.24	−0.05	−0.20	24.31	−0.20	24.28
2001-12-08	1073	24.27	−0.10	−0.22	24.29	−0.22	24.25
2001-12-13	1078	24.35	−0.16	−0.16	24.28	−0.15	24.26
2002-02-09	1136	23.94	0.00	−0.17	24.07	−0.17	24.05
2002-03-09	1164	24.08	−0.09	−0.14	24.10	−0.14	24.09

Table A.5: Photometric solutions in the *V*-band.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2002-10-06	1375	24.56	−0.38	−0.15	24.22	−0.15	24.21
2002-10-07	1376	24.34	−0.21	−0.13	24.22	−0.13	24.22
2002-10-09	1378	24.06	0.00	−0.15	24.19	−0.15	24.18
2002-10-10	1379	24.30	−0.18	−0.16	24.20	−0.15	24.19
2002-10-12	1381	24.21	−0.13	−0.16	24.19	−0.16	24.17
2004-01-11	1837	24.18	−0.12	−0.15	24.17	−0.15	24.15
2004-01-15	1841	24.19	−0.15	−0.15	24.14	−0.15	24.13
2004-01-21	1847	24.15	−0.11	−0.15	24.15	−0.15	24.14
2004-10-02	2102	24.06	0.00	−0.11	24.19	−0.11	24.20
2004-10-09	2107	-	-	-	23.92	−0.26	23.86
2005-09-30	2465	24.24	−0.10	−0.17	24.25	−0.17	24.21
2005-10-30	2495	24.36	−0.20	−0.20	24.23	−0.18	24.20
2006-07-01	2739	-	-	-	24.24	−0.16	24.22

Table A.6: Photometric solutions in the R-band. The colour term corresponds to the Johnson-Cousins colour $V - R$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = −0.07 and the default value for the colour term in the one-parameter fit is CT1 = 0.00.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
1999-12-04	338	24.47	0.00	−0.06	24.55	−0.06	24.51
2000-03-29	454	24.46	0.00	−0.09	24.55	−0.09	24.50
2000-03-30	455	24.39	0.00	−0.05	24.47	−0.05	24.44
2000-03-31	456	24.32	0.00	−0.03	24.40	−0.03	24.39
2000-04-05	461	24.49	−0.06	−0.02	24.50	−0.02	24.49
2000-04-06	462	24.44	0.00	−0.02	24.52	−0.02	24.51
2000-07-27	574	24.54	0.00	−0.12	24.63	−0.12	24.55
2000-08-25	603	25.04	−0.38	−0.15	24.59	−0.09	24.49
2000-08-26	604	24.88	−0.27	−0.15	24.63	−0.14	24.47
2000-08-27	605	25.04	−0.38	−0.13	24.66	−0.12	24.50
2000-12-25	725	24.46	0.00	−0.01	24.54	−0.01	24.53
2000-12-26	726	24.54	−0.08	−0.02	24.53	−0.01	24.52
2001-02-20	782	24.45	−0.09	−0.02	24.42	−0.02	24.41
2001-02-22	784	24.36	0.00	−0.04	24.44	−0.04	24.42
2001-02-23	785	24.42	0.00	−0.03	24.51	−0.04	24.49
2001-03-26	816	24.67	−0.28	−0.03	24.28	0.06	24.31
2001-03-27	817	24.42	0.00	−0.07	24.51	−0.07	24.47
2001-06-20	902	24.49	0.00	−0.06	24.60	−0.07	24.56
2001-06-21	903	24.49	0.00	−0.05	24.60	−0.05	24.57
2001-06-26	908	24.18	0.00	0.29	24.28	0.29	24.41
2001-06-27	909	24.70	−0.11	−0.12	24.64	−0.12	24.55
2001-06-29	911	24.65	−0.08	−0.12	24.64	−0.12	24.50
2001-06-30	912	24.39	0.00	−0.02	24.48	−0.02	24.47
2001-08-20	963	24.64	−0.07	−0.10	24.64	−0.10	24.54
2001-08-21	964	24.67	−0.09	−0.10	24.64	−0.10	24.54
2001-11-12	1047	24.39	−0.00	−0.01	24.47	−0.01	24.47
2001-11-16	1051	24.58	−0.06	−0.11	24.59	−0.11	24.52
2001-11-17	1052	24.62	−0.11	−0.09	24.58	−0.10	24.53
2001-11-19	1054	24.53	0.00	−0.08	24.61	−0.08	24.57
2001-12-09	1074	24.50	0.00	−0.09	24.58	−0.09	24.54
2001-12-11	1076	24.55	−0.05	−0.05	24.57	−0.05	24.54
2003-04-05	1556	-	-	-	24.47	−0.05	24.43

Table A.6: Photometric solutions in the *R*-band.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2003-04-06	1557	24.40	0.00	−0.06	24.48	−0.06	24.44
2003-04-21	1572	24.58	−0.08	−0.12	24.56	−0.12	24.51
2004-10-04	2104	24.26	0.00	−0.06	24.44	−0.06	24.40
2004-10-09	2109	24.32	0.00	−0.08	24.40	−0.08	24.36
2004-10-10	2110	-	-	-	24.56	−0.12	24.49

Table A.7: Photometric solutions in the I-band. The colour term corresponds to the Johnson-Cousins colour $R-I$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = −0.10 and the default value for the colour term in the one-parameter fit is CT1 = 0.11.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
1999-11-04	308	23.16	0.00	0.28	23.27	0.28	23.37
1999-11-07	311	26.18	−2.61	0.27	23.31	0.27	23.40
1999-11-08	312	23.24	−0.00	0.24	23.36	0.24	23.44
2000-02-26	422	22.97	0.00	0.29	23.09	0.28	23.18
2000-03-29	454	25.13	−1.66	0.16	23.11	0.10	23.11
2000-03-30	455	23.59	−0.50	0.28	23.12	0.27	23.21
2000-03-31	456	22.87	0.00	0.32	22.99	0.32	23.11
2000-04-01	457	22.99	−0.00	0.22	23.11	0.22	23.17
2000-04-05	461	23.00	0.00	0.28	23.13	0.28	23.22
2000-04-06	462	23.19	−0.12	0.31	23.17	0.31	23.28
2000-07-04	551	23.10	−0.04	0.35	23.16	0.35	23.44
2000-07-28	575	23.04	−0.00	0.23	23.16	0.23	23.24
2000-07-31	578	23.06	0.00	0.26	23.18	0.26	23.28
2000-08-01	579	23.08	0.00	0.28	23.21	0.28	23.32
2000-08-03	581	22.81	0.00	0.26	22.92	0.26	23.02
2000-11-29	699	23.13	−0.09	0.31	23.15	0.31	23.26
2000-12-25	725	23.04	0.00	0.32	23.15	0.32	23.27
2000-12-26	726	23.04	−0.01	0.31	23.17	0.30	23.28
2001-02-21	783	23.07	−0.18	0.28	22.97	0.29	23.06
2001-03-24	814	23.09	−0.11	0.29	23.08	0.28	23.18
2001-03-25	815	22.95	0.00	0.29	23.08	0.29	23.18
2001-06-19	901	23.15	−0.09	0.29	23.17	0.29	23.28
2001-06-20	902	23.77	0.00	−1.05	23.88	−0.99	23.29
2001-06-21	903	23.18	−0.09	0.29	23.20	0.29	23.30
2001-06-22	904	23.29	0.00	−1.02	23.40	−1.01	22.77
2001-06-24	906	23.02	0.00	0.34	23.14	0.36	23.32
2001-06-25	907	23.10	−0.08	0.37	23.13	0.37	23.31
2001-06-26	908	23.04	0.00	0.27	23.17	0.29	23.27
2001-06-29	911	23.21	−0.14	0.38	23.16	0.37	23.46
2001-06-30	912	23.05	−0.08	0.43	23.08	0.43	23.39
2001-07-20	932	25.32	−1.97	0.48	22.75	0.57	23.31
2001-07-21	933	22.59	0.00	0.60	22.73	0.59	23.27

Table A.7: Photometric solutions in the I-band.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2001-07-22	934	23.08	0.00	0.33	23.20	0.34	23.45
2001-07-23	935	23.64	−0.50	0.35	23.11	0.35	23.41
2001-07-24	936	23.07	0.00	0.34	23.19	0.34	23.44
2001-07-25	937	23.07	0.00	0.36	23.21	0.36	23.52
2001-07-26	938	23.13	−0.05	0.35	23.20	0.35	23.50
2001-07-27	939	22.66	0.00	0.40	22.81	0.40	23.18
2001-08-19	962	23.08	0.00	0.25	23.21	0.25	23.29
2001-08-21	964	23.00	−0.02	0.41	23.10	0.42	23.41

Table A.8: Photometric solutions in the I_{EIS}-band. The colour term corresponds to the Johnson-Cousins colour $R - I$. The default value for the extinction coefficient in the one- and two-parameter fits is EXT1 = EXT2 = 0.00 and the default value for the colour term in the one-parameter fit is CT1 = 0.03.

night	ID	ZP3	EXT3	CT3	ZP2	CT2	ZP1
2002-03-10	1165	23.27	0.00	0.06	23.27	0.06	23.28
2002-03-11	1166	23.33	0.00	0.06	23.33	0.06	23.35
2002-03-12	1167	23.36	0.00	0.06	23.36	0.06	23.38
2002-03-13	1168	23.39	−0.02	0.05	23.37	0.06	23.38
2005-07-28	2401	25.08	−1.45	0.06	23.35	0.06	23.37
2006-07-01	2739	23.03	0.00	−0.06	23.03	−0.06	22.98
2006-08-27	2796	23.37	0.00	0.08	23.37	0.08	23.44
2006-09-09	2809	-	-		23.26	0.21	23.34
2006-09-12	2812	23.25	0.00	0.08	23.25	0.08	23.32

A.2 Colour-colour plots of stars

In Figs. A.1 to A.9 all colour-colour plots of stars in the DPS are displayed. These were used for a check of the absolute photometric calibration.

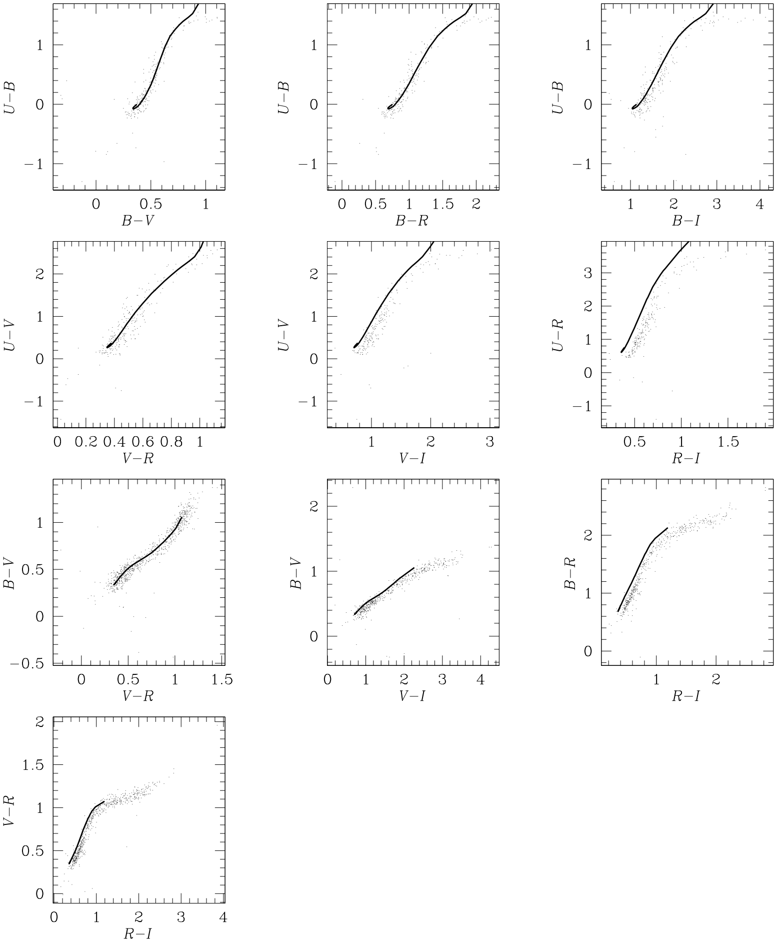

Figure A.1: Same as Fig. 3.4 but for all available colours in the field Deep1a.

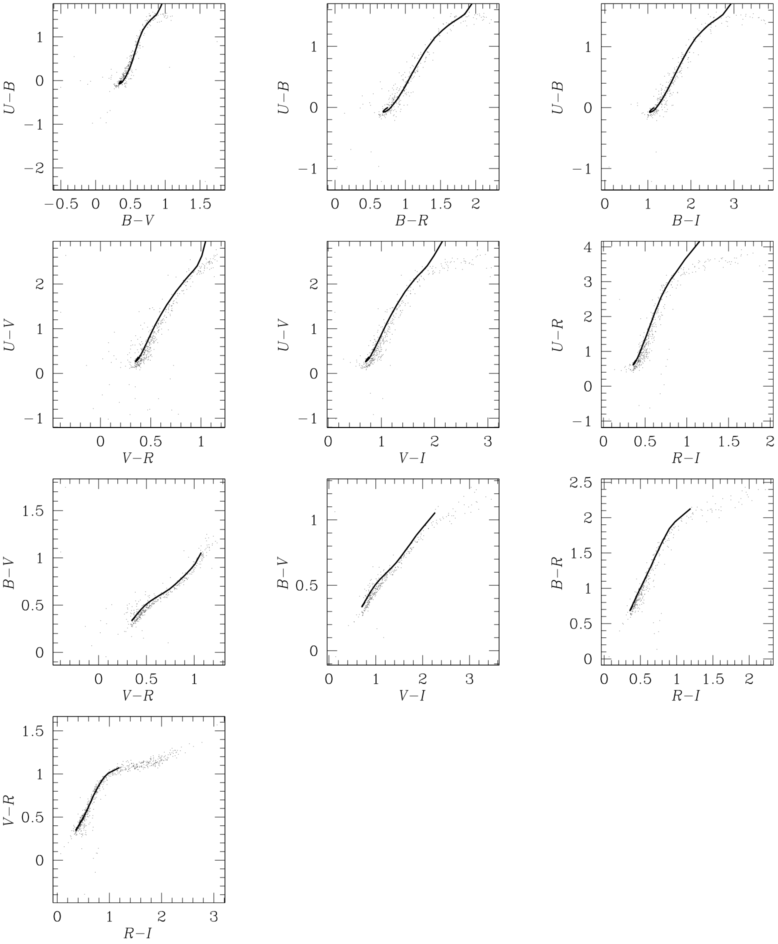

Figure A.2: Same as Fig. 3.4 but for all available colours in the field Deep1b.

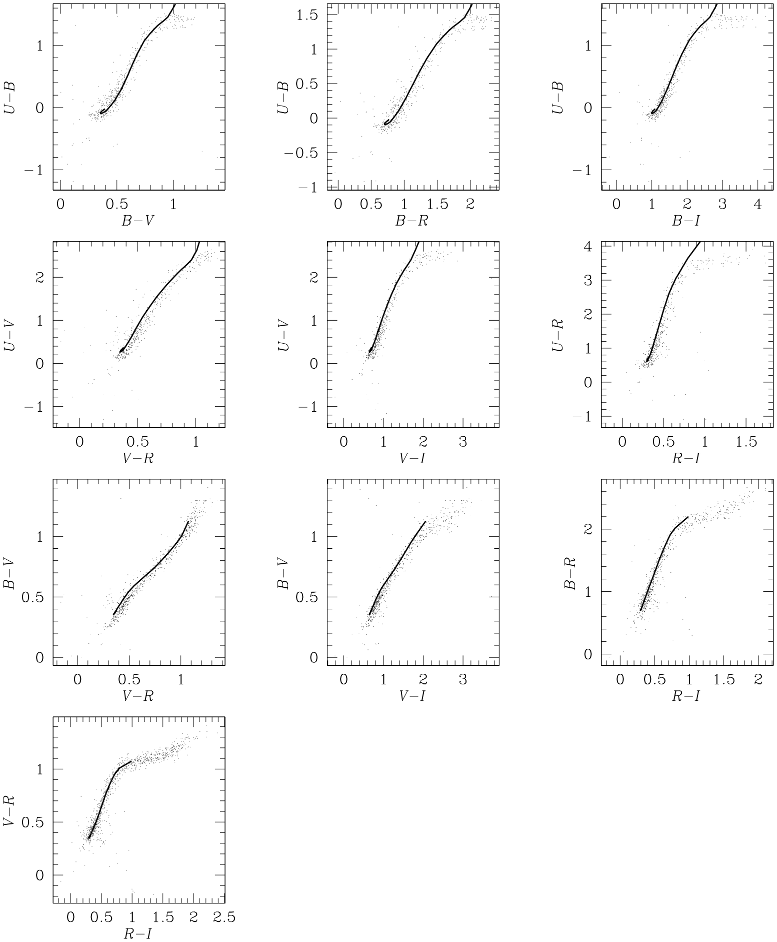

Figure A.3: Same as Fig. 3.4 but for all available colours in the field Deep1c.

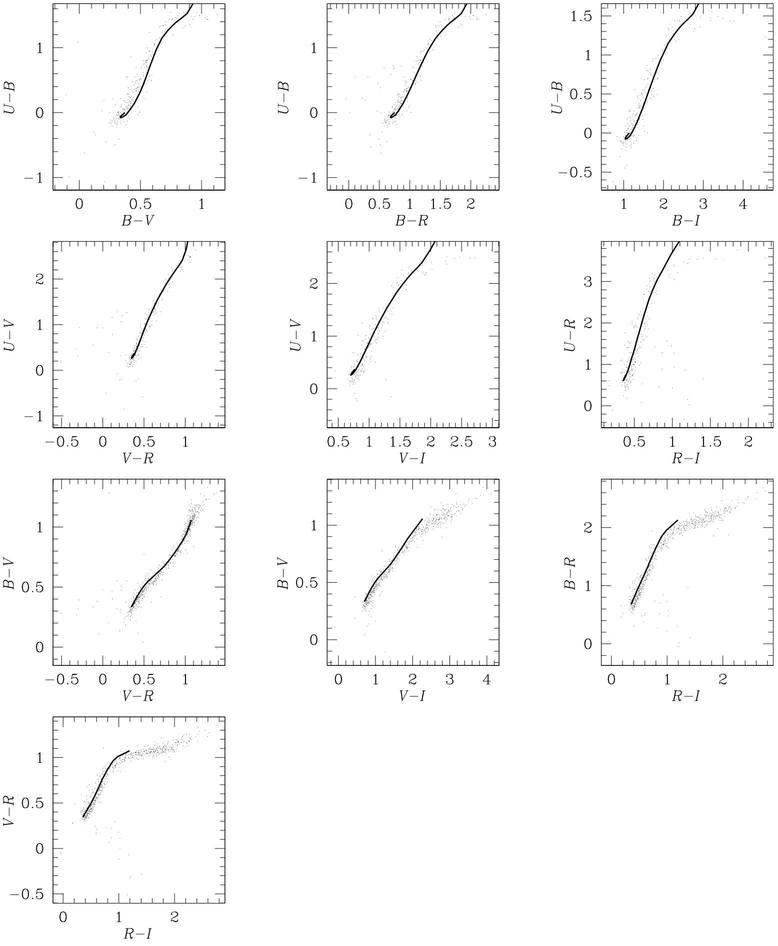

Figure A.4: Same as Fig. 3.4 but for all available colours in the field Deep2b.

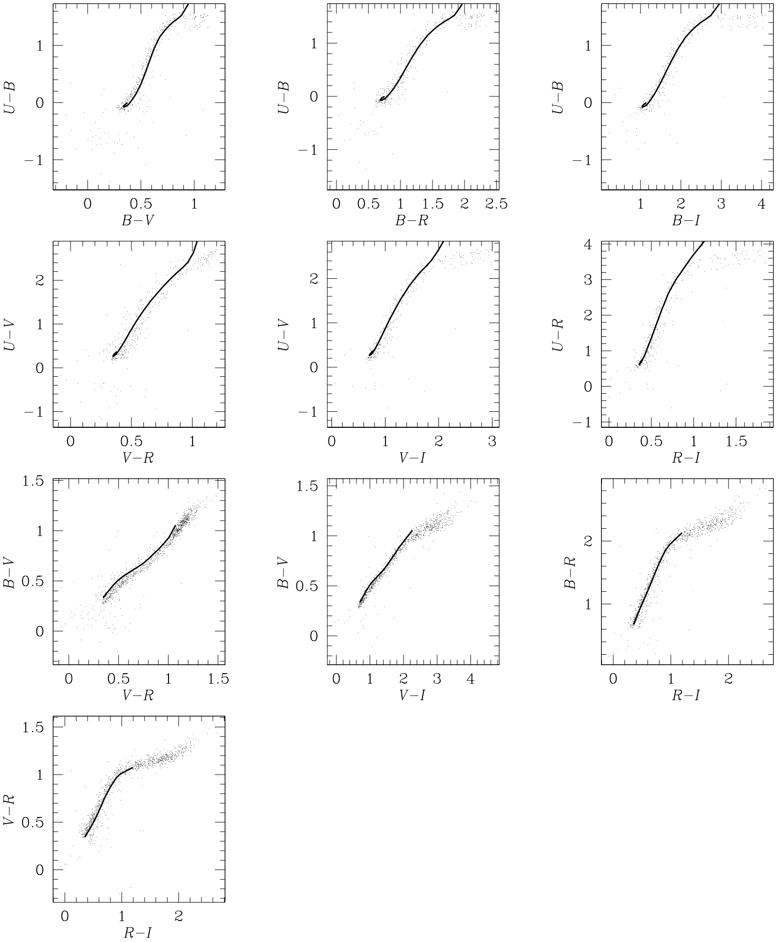

Figure A.5: Same as Fig. 3.4 but for all available colours in the field Deep2c.

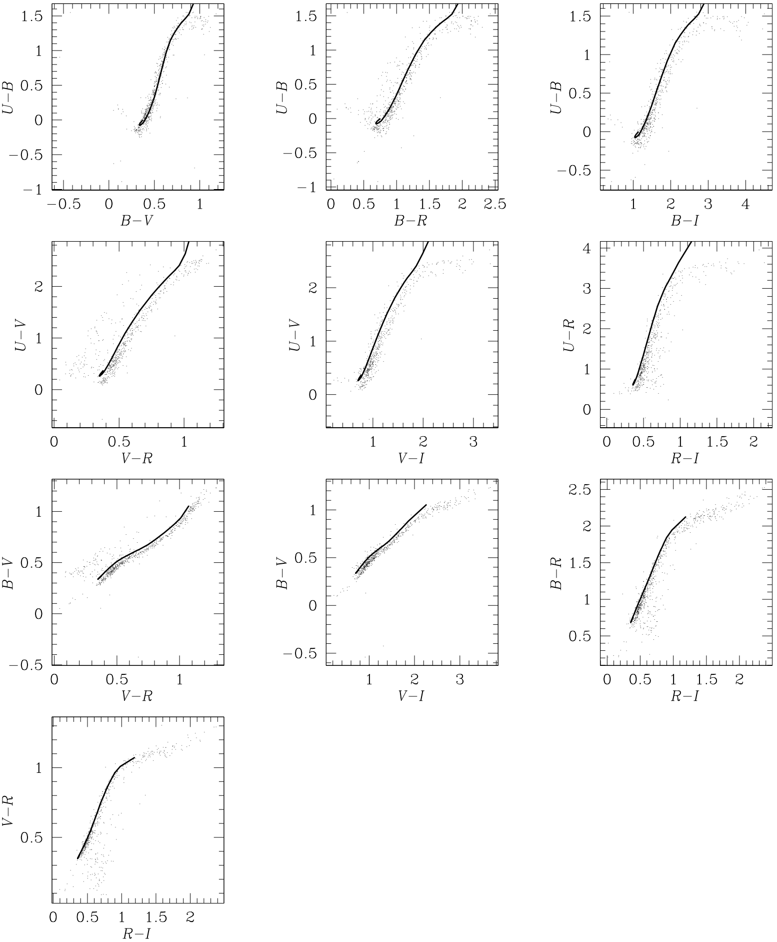

Figure A.6: Same as Fig. 3.4 but for all available colours in the field Deep3a.

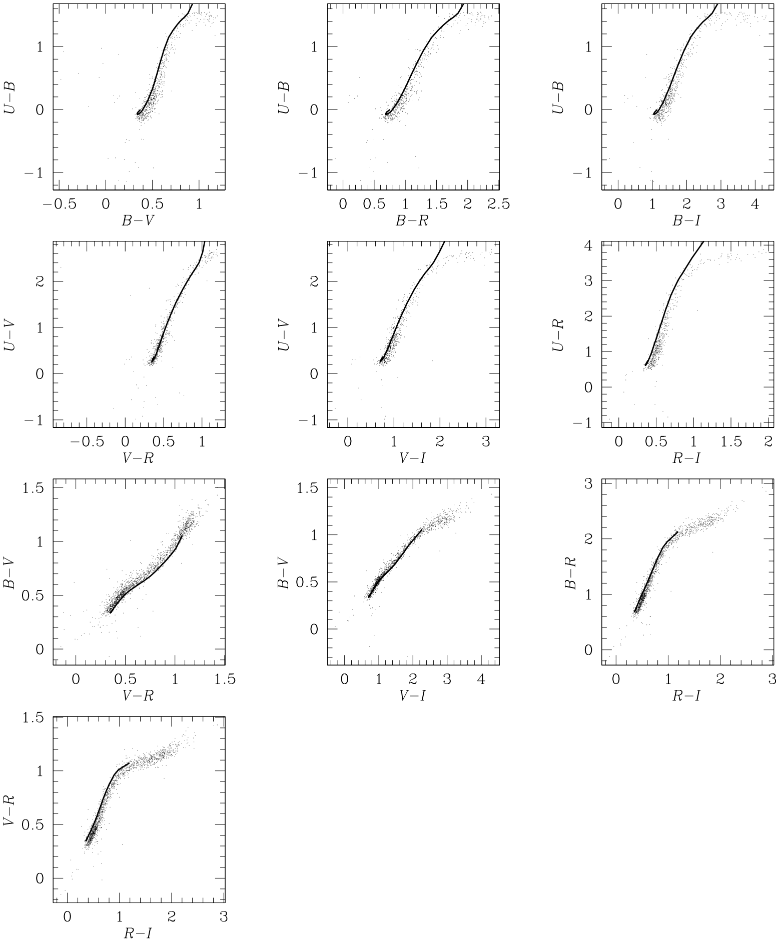

Figure A.7: Same as Fig. 3.4 but for all available colours in the field Deep3b.

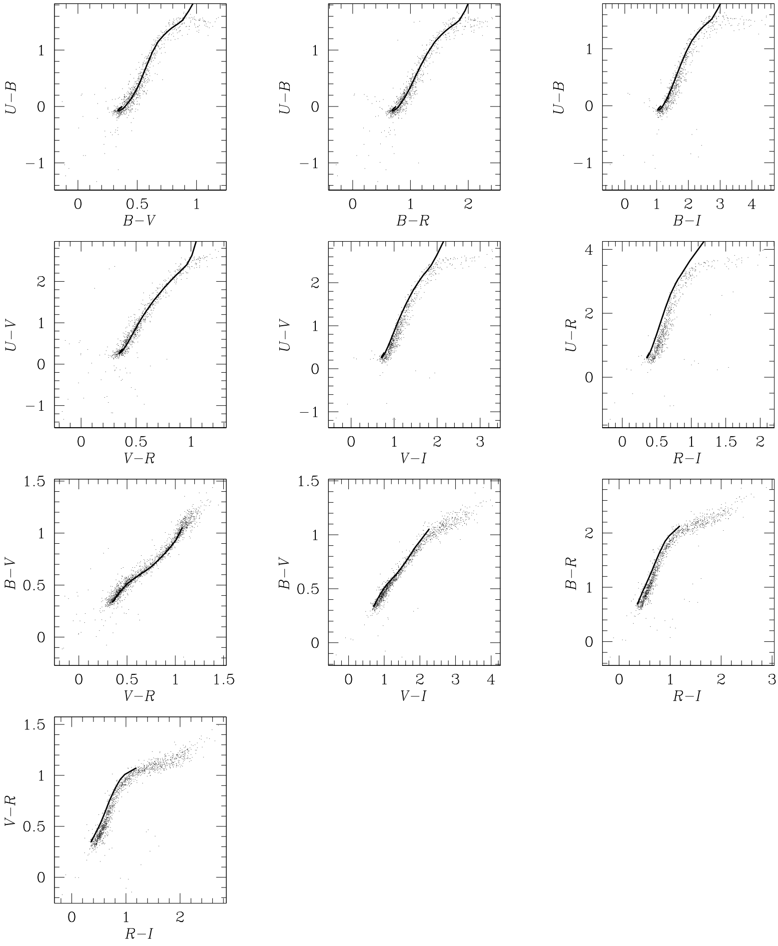

Figure A.8: Same as Fig. 3.4 but for all available colours in the field Deep3c.

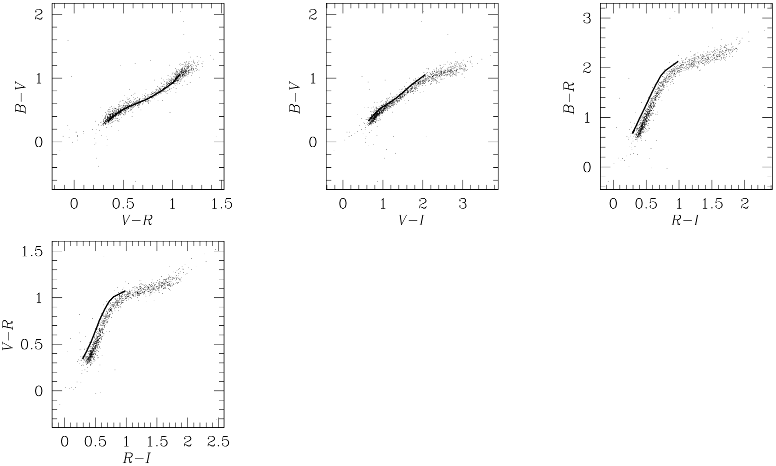

Figure A.9: Same as Fig. 3.4 but for all available colours in the field Deep3d.

B

FITS header

In the following we show an excerpt of one of our image headers to describe the special keywords inserted by the reduction pipeline[1]. EXPTIME is the sum of the exposure times of all images that entered the coaddition. GAIN is the instrumental gain (2.0) multiplied by EXPTIME. MAGZP is the Vega magnitude zeropoint which is to be used for converting counts into magnitudes by: $\mathrm{mag} = \mathrm{MAGZP} - 2.5 \cdot \log(\mathrm{counts})$. The COND keywords contain the filter conditions that were applied to the image catalogues before coaddition. Within these conditions AUTO_ZP represents a filtering on the single images' relative zeropoints, BACKGR a filtering on the single images' background flux, and SEEING a filtering on the single images' measured seeing. The NIGHT keywords summarise the absolute photometric calibration containing pairs of GaBoDS IDs and chosen solutions, e.g. 903 2 means that in the night number 903 (2001-06-21) the two-parameter fit was chosen.

In Table B.1 the important FITS header keywords of our released images are summarised.

```
====> file Deep3c/Deep3c_I.D3CA.swarp.fits (main) <====
...
...
...
EXPTIME =                25193 / total Exposure Time
GAIN    =                50386 / effective GAIN for SExtractor
MAGZP   =              23.0005 / Vega Magnitude Zeropoint
```

---

[1] The image headers of the files released via the ESO archive at `http://archive.eso.org/archive/eso_data_products.html` are slightly different and contain a lot of additional keywords in order to comply with archive standards.

```
SEEING  =              1.02218 / measured image Seeing
COMMENT
COMMENT Conditions on the input images:
COND1   = '((SEEING<2)AND(AUTO_'
COND2   = 'ZP>0));'
DUMMY8  =                    0 / DUMMY keyword
DUMMY9  =                    0 / DUMMY keyword
DUMMY10 =                    0 / DUMMY keyword
NIGHT1  = '903 2, 906 2, 907 2,'
NIGHT2  = '908 2, 911 2,'
NIGHT3  = ', 911 0, 912 0,'
...
...
...
```

Table B.1: FITS header keywords for the released images. The astrometric keywords from RADECSYS to CDELT2 are described in Greisen & Calabretta (2002).

Keyword	Unit	comment
RADECSY		astrometric system
CTYPE1		WCS projection type for x-axis
CUNIT1		x-axis unit
CRVAL1	deg	world x-coordinate of reference pixel
CRPIX1	pix	reference pixel on x-axis
CDELT1	deg/pix	pixel step along x-axis
CTYPE2		WCS projection type for y-axis
CUNIT2		y-axis unit
CRVAL2	deg	world y-coordinate of reference pixel
CRPIX2	pix	reference pixel on y-axis
CDELT2	deg/pix	pixel step along y-axis
EXPTIME	sec	total exposure time
GAIN		effective GAIN (instrumental gain [2.0] × EXPTIME)
MAGZP	mag	Vega magnitude zeropoint
SEEING	arcsec	measured image seeing
COND1-5		condition on the input images entering coaddition
NIGHT1-3		GaBoDS IDs and solutions for nights included in phot. solution

Bibliography

Adelberger, K. L., Steidel, C. C., Giavalisco, M., et al. 1998, ApJ, 505, 18

Adelberger, K. L., Steidel, C. C., Pettini, M., et al. 2005, ApJ, 619, 697

Allen, P. D., Moustakas, L. A., Dalton, G., et al. 2005, MNRAS, 360, 1244

Arnouts, S., Vandame, B., Benoist, C., et al. 2001, A&A, 379, 740

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Barlow, R. 1989, Statistics. A guide to the use of statistical methods in the physical sciences (The Manchester Physics Series, New York: Wiley, 1989)

Benítez, N. 2000, ApJ, 536, 571

Benítez, N., Ford, H., Bouwens, R., et al. 2004, ApJS, 150, 1

Bertin, E. 2003, SWarp user's guide, `http://terapix.iap.fr`

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393

Blanton, M. R., Dalcanton, J., Eisenstein, D., et al. 2001, AJ, 121, 2358

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476

Bouché, N. & Lowenthal, J. D. 2004, ApJ, 609, 513

Bruzual , A. G. & Charlot, S. 1993, ApJ, 405, 538

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682

Coe, D., Benítez, N., Sánchez, S. F., et al. 2006, AJ, 132, 926

Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, ApJS, 43, 393
Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
Collister, A. A. & Lahav, O. 2004, PASP, 116, 345
Cooray, A. & Sheth, R. 2002, Phys. Rep., 372, 1
Csabai, I., Budavári, T., Connolly, A. J., et al. 2003, AJ, 125, 580
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Edmondson, E. M., Miller, L., & Wolf, C. 2006, MNRAS, 371, 1693
Einstein, A. 1916, Annalen der Physik, 49, 769
Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, Astronomische Nachrichten, 326, 432
Etherington, J. M. H. 1933, Phil. Mag., 15, 761
Fioc, M. & Rocca-Volmerange, B. 1997, A&A, 326, 950
Foucaud, S., McCracken, H. J., Le Fèvre, O., et al. 2003, A&A, 409, 835
Franx, M., Labbé, I., Rudnick, G., et al. 2003, ApJ, 587, L79
Fruchter, A. S. & Hook, R. N. 2002, PASP, 114, 144
Gabasch, A. 2004, PhD Thesis, `http://www.usm.uni-muenchen.de/people/gabasch/phdthes.html`
Gabasch, A., Bender, R., Seitz, S., et al. 2004, A&A, 421, 41
Gawiser, E., van Dokkum, P. G., Herrera, D., et al. 2006, ApJS, 162, 1
Giavalisco, M. & Dickinson, M. 2001, ApJ, 550, 177
Giavalisco, M., Steidel, C. C., Adelberger, K. L., et al. 1998, ApJ, 503, 543
Girardi, L., Bertelli, G., Bressan, A., et al. 2002, A&A, 391, 195
Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895
Greisen, E. W. & Calabretta, M. R. 2002, A&A, 395, 1061
Hamana, T., Ouchi, M., Shimasaku, K., Kayo, I., & Suto, Y. 2004, MNRAS, 347, 813
Harrison, E. R. 1970, Phys. Rev. D, 1, 2726
Heidt, J., Appenzeller, I., Gabasch, A., et al. 2003, A&A, 398, 49
Hildebrandt, H., Bomans, D. J., Erben, T., et al. 2005, A&A, 441, 905
Hildebrandt, H., Erben, T., Dietrich, J. P., et al. 2006, A&A, 452, 1121

Hildebrandt, H., Erben, T., Dietrich, J. P., Schirmer, M., & Schneider, P. 2007a, in ESO Calibration Workshop 2007

Hildebrandt, H., Pielorz, J., Erben, T., et al. 2007b, A&A, 462, 865

Hildebrandt, H., Wolf, C., & Bénitez, N. 2007c, submitted to A&A

Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841

Kashikawa, N., Yoshida, M., Shimasaku, K., et al. 2006, ApJ, 637, 631

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 303, 188

Kinney, A. L., Calzetti, D., Bohlin, R. C., et al. 1996, ApJ, 467, 38

Koch, A., Odenkirchen, M., Grebel, E. K., & Caldwell, J. A. R. 2004, Astronomische Nachrichten, 325, 299

Koo, D. C. 1999, in ASP Conf. Ser. 191, ed. R. Weymann, L. Storrie-Lombardi, M. Sawicki, & R. Brunner, 3

Landolt, A. U. 1992, AJ, 104, 340

Landy, S. D. & Szalay, A. S. 1993, ApJ, 412, 64

Le Fèvre, O., Vettolani, G., Paltani, S., et al. 2004, A&A, 428, 1043

Lee, K.-S., Giavalisco, M., Gnedin, O. Y., et al. 2006, ApJ, 642, 63

Limber, D. N. 1953, ApJ, 117, 134

Madau, P. 1995, ApJ, 441, 18

Moles, M., Alfaro, E., Benítez, N., et al. 2005, astro-ph/0504545

Monet, D. B. A., Canzian, B., Dahn, C., et al. 1998, VizieR Online Data Catalog, 1252, 0

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996, ApJ, 462, 563

Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493

Noll, S., Mehlert, D., Appenzeller, I., et al. 2004, A&A, 418, 885

Ouchi, M., Hamana, T., Shimasaku, K., et al. 2005, ApJ, 635, L117

Ouchi, M., Shimasaku, K., Okamura, S., et al. 2001, ApJ, 558, L83

Ouchi, M., Shimasaku, K., Okamura, S., et al. 2004, ApJ, 611, 685

Peacock, J. A. 1999, Cosmological Physics (Cambridge University Press)

Peacock, J. A. & Dodds, S. J. 1996, MNRAS, 280, L19

Peebles, P. J. E. 1980, The large-scale structure of the universe (Princeton University Press)

Porciani, C. & Giavalisco, M. 2002, ApJ, 565, 24

Prandoni, I., Gregorini, L., Parma, P., et al. 2000, A&AS, 146, 31

Press, W. H. & Schechter, P. 1974, ApJ, 187, 425

Radovich, M. 2002, ASTROMETRIX, `http://terapix.iap.fr`

Robertson, H. P. 1935, ApJ, 82, 284

Schechter, P. 1976, ApJ, 203, 297

Scranton, R., Ménard, B., Richards, G. T., et al. 2005, ApJ, 633, 589

Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119

Simon, P. 2006, astro-ph/0609165

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, MNRAS, 341, 1311

Spergel, D. N., Bean, R., Dore', O., et al. 2006, astro-ph/0603449

Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, ApJ, 492, 428

Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, ApJ, 519, 1

Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996, ApJ, 462, L17

Steidel, C. C. & Hamilton, D. 1993, AJ, 105, 2017

Stetson, P. B. 2000, PASP, 112, 925

van Dokkum, P. G., Förster Schreiber, N. M., Franx, M., et al. 2003, ApJ, 587, L83

Walker, A. G. 1936, Proc. Lond. Math. Soc.(2), 42, 90

Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, A&A, 421, 913

Wolf, C., Meisenheimer, K., & Röser, H.-J. 2001, A&A, 365, 660

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579

Zacharias, N., Urban, S. E., Zacharias, M. I., et al. 2004, AJ, 127, 3043

Zeldovich, Y. B. 1972, MNRAS, 160, 1P

Acknowledgement

Many thanks to:

- Peter Schneider for being the supervisor of my thesis,
- Thomas Erben for being the co-supervisor,
- Bodo Ziegler for being the second referee,
- the two other members of the exam committee for their time,
- Jörg Dietrich, Christian Wolf, Ludovic van Waerbeke, Patrick Simon, Jasmin Pielorz, Tim Eifler, Narciso Benítez,
- Mischa Schirmer, Marco Hetterscheidt, Tim Schrabback-Krahe, Marusa Bradac, Philippe Heraudeau, Martha Milkeraitis, Kathy Schrüfer, Oliver Cordes, Oliver Czoske,
- Leo Girardi, Remco Slijkhuis, Clemens Trachternach, Olaf Schmithüsen, Lutz Haberzettl,
- all current and former members of the lensing group and of the Argelander-Institut für Astronomie,
- my friends,
- my family.

This work was supported by the German Ministry for Education and Science (BMBF) through the DLR under the project 50 OR 0106, by the BMBF through DESY under the project 05 AV5PDA/3, and by the Deutsche Forschungsgemeinschaft (DFG) under the projects SCHN342/3-1 and ER327/2-1.