Photometric Redshifts and Properties of Galaxies from the Sloan Digital Sky Survey

Natascha Greisel

Photometric Redshifts and Properties of Galaxies from the Sloan Digital Sky Survey

Natascha Greisel

Dissertation der Fakultät für Physik
Dissertation of the Faculty of Physics

der Ludwig-Maximilians-Universität München
at the Ludwig Maximilian University of Munich

für den Grad des
for the degree of

Doctor rerum naturalium

vorgelegt von Natascha Greisel
presented by

aus München, Deutschland (Germany)
from

Munich, 18 December 2014

1^{st} Evaluator: Prof. Dr. Ralf Bender

2^{nd} Evaluator: Prof. Dr. Jochen Weller

Date of the oral exam: 25 February 2015

Zusammenfassung

Die Bestimmung kosmologischer Rotverschiebungen ist für viele Untersuchungen in der Kosmologie und extragalaktischen Astronomie essentiell. Dies sind z.B. Analysen der großskaligen Struktur des Universums, des Gravitationslinseneffekts oder der Galaxienentwicklung. In der Kosmologie statistisch aussagekräftige Volumina werden heutzutage meist durch Breitbandfilter beobachtet. Mit diesen photometrischen Beobachtungen kann man nur Aussagen über die grobe Form des Spektrums machen, weshalb man sich statistischer Mittel bedienen muss um die Rotverschiebung zu schätzen. Eine Methode zur Bestimmung photometrischer Rotverschiebungen ist es, Flüsse von Modellgalaxien in den Filtern bei verschiedenen Rotverschiebungen z vorherzusagen und mit den beobachteten Flüssen zu vergleichen (*template fitting*). Danach wird eine Likelihood-Analyse durchgeführt, in der die Vorhersagen mit den Beobachtungen verglichen werden, um die Wahrscheinlichkeitsdichte $P(z)$ und die wahrscheinlichste Rotverschiebung zu bestimmen. Um mit template fitting Methoden möglichst genaue Ergebnisse zu erzielen, müssen die Modellspektren, sowie die zugehörigen a priori Wahrscheinlichkeiten sorgfältig ausgewählt werden.

In dieser Arbeit nutze ich photometrische und spektroskopische Daten leuchtkräftiger roter Galaxien (LRGs) aus dem *Sloan Digital Sky Survey* (SDSS). Ich untersuche die Genauigkeit photometrischer Rotverschiebungen, die mit Modellspektren erreicht werden, welche von mir speziell für den Satz LRGs bei $z \lesssim 0.5$ (SDSS-II) entwickelt wurden, und vergleiche sie mit publizierten Ergebnissen. Diese Modelle wurden ohne Informationen aus Wellenlängenbereichen, die kurzwelliger als das SDSS u Band sind, erstellt. Die Galaxieneigenschaften, die wir aus den Modellen für den UV Bereich vorhersagen können decken sich allerdings mit denen aus anderen Beobachtungen. Darüber hinaus analysiere ich die sich daraus ergebenden Eigenschaften der am besten fittenden Modellspektren und vergleiche sie mit den spektroskopischen Daten. Aus den Ergebnissen ist abzuleiten, dass leuchtschwächere rote Galaxien bei niedriger Rotverschiebung im Mittel größere Anzeichen von Sternentstehung zeigen als leuchtkräftige, was durch Analysen der Spektren bestätigt wird. Überdies können wir einen Abfall im UV Fluss von höheren zu niedrigeren Rotverschiebungen hin beobachten, welcher durch die Alterung der Galaxienpopulation erzeugt wird.
Desweiteren generiere ich Modellspektren für leuchtkräftige rote Galaxien aus SDSS-III bei höheren Rotverschiebungen $0.45 \leq z \leq 0.9$. Ich modifiziere hierzu die Form theoretischer spektraler Energieverteilungen um die Farben der untersuchten Galaxien mit den Modellen bestmöglich wiedergeben zu können. Ich reduziere die Dimension des Raums, der durch die Farben und absoluten Helligkeiten aufgespannt wird, auf zwei Dimensionen mit einer selbstorganisierenden Karte. Diese wird mit einem k-means Algorithmus partitioniert indem wir Häufungspunkte der Daten identifizieren. Aus den sich ergebenden Partitionen selektiere ich einzelne Modellspektren, die die zugrundeliegenden Galaxien repräsentieren. Eine Auswahl aus den erstellten Modellen wird danach für die Schätzung photometrischer Rotverschiebungen verwendet, deren Genauigkeit über die von SDSS publizierten Ergebnisse hinausgeht.

Abstract

The determination of photometric redshifts is essential for many subjects in cosmology and extragalactic astronomy, like the large scale structure of the Universe, gravitational lensing, or galaxy evolution. If the spectral energy distribution (SED) of a galaxy is measured with high enough spectral resolution, the redshift can be easily derived through the absorption and emission lines which are created by the elements in the galaxy. However, currently more telescopes are equipped with large cameras with charged coupled devices (CCDs) that observe the sky through optical filters. With these photometric observations it is possible to detect much fainter astronomical objects than with spectroscopy. Furthermore, photometric observations are less time consuming and cheaper in comparison, wherefore they are preferentially used for observations of statistical meaningful cosmological volumes. Nonetheless, photometric data, which are often gained by observations through broadband filters, are not as precisely resolved as spectra. Therefore one does not have information about the accurate position in wavelength of spectral lines, but only about the overall shape of the SED. This is the reason why so-called photometric redshifts have to be derived by statistical means.
One approach to estimate the redshift through photometry alone are template fitting methods which compare the fluxes predicted by model spectra with the observations. After that, a likelihood analysis is performed with which a probability density function $P(z)$ and the most probable value of z can be derived. To achieve high accuracies with photometric redshift template fitting techniques, the model spectra as well as their corresponding prior probabilities have to be chosen carefully.

In this work I use photometric and spectroscopic data of luminous red galaxies from the *Sloan Digital Sky Survey* (SDSS). I analyze the precision of photometric redshifts estimated with model SEDs specifically designed to match the set of luminous red galaxies of SDSS-II at redshifts $z \lesssim 0.5$ in color and I compare them with published results. These models were created without information on their properties at wavelengths shorter than the SDSS u band. However, the galaxy UV characteristics derived from the model SEDs match those of other observations. Furthermore, I investigate the SED properties derived from the best fitting models with respect to spectroscopic data as functions of redshift and luminosity. At lower redshifts less luminous galaxies from our sample on average show increased signs of star formation in comparison to galaxies with higher luminosities. This is supported by analyses of the line strengths in the spectra. Moreover, star formation activity increases with increasing redshift which is caused by the aging of the galaxy population from higher to lower redshifts. I also generate model spectra for red galaxies from the SDSS-III located at even higher redshifts $0.45 \leq z \leq 0.9$. For this I modify the shape of theoretical spectra to match the data of the analyzed galaxies to a better extent. The multidimensional space defined by the colors and the absolute magnitude of the galaxies is reduced to two dimensions through a self-organizing map. The map is then partitioned by a k-means algorithm which identifies clusters in the data. From the cluster cells I select model spectra which represent the galaxies from within the same cell. A selection of the models is then used as a template set for photometric redshift estimation. I find that our models improve the redshift accuracy in comparison to the results published by SDSS.

Contents

Zusammenfassung **vii**

Abstract **ix**

Contents **xii**

List of Figures **xv**

List of Tables **xvii**

1 Preface **1**

2 Photometric Redshifts **7**
2.1 Introduction to Astronomical Observations 7
2.2 Photometric Redshifts 11
2.2.1 Empirical Methods 11
2.2.2 Template Fitting 13

3 SDSS Data **21**
3.1 The SDSS-II LRG Sample 24
3.2 The BOSS LOWZ and CMASS Samples 27

4 Model SEDs **31**
4.1 Galaxies 31
4.1.1 Elliptical Galaxies 32
4.2 Model Spectral Energy Distributions 37
4.2.1 Galaxy Stellar Population Evolution 37
4.2.2 Model SEDs and LRG Data 40

5 Novel SED Templates for SDSS-II LRGs **51**
5.1 SEDs for LRGs by SED Fitting 51
5.2 Selection of Novel LRG Templates 53
5.3 Photo-z Precision with Novel LRG Templates 55
5.4 Comparison to SDSS Database Results 57
5.5 Properties of the Novel Template Set 62

5.5.1 UV colors of the Novel Templates and M09 Models 62
5.5.2 Differences in the SEDs within $z - M_R$ Bins 64
5.6 Summary 70

6 Templates and Photo-zs for CMASS Galaxies **71**
6.1 The BOSS CMASS Sample 72
6.1.1 Colors of M09 Models and BOSS CMASS Galaxies 72
6.2 New SED Templates 73
6.2.1 Generating Model SEDs for CMASS Galaxies by SED Fitting 75
6.2.2 Selection of Best Fitting SEDs for the New Template Set 83
6.3 Photometric Redshifts 90
6.3.1 Photometric Redshifts with the Novel Template SEDs 90
6.3.2 Comparison to SDSS Photometric Redshifts 98
6.3.3 Deviations in Color Predictions 103
6.4 Summary 105

7 Photo-zs from DES-SV Data **107**

8 Physical Properties of a Lensed High-z System **117**

9 Summary and Conclusion **127**

Bibliography **143**

A Photo-z Quality Metrics **145**

B SED Fitting the SDSS-II LRG Data **149**
B.1 SED Fitting Results with Different Model Sets 149
B.2 Color-Color Relations of the New Models 150

C SED Fitting Results for the BOSS CMASS Sample **161**

D SED Fitting Results for the Lensed High-z System **163**

E SDSS `CasJobs` Queries **173**
E.1 SDSS-II Spectroscopic LRG Sample 173
E.2 SDSS-III BOSS CMASS Galaxy Sample 175

F Creation of SSP and CSP Models **179**
F.1 Generation of BC03 Models 179
F.2 Generation of Maraston CSPs 181

Acknowledgments **183**

List of Figures

1.1 History of the Universe 2
1.2 Hierarchical Structure Formation 3

2.1 Example for Spectroscopic and Photometric Data 9
2.2 Prior Probabilities 17
2.3 `PhotoZ` Output Example 18

3.1 SDSS Filter Curves 22
3.2 Spectrum of Object 7304 23
3.3 SDSS DR7 Footprint 24
3.4 SDSS DR10 Footprint 25
3.5 LRG Selection Cuts 26
3.6 z versus Colors of SDSS-II LRG Sample 28
3.7 BOSS Spectroscopic Redshift Distribution 29

4.1 Hubble Tuning Fork 32
4.2 Isophotes Elliptical Galaxies 34
4.3 Image of an Elliptical Galaxy 36
4.4 Hertzsprung-Russell Diagram 37
4.5 SEDs of BC03 SSP Models at Various Ages 38
4.6 SEDs of BC03 SSP Models at Various Metallicities and Ages 39
4.7 z versus Colors of SDSS-II LRG Sample and *PZstandard* Model Set 41
4.8 z versus Colors of SDSS-II LRG Sample and CWW Models 43
4.9 z versus Colors of SDSS-II LRG Sample and BC03 Models 45
4.10 z versus Colors of SDSS-II LRG Sample and `LePhare` Model Subset 47
4.11 z versus Colors of SDSS-II LRG Sample and M09 Model 49

5.1 Example Output of `SEDfit` 52
5.2 Color-Color Plot of LRGs and SED Fitting Results 54
5.3 Redshift versus Colors of LRG Models 56
5.4 Photo-z Results with Novel LRG Templates 58
5.5 Photo-z Results of SDSS Template Fitting 59
5.6 Photo-z Results with SDSS ANN(D1) 60
5.7 Photo-z Results with SDSS ANN(CC2) 61

5.8 UV Colors of Novel LRG Models at $0.05 \leq z_{\rm spec} \leq 0.12$ . . . 63
5.9 z versus M_R of LRG Models . . . 65
5.10 Superposition of LRG SEDs for $0.0 \leq z_{\rm spec} \leq 0.1$ . . . 66
5.11 Superposition of LRG SEDs for $-24.5 \leq M_R \leq -22.7$ . . . 67
5.12 z versus M_R of Star Forming LRGs . . . 68
5.13 Fraction of Star Forming LRGs as a Function of Luminosity . . . 68
5.14 Stacked Spectra of LRGs from $z < 0.1$ . . . 69

6.1 CMASS Spec-z Distribution and M_r . . . 73
6.2 z versus Colors of CMASS Sample and M09 Model . . . 74
6.3 Color vs. Color for CMASS Galaxies at $z \sim 0.5$ and $z \sim 0.5$ and BC03 . . . 77
6.4 Color vs. Color for CMASS Galaxies at $z \sim 0.7$ and $z \sim 0.8$ and BC03 . . . 78
6.5 SED of G13 and Modifications by $\lambda^{-\beta}$ . . . 79
6.6 Color vs. Color of CMASS Data and Modified Model SEDs . . . 80
6.7 Color vs. Color for CMASS Galaxies at $z \sim 0.5$ and $z \sim 0.6$ and BC03 $\lambda^{-\beta}$ . . . 81
6.8 Color vs. Color for CMASS Galaxies at $z \sim 0.7$ and $z \sim 0.8$ and BC03 $\lambda^{-\beta}$ . . . 82
6.9 Δcol for BC03 and BC03 $\lambda^{-\beta}$ models at $0.48 \leq z \leq 0.52$ . . . 84
6.10 Δcol for BC03 and BC03 $\lambda^{-\beta}$ models at $0.58 \leq z \leq 0.62$ . . . 85
6.11 Δcol for BC03 and BC03 $\lambda^{-\beta}$ models at $0.68 \leq z \leq 0.72$ . . . 86
6.12 Δcol for BC03 and BC03 $\lambda^{-\beta}$ models at $0.78 \leq z \leq 0.82$ . . . 87
6.13 SOM Example for $0.48 \leq z \leq 0.52$ . . . 88
6.14 k-Means Clustering of the SOM for $0.48 \leq z \leq 0.52$ . . . 89
6.15 Novel CMASS Template SEDs . . . 91
6.16 Colors of the Novel CMASS Template SEDs . . . 93
6.17 Photo-z Results with the Novel CMASS Template SEDs . . . 95
6.18 Δz_{rf} Distribution of the Novel CMASS Template SEDs . . . 96
6.19 Δmag of the Novel CMASS Template SEDs . . . 97
6.20 Photo-z Results of SDSS-KF . . . 99
6.21 Δz_{rf} Distribution of SDSS-KF . . . 100
6.22 Photo-z Results of SDSS-RF . . . 101
6.23 Δz_{rf} Distribution of SDSS-RF . . . 102
6.24 Δcol with photo-z Results from this Work and SDSS-KF . . . 104

7.1 Model SEDs used for the DES-SV . . . 110
7.2 Example for Prior Adaption . . . 111
7.3 $z_{\rm spec}$ versus $z_{\rm phot}$ for *main-main* Test on DES-SV Data . . . 113
7.4 $N(z)$ for *main-main* Test on DES-SV Data . . . 114
7.5 Δmag for *main-main* Test on DES-SV Data . . . 115

8.1 CLASH Filter Transmission Curves . . . 118
8.2 Image of RXC J2248.7-4431 . . . 119
8.3 Photo-z Results of ID1-4. . . . 120
8.4 SED Fitting Age and Mass of Mo14, R11, Z12, and B14 . . . 124
8.5 Best UV Slope and Age Distributions . . . 125

B.1 Color vs. Color of LRGs, BC03 and M09/11 Models at $z \sim 0.02$ and 0.1 . . . 151
B.2 Color vs. Color of LRGs, BC03 and M09/11 Models at $z \sim 0.2$ and 0.3 . . . 152
B.3 Color vs. Color of LRGs, BC03 and M09/11 Models at $z \sim 0.4$ . . . 153

B.4 χ^2 Distribution from SED Fitting with BC03, M09 and M11 154
B.5 Color vs. Color for LRGs and BC03 Models for $0.0 \leq z \leq 0.04$ 155
B.6 Color vs. Color for LRGs and BC03 Models for $0.08 \leq z \leq 0.12$ 156
B.7 Color vs. Color for LRGs and BC03 Models for $0.18 \leq z \leq 0.22$ 157
B.8 Color vs. Color for LRGs and BC03 Models for $0.28 \leq z \leq 0.32$ 158
B.9 Color vs. Color for LRGs and BC03 Models for $0.38 \leq z \leq 0.42$ 159

C.1 χ^2 Distribution from SED Fitting with BC03, BC03 $\lambda^{-\beta}$, M09 and M11 . . . 162

D.1 SED Fitting Results for ID2 with $\tau < 0$ Models 164
D.2 SED Fitting Results for ID3 with $\tau < 0$ Models 165
D.3 SED Fitting Results for ID2 with $\tau > 0$ Models 166
D.4 SED Fitting Results for ID3 with $\tau > 0$ Models 167
D.5 SED Fitting Results for ID2 with SSP Models 168
D.6 SED Fitting Results for ID3 with SSP Models 169
D.7 SED Fitting Results for ID2 with All Models 170
D.8 SED Fitting Results for ID3 with All Models 171
D.9 PDF of Model Age of ID2&3 with Different SFH Models 172

List of Tables

6.1 Physical Parameters of the Novel CMASS Template SEDs 92
6.2 Summary of Photo-z Quality Metrics 103
6.3 Significance Test for Outlier Classification. 103

7.1 DES-SV Zero Point Offsets 109
7.2 Photo-z Results for *main-main* Test on DES-SV Data 112
7.3 Photo-z Results for *deep-deep* and *deep-main* Tests on DES-SV Data 114

8.1 Best Fitting Parameters for Lensed System in RXC J2248 122
8.2 Best Fitting Parameters for the Objects of R11 and Z12 123

E.1 Asinh Magnitude Softening Parameters 175

Chapter 1

Preface

Throughout history mankind has gazed into the nightly sky wondering what might be its nature. It has been a long way to understand that there are many stars just like our sun, and to accept that there are many structures holding accumulations of stars just like our Milky Way, furthermore, that these galaxies are part of bigger structures known as the cosmic web, and that our location in the Universe is not unique. With the advancements in technology our telescopes and instruments are able to observe celestial objects up to unprecedented distances and enable surveys to cover the full sky. However, the quest of understanding our Universe and its constituents is far from over.

We observe that matter in the Universe is homogeneously and isotropically distributed on scales larger than a few hundred megaparsecs, out at least as far as the present horizon scale of 10^{28} cm. The horizon encompasses a volume which is causally connected to us and which is determined by the distance which light and other signals have had time to travel to us. On smaller scales however, matter is no longer homogeneously and isotropically distributed but accumulated in environments like stellar systems, galaxies, and galaxy clusters.
We show the history of the Universe schematically in Figure 1.1. After the Big Bang the Universe entered a stage of accelerated expansion, known as inflation, which ended in a smooth "graceful exit". Inflation preserves homogeneous patches and expands them onto the present horizon scale, while stretching quantum fluctuations to galactic scales. These small inhomogeneities are the seeds of structure formation. The primordial Universe was very dense and hot, such that photons could not travel freely through the plasma, but were scattered shortly after emission. As the Universe expanded, it cooled down. At the time of recombination, after $\sim 378,000\,\mathrm{yr}$, or at 3600 K, protons could capture electrons and form neutral hydrogen atoms. At this stage the photons decoupled and the Universe became transparent to radiation. These photons are still observable today in the cosmic microwave background (CMB), and exhibit a blackbody spectrum of a much smaller temperature $T \sim 2.7\,\mathrm{K}$, where the decrease in temperature is due to the past expansion of the Universe. The variations in the CMB is believed to trace variations in the matter density and reveal that the Universe was homogeneous and isotropic at the time of recombination with $\Delta T/T \sim 10^{-5}$. The nonlinear structure of today's Universe is caused by these variations in the matter density and the effect of gravity, which attracts matter to regions of higher density such that inhomogeneities grow with time.

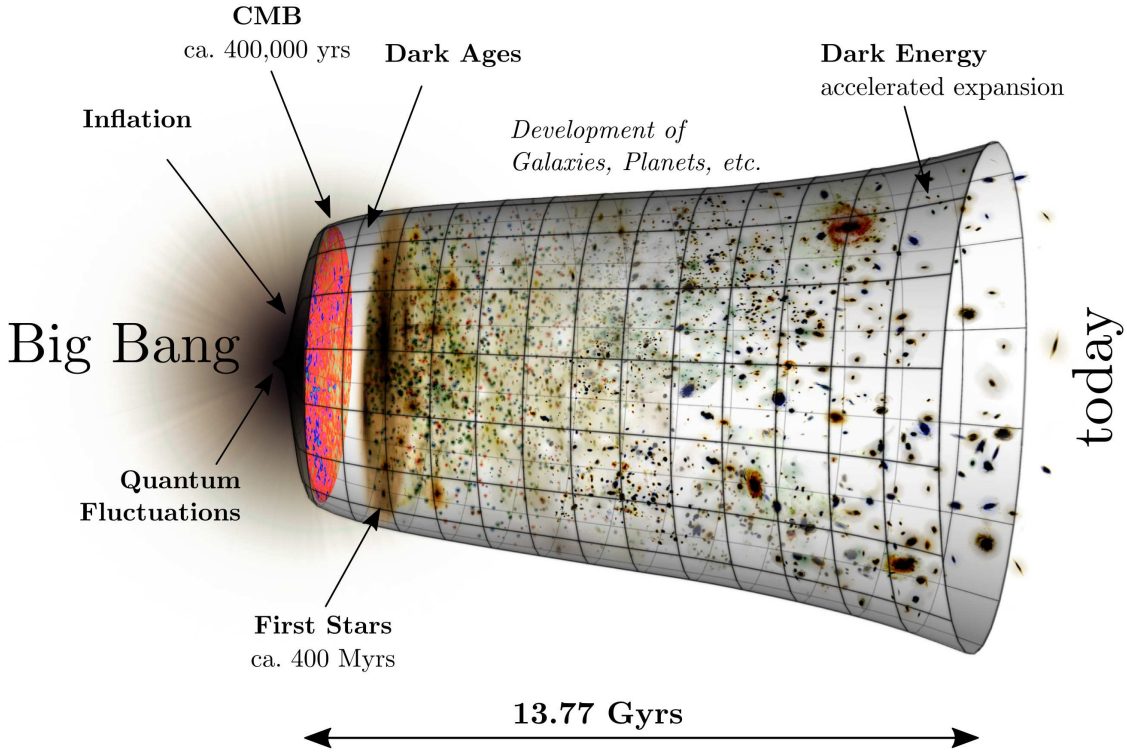

Figure 1.1: Schematic image[1] of the history of the Universe. The cylindric hull in the image qualitatively illustrates the expansion of the Universe as a function of time.

Structures in the Universe have formed hierarchically from low mass objects to higher masses. The first structures to form are believed to be quasars and population III stars which are devoid of metals (except for Lithium from the early Universe). Stars are only small structures that accumulate in stellar clusters and galaxies, which themselves form galaxy groups and clusters. Finally, the biggest structures known are sheets and filaments leaving sparsely populated voids and defining the foam-like large-scale structure (LSS) of the matter distribution (see Fig. 1.2). The biggest structures denote the so-called *End of Greatness*, because on scales larger than that, the matter distribution is homogeneous and isotropic.

We are able to observe only stars and gas on account of their ability to interact electromagnetically. However, the luminous baryonic matter constitutes only a small fraction of the total mass present in the Universe. The dominant part of the matter content is a component that interacts gravitationally but not electromagnetically, therefore it is invisible and was dubbed *dark matter*. It was first postulated to explain the high rotational velocity of spiral galaxies, as well as the orbital velocities of galaxies in clusters, which both demanded a higher mass than that of the observable luminous matter. After this discovery, many different independent observations could finally be explained by a non-luminous mass component (like structure formation). A very important tool for the analysis of the dark matter distribution is gravitational lensing, which uses the effect that space-time is curved in the presence of a

[1] image credit NASA/WMAP Science team `http://science.nasa.gov/missions/wmap/`; modified

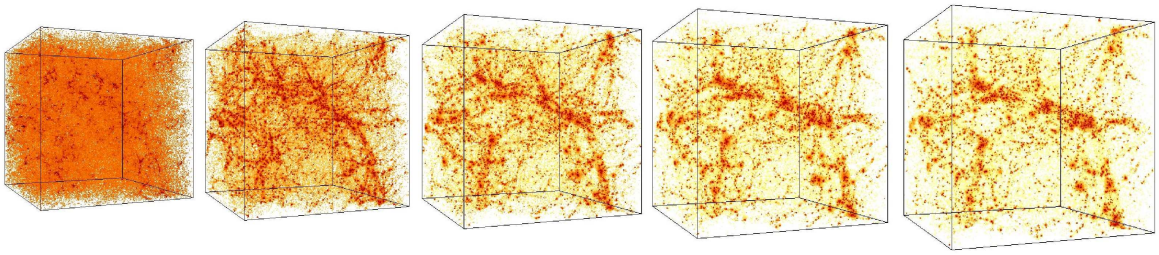

Figure 1.2: Simulation of hierarchical structure formation (from left to right) in a box of 43 Mpc size[2]. The simulation is sampled at different ages of the Universe from 0.5 Gyr until now (13.7 Gyr).

gravitational field, and therefore light is bent near mass overdensities. The nature and properties of dark matter are nowadays among the topics of greatest interest in cosmology and particle physics.

The Cosmological Redshift

In 1927 Georges Edouard Lemaître first proposed the theory of an expanding Universe. At that time *General Relativity* (Einstein, 1916) was already established, and different cosmological models were proposed, all of which considered a static Universe. Lemaître also derived the so-called *Hubble-law* often misattributed to Edwin Hubble who published it only two years later. The Hubble law applies to a homogeneous and isotropic Universe and reads

$$\mathbf{v} = H(t) \cdot \mathbf{r}, \tag{1.1}$$

with the *Hubble parameter* H. It states that the recession velocities of galaxies are larger at greater distances.
Curiously, the Universe is not just expanding, but entered a stage of accelerated expansion shortly before the present time, which was confirmed from the observation of the light curves of supernovae type Ia (e.g., Perlmutter et al., 1999). Before that, the Universe was matter dominated, a period needed for structure formation to take place. The accelerated expansion of the Universe is one of the greatest mysteries of contemporary cosmology (aside from the nature of dark matter) and one of the most heavily researched subjects. For an explanation, a component in the mass-energy content of the Universe was introduced which has negative pressure and is driving the expansion. This *dark energy* is the prevalent constituent of the Universe, making up $\sim 70\,\%$ of the mass-energy content, where the rest consists of dark matter and baryonic matter (e.g., Planck Collaboration et al., 2014).

The expansion of the Universe leads to a redshift of light. This *cosmological redshift* is therefore not a result of a motion of galaxies away from each other, but a consequence of the expansion of space-time itself. Features in the spectrum of a galaxy appear at higher wavelengths λ for an observer not located in the rest frame of the galaxy. Furthermore, the greater the distance of a galaxy from the observer, the greater is the redshift of the spectrum, since the journey of the photon is longer (cf. Eq. 1.1), and the Universe expands by a larger

[2]Simulations were performed at the National Center for Supercomputer Applications by Andrey Kravtsov (The University of Chicago) and Anatoly Klypin (New Mexico State University). Visualizations by Andrey Kravtsov. Image was modified.

factor during the light travel time.
We introduce the *redshift parameter* z that represents the fractional shift in the wavelength of a photon emitted by a galaxy at time $t_{\rm em}$ and observed at $t_{\rm obs}$:

$$z + 1 \equiv \frac{\lambda_{\rm obs}}{\lambda_{\rm em}} = \frac{a(t_{\rm obs})}{a(t_{\rm em})}, \tag{1.2}$$

where $a(t)$ is the scale factor which describes the distance between observers at time t. The second equality can be easily derived by integration over the geodesic (the path in space-time with the shortest length) equation of a photon (cf. Mukhanov, 2005).
In the following, we will use the term *redshift* unambiguously for *redshift parameter*, because we will not concern ourselves with Doppler redshifts caused by the peculiar velocities of galaxies. From Equation 1.2 we can infer that z directly corresponds to $t_{\rm em}$. Therefore, we can express time-dependent quantities as functions of z which is directly measurable, in contrast to t for which we have to know the cosmology. To derive a relation between redshift and time we can combine the energy conservation equation and the Friedmann equation following the textbook of Mukhanov (2005):

$$t = \int_z^\infty \frac{dz}{H\,(z)\,(1+z)}, \tag{1.3}$$

where $z \to \infty$ corresponds to $t = 0$. Knowing the light-travel time, we can easily derive also the distance to the light source, the so-called *light travel time distance* and thereby use z also as a measure of distance.

Redshift Measurements in Cosmology

Observing galaxies that are located at higher distances from us also means that we observe backwards in time. As redshift measurements can be translated into distances, assuming a cosmology and therefore $H(z)$, they carry information about the three-dimensional nature of the Universe as well as its history. Since greater distances cannot be measured directly this makes redshift information a very important prerequisite in many cosmological contexts.
Redshift information is most directly used in analyses of the large scale structure of the Universe. Apart from the angular coordinates at which galaxies are observed, z holds the information about the third spatial dimension. From the three-dimensional distribution of galaxies their two-point and three-point correlation functions can be computed. Thereby, we can investigate the clustering of galaxies as a function of their properties, like luminosity, colors, or mass. Also, we can compare to simulations and evaluate dark matter models on the basis of the formation and evolution of galaxies (located within their dark matter halos), and on their clustering. Furthermore, with redshift information one can determine the scale of *Baryon Acoustic Oscillations* (BAOs) which are imprinted on today's matter distribution. BAOs originate from sound waves in the primordial Universe before the photons decouple and can travel freely. These sound waves are created by the pressure in overdense regions in the plasma that drove the baryonic matter spherically outwards from the center of the region before the decoupling of photons. The sound horizon scale at recombination is preserved by the CMB (see above), and was increased afterwards by the expansion of the Universe. By comparing with the sound horizon today which is imprinted in the large scale structure, we get information about the expansion history of the Universe and therefore dark energy.

Another important area of application of redshift information is in the context of gravitational lensing. General Relativity predicts a curvature in space-time around gravitating objects. Therefore, light paths are bent in gravitational fields and distorted images or even multiple images of objects behind gravitating masses can be observed. The space-time curvature is not only caused by baryonic matter but also dark matter, which makes gravitational lensing an important tool for the detection of dark matter and for analyses of its distribution around galaxies and in galaxy clusters. The distances of the lenses and the sources enter the lens equation and have to be determined by redshift measurements (Eq. 8 in Schneider et al., 2006).

As shown in Equation 1.3, the redshift can be regarded as a measure of time as well. When we analyze galaxy properties such as age, metallicity, colors, but also morphology, as a function of redshift we can analyze the evolution of galaxies. Looking backwards in time, we get information about, e.g., their stellar content as a function of time and therefore their star formation history. We will carry out such an analysis in Section 5.5.2, where we observe that the spectra of our investigated galaxies behave as spectra of aging galaxy populations with decreasing redshift.

There are many applications in cosmology and galaxy evolution where redshift information is crucial, wherefore we aim to determine z as precisely as possible. A further goal is that redshifts have very low systematical errors, while keeping statistical errors small. This is especially a challenge when redshifts cannot be measured precisely but have to be estimated statistically. The statistical derivation of redshifts from photometric data and the analyses of physical properties of the galaxies are the main topics of this work. In the next chapter we first give a brief overview of astronomical observations and then introduce the concept of estimating redshifts from photometry. We will detail template fitting techniques, which will be used particularly in this work. Since we will work mainly with data from the Sloan Digital Sky Survey (SDSS), we will present the survey and the selection of the galaxy samples in Chapter 3. After that, we shortly discuss galaxies in Chapter 4, concentrating on elliptical galaxies which are the predominant focus in this work. In Section 4.2 we describe the approach of galaxy stellar population synthesis to create model spectra of galaxies, and we analyze how well existing model spectra can describe the luminous red galaxies (LRGs) from the SDSS. After that we describe how we create model spectral energy distributions that are generated specifically for the Sloan LRG sample in Chapter 5. We will use these in the estimation of redshifts from photometry in Section 5.3, and compare with the results of other techniques in Section 5.4. To conclude this Chapter we analyze the novel models in terms of their spectra and properties in the ultraviolet (UV) wavelength range (Section 5.5). In Chapter 6 we create model spectra for SDSS galaxies at higher redshifts. Since the usual semi-empirical models cannot match the colors of the data, we modify the slope of the continuum of the spectral energy distributions. We then generate and select a small set of models for redshift estimation and compare our results with those from SDSS. We turn our focus to the redshift estimation with Dark Energy Survey science verification data in Chapter 7. Finally, we present the results of estimating the physical properties of a high redshift, gravitationally lensed system in Chapter 8, which was the focus of a strong lensing analysis. In the last chapter (Chap. 9) we give a summary and conclusions.

Throughout this thesis we adopt a ΛCDM cosmology, where Λ denotes the cosmological

constant that accounts for dark energy, and the universe contains cold dark matter (CDM). We assume $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the cosmological parameters[3].

[3]These are not the latest values from the Planck Collaboration et al. (2014) which read $\Omega_m \approx 0.32$ and $H_0 = 67.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The difference in values would alter the rest frame results in this thesis only at a negligible level, which is much smaller than the errors of the derived rest frame properties.

Chapter 2

Photometric Redshifts

In this chapter we introduce the concept of photometric redshifts which are the main subject of this thesis. Before that, we describe the concept of today's astronomical observations in Section 2.1 and give an introduction on observational data. We explain why many surveys take photometric data and motivate why we estimate redshifts from photometry. After that, we introduce two main concepts that are being used in photometric redshift determination (Sec. 2.2) and detail the photometric redshift code used in this work. Finally, at the end of Section 2.2.2 we briefly mention also a technique that is applicable only to high redshift objects.

2.1 Introduction to Astronomical Observations

Astronomical objects, such as stars and galaxies, can be detected explicitly by observing the light (or photons) they emit. We only are able to observe a fraction of the total amount of radiative energy emitted by the source, since the light is emitted in all directions. Therefore, the total observed radiative energy of an object is measured as energy per time and unit area, that is $\frac{\mathrm{W}}{\mathrm{m}^2}$ in SI units, or $\frac{\mathrm{erg}}{\mathrm{cm}^2\,\mathrm{s}}$ in cgs units, where $1\,\mathrm{erg} = 10^{-7}\,\mathrm{Joule}$. As a function of wavelength λ (or frequency ν) we get the *spectral flux density* f_λ (or f_ν), which we will term *flux* for simplicity in the following. Unless stated otherwise *flux* will be used interchangeably for f_λ. Observing the flux of an object as a function of wavelength or frequency, yields the *spectrum*, or *spectral energy distribution* (SED) of that object. It is usually given in units of $[f_\lambda] = 1\,\frac{\mathrm{erg}}{\mathrm{cm}^2\,\mathrm{s}\,\mathrm{\AA}}$, or $[f_\nu] = 1\,\frac{\mathrm{erg}}{\mathrm{cm}^2\,\mathrm{s}\,\mathrm{Hz}} = 10^{23}\,\mathrm{Jansky} = 10^{23}\,\mathrm{Jy}$.
However, astronomers most often use *magnitudes* instead of fluxes to describe the brightness of astronomical sources. The origin of the magnitude system dates back ~ 2000 years to the Greek astronomer Hipparchus, or the Alexandrian astronomer Claudius Ptolemy (depending on the reference), who defined the brightest stars as first class (or magnitude) to barely visible stars to sixth class (or magnitude). The human eye was believed to perceive brightnesses logarithmically and therefore in the middle of the 19th century the astronomer N. Pogson introduced the $\log_{10}$ magnitude system still used today (Pogson, 1856), where he defined a star of magnitude 1 to be 100 times brighter than a magnitude 6 star. Afterwards, it was discovered that some magnitude 1 stars were brighter than others, and furthermore stars fainter than mag = 6 were found, such that the magnitude range had to be extended outside the range of $1 - 6$. E.g., Vega (or αLyrae) has magnitude zero which is unintuitive, but can

serve as a standard ruler (see below).
The difference in magnitudes $m_{\star,i}$ of two astronomical objects can be expressed in terms of their fluxes $f_{\star,i}$

$$m_{\star,1} - m_{\star,2} = -2.5 \cdot \log_{10}\left(\frac{f_{\star,1}}{f_{\star,2}}\right), \tag{2.1}$$

where the factor in front of the logarithm is approximately the fifth root of 100 due to Pogson's magnitude system. Using Vega as second star in this equation yields

$$m_{\star} = -2.5 \cdot \log_{10}\left(\frac{f_{\star}}{f_{\alpha\mathrm{Lyr}}}\right), \tag{2.2}$$

the magnitude of a star in the Vega system. Therefore, the magnitudes of any survey can be calibrated by the observed fluxes on Vega, which serve as zero points. Nowadays however, a calibration to a theoretical AB star (Gunn & Peterson, 1965) is more commonly used, since the observational results of Vega can vary from survey to survey and a universal calibration is difficult. This theoretical AB star is defined to have zero magnitude and $f_\nu = 3631\,\mathrm{Jy}$, wherefore $m_{\star,\mathrm{AB}} = -2.5 \cdot \log_{10}(f_{\nu,\star}/3631\,\mathrm{Jy})$. Throughout this work, magnitudes are given in the AB system.
Often we want to compare the intrinsic brightnesses of galaxies located at different redshifts (i.e., distances) with each other. We therefore compute the *absolute magnitude* which is the magnitude the object would have were it located at 10 pc. 1 Parsec ($1\,\mathrm{pc} \equiv 3.08 \cdot 10^{16}\,\mathrm{m}$) is the distance at which one astronomical unit (the distance of the earth to the sun) is observed at an angle of $1''$. The absolute magnitude M is then computed by

$$M = m - 5 \cdot \log_{10}\left(\frac{r}{10\,\mathrm{pc}}\right), \tag{2.3}$$

where r is the distance of the object, and therefore depends on the redshift and cosmology.
The magnitude system has some caveats compared to fluxes. Fluxes of zero cannot be expressed in terms of magnitude since the logarithm of zero diverges. Furthermore, also negative fluxes, which can arise due to errors in the measurement, cannot be converted into the magnitude system. Lupton et al. (1999) circumvent this by introducing arsinh magnitudes, sometimes informally called *"luptitudes"*, which use the arsinh instead of the logarithm in Equation 2.2. They are used in the Sloan Digital Sky Survey (SDSS) discussed in Chapter 3 and we will give a description in Appendix E.
A galaxy *color* is defined by the flux ratio, or the magnitude difference between two wavelength ranges:

$$m_i - m_j = -2.5 \cdot \log_{10}\frac{f_i}{f_j}. \tag{2.4}$$

i and j denote the wavelength ranges, i.e., the optical filters through which the object has been observed.

After this introduction on fluxes, magnitudes, and colors we show an exemplary spectrum of a galaxy at redshift $z = 0.46$ from the SDSS spectroscopic survey (ID 7304 in our catalog, or SDSS object ID 587725505018986933) in the upper panel of Fig. 2.1. If the reader would like to access further information about our exemplary object, we would like to point her/him to the SDSS-II Sky Server Object Explorer[1]. We can see from Fig. 2.1 that the flux is not a

[1] `http://skyserver.sdss2.org/dr7/en/tools/explore/obj.asp?id=587725505018986933`

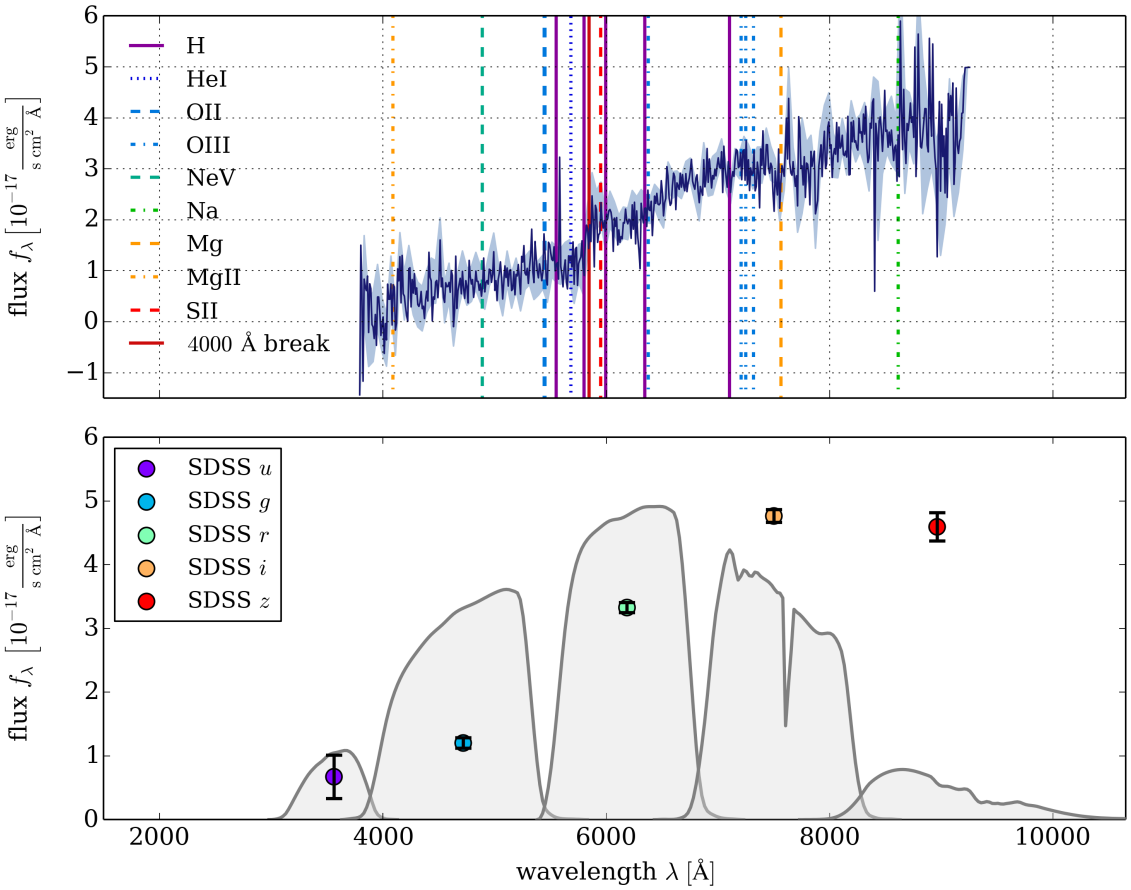

Figure 2.1: *Upper Panel:* Spectrum of a red galaxy from the SDSS-II spectroscopic survey (SDSS object ID 587725505018986933, or 7304 in our catalog) located at $z = 0.46$. The light blue area shows the error of the measurement. Selected spectral lines are indicated by vertical lines (see legend for identification). The 4000 Å break is highlighted by a solid red line. Negative flux values are, of course, not physical but due to noise in the background. The data is only available for wavelengths between 3800 Å and 9200 Å which is covered by the SDSS-II spectrographs. *Lower Panel:* The same galaxy but observed through the five SDSS filters (u, g, r, i, and z). The filter curves are plotted in gray and scaled such that they are visible in the plot. Error bars denote the photometric measurement uncertainties.

smooth distribution as a function of wavelength, but instead exhibits many sharp variations and even stronger features which originate from the chemical composition of the observed galaxy. These features are elicited by *spectral lines* which are created by the absorption or emission of photons by chemical elements present in the galaxy. From atomic physics we know the (discrete) energies of electron transitions (and hence the wavelengths) within the atoms of several elements and can therefore associate the observed absorption or emission lines with their corresponding elements. By identifying the lines and continuum features in the spectrum and relating their positions in the observed wavelength λ_{obs} with their rest frame (or emitted frame) wavelength λ_{em} through $\lambda_{\mathrm{obs}}/\lambda_{\mathrm{em}} = 1+z$, one can determine the redshift z. The error in the redshift measurement thus only depends on the measurement error of the spectrum, and therefore the position of the spectral lines, as well as the resolution of the spectrum $\lambda/\Delta\lambda$.

However, today statistically meaningful volumes of the Universe are being probed by photometric surveys. These are well-known projects like the *Sloan Digital Sky Survey* (SDSS; York et al., 2000; Abazajian et al., 2009; Ahn et al., 2014), the *Panoramic Survey Telescope and Rapid Response System* (Pan-STARRS; Kaiser et al., 2000, 2010), the *Dark Energy Survey* (DES; Flaugher, 2005; Flaugher et al., 2012; Diehl & Dark Energy Survey Collaboration, 2012), or the planned *Large Synoptic Survey Telescope* (LSST; Tyson et al., 2003) that use ground-based telescopes, as well as the *Euclid* mission (Laureijs et al., 2011) of the European Space Agency that will survey the Universe from space. The telescopes and cameras are equipped with the latest detector technology and are able to survey large areas of the sky with deep multicolor photometry in several broad passbands. In contrast to spectroscopic surveys, photometric surveys can observe ~ 100 times more astronomical sources in the same time, while the instrument costs are significantly lower. Also, the signal-to-noise is lower in spectroscopy where the flux is distributed over several pixels, and the sensitivity varies reciprocal to the resolution. With imaging one can detect much deeper (less luminous) sources in the sky, which also minimizes selection effects. For these reasons most current surveys are photometric.

Nevertheless, photometric surveys have caveats compared to spectroscopic observations. By observing through photometric filters one has limited information about the spectrum of an object, since spectral lines cannot be resolved by broad-band filters. The fluxes in the passbands give an idea about the overall shape of the SED, but no information on the positions of features that are narrow in wavelength and therefore need a higher resolution in order to be detected. Figure 2.1 gives an example of the difference between spectroscopic and photometric observations. While the upper panel shows the galaxy spectrum, in the lower panel the photometric fluxes in the five SDSS filters (u, g, r, i, and z), are plotted along with the filter throughput curves. Features in the SDSS filter curves are due to atmospheric extinction, whereas the overall shape of the whole filter system (low throughput at lower and higher wavelengths) is due to the detector quantum efficiency. From the data plotted in the lower panel of Figure 2.1 we may not be able to infer the position of specific spectral lines from the upper panel, but we can observe features in the overall shape that the underlying SED must exhibit. I.e., we can see that there is a big increase in flux between the g and the r band, therefore this feature must be visible also in the SED of the galaxy. We can associate this step in flux with the *4000 Å break* (described in more detail in Section 4.1.1) which is due to the accumulation of absorption lines and the stellar population. These break features in the observed fluxes and the overall shape of the SED gives us a means of determining the

redshift of a galaxy by photometric data alone, a so-called *photometric redshift*.

2.2 Photometric Redshifts

It has been more than 50 years since the inception of photometric redshifts (photo-zs). Baum (1962) estimated the first redshifts of galaxies by comparing the mean galaxy spectra of six bright elliptical galaxies in the Virgo cluster with photometry in 9 medium-band filters from the cluster 3C395, which was extracted from photographic plates. The cluster redshift estimated such to be $z \approx 0.44$ is close to the actual value of 0.46. This process is known today as *SED* or *Template Fitting* (see Sec. 2.2.2).
It was Puschell et al. (1982) who, by estimating redshifts of faint radio galaxies through broad-band photometry, were the first to use the term *photometric redshift*. In the paper of Loh & Spillar (1986), who also used SED fitting to get redshifts through photometry in six medium bands, the term *photometric redshift* first entered a paper title. Since then, a huge amount of papers have been published on the topic and various methods to estimate photo-zs have been presented. Koo (1999) suggests the following definition for photometric redshifts:

> [...] photometric redshifts are those derived from only images or photometry with spectral resolution $\lambda/\Delta\lambda < 20$. My choice of 20 is intended to exclude redshifts derived from slit and slitless spectra, narrow-band images, ramped-filter images, Fabry-Pérot images, Fourier transform spectrometers, etc.

As mentioned before, there are several methods of determining photometric redshifts applied by the astronomical community these days. These techniques are usually divided into two main categories: *Template Fitting* and *Empirical Methods*. We will discuss both approaches in the following Sections (2.2.1 and 2.2.2). We will start, although not chronologically correct, with empirical methods and will then turn to the template fitting technique afterwards, since we will work with a template fitting code in this thesis. This code will be described in Section 2.2.2.

2.2.1 Empirical Methods

Generally speaking, empirical methods derive a relation between photometric observables of a galaxy and their redshift. When this relation is established, it can be used to map the same observables of galaxies without measured redshift information to z values or a probability distribution function (PDF) $P(z)$ from which z can then be extracted. If a $P(z)$ is derived, the resulting photo-z value is usually determined by the mode, median, or mean of the distribution. The photometric observables used in empirical approaches are typically colors and/or magnitudes or fluxes, but can also be other observables, like half-light radii, galaxy morphology, etc..
Earliest approaches with empirical methods were made by Connolly et al. (1995), who identified iso-z layers in the multidimensional color space and used polynomial fitting to derive redshifts. Shortly afterwards, Brunner et al. (1997) introduced piece-wise fitting of 2nd order polynomials of galaxy colors, and Wang et al. (1998) derived redshifts using a linear function of three photometric colors. Csabai et al. (2003) performed linear fits to k nearest neighbors in color space for the SDSS early data release (see results on SDSS data in Sec. 5.4).

We are living in the digital age and processor power and memory capacity has increased roughly exponentially during the last decades. Therefore, computationally expensive codes became feasible and resulted in the birth of machine learning techniques which can be applied to photo-z estimation. The first published machine learning photometric redshift algorithm is `ANNz` (Collister & Lahav, 2004), which uses an artificial neural network (ANN) for the estimation of photometric redshifts. Since then, many more machine learning techniques were used for photo-zs:

- **Support Vector Machines**, e.g., Wadadekar (2005) and Zheng & Zhang (2012)
- **Relevance Vector Machines**, e.g., Sánchez et al. (2014), Rau et al. (in prep.)
- **Boosted Decision Trees**, e.g., Gerdes (2009, `ArborZ`)
- **Gaussian Processes**, e.g., Way et al. (2009)
- **Diffusion Maps**, e.g., Richards et al. (2009) and Freeman et al. (2009)
- **Random Forests**, e.g., Carliles et al. (2010, results on SDSS data see Sec. 6.3.2) and Carrasco Kind & Brunner (2013, in combination with decision trees, `TPZ`)
- **Self Organizing Maps**, e.g., Carrasco Kind & Brunner (2014b, `SOMz`)

We will not go into a review of the different techniques because, as mentioned before, this work will concentrate on photo-z estimation with template fitting. Nevertheless, we want to emphasize some subtleties of empirical methods, and point out advantages and disadvantages in the following.
As explained above, empirical methods *"learn"* or *train* a relation between photometric observables and spectroscopic redshifts (spec-zs), in order to apply this relation to objects with unknown z. These estimators are in general regression methods, which estimate the conditional mean of the redshift given photometric observables $\boldsymbol{\phi}$. Instead of predicting the conditional mean only, which is just a statistic of the underlying conditional PDF $P(z|\boldsymbol{\phi})$, some efforts were made in determining the whole PDF (Carrasco Kind & Brunner, 2013; Bonnett, 2013; Rau et al., in prep.). Since the point predictions and the PDF are constructed on the basis of a training sample with known spectroscopic redshifts, it is of paramount importance for accurate photo-zs that the provided training set is a good representation of the galaxies in the query set. This is not an simple task, since, as mentioned earlier, spectroscopy is only feasible for more luminous objects and a spectroscopic sample therefore is bound to suffer from selection effects. On the bright side, effects of dust or spectral evolution are taken into account explicitly, and the physical properties of the underlying galaxy sample do not have to be known. Furthermore, precise knowledge about the photometric system of the telescope is not necessary (unlike template fitting where this knowledge is of great importance). Also, the data need not necessarily be well calibrated as long as the training and the query set suffer from the same calibration errors. The downside of this feature is that one cannot simply apply an already learned parameter-redshift relation to any other data set, since the photometry of an object varies with different telescope and camera combinations. Therefore, one has to train the empirical code again with a representative training set of that specific survey. Another downside is that most of the empirical codes currently available do not take photometric measurement uncertainties into account.
While most of the (dis-)advantages of template fitting methods have implicitly been covered

in the previous paragraph, the next section will provide more detailed information about template fitting codes in general and, specifically, the code used in this work.

2.2.2 Template Fitting

The idea of template fitting has already been briefly broached at the beginning of this chapter, where we explained how Baum (1962) estimated the first photometric redshifts. The basics of the technique have not changed significantly since Baum's first attempts. Simply speaking, template fitting searches for the minimum value of the difference between observed colors and synthetic colors derived from model (or template) SEDs. Model SEDs are being redshifted through $\lambda(z) = \lambda_{\text{rest frame}}(1+z)$ for various values of z, equivalently to if we would move a galaxy emitting that model spectrum further away from us (increasing the redshift) and observe it at these distances. Then the spectrum is projected through the filter functions of the survey, e.g., the SDSS u, g, r, i, z filter throughputs (see also the lower panel of Fig. 2.1) to obtain a simulated photometric observation of a galaxy with that SED. To predict the fluxes of an SED in a filter i it is multiplied by the filter throughput and integrated over:

$$f_i = \int_{-\infty}^{\infty} \text{thrpt}_i(\lambda)\, f_\lambda(\lambda)\, d\lambda.$$

The synthetic colors are then calculated and compared to the colors of the observation, which yields a χ^2 value for each redshift step and model SED. The best model-redshift combination is then obtained by searching for the minimal value in the χ^2 distribution:

$$\chi^2_T(z) = \sum_{i \in \{\text{filter}\}} \left[\frac{f_{\text{obs},i} - a_T \cdot f_{T,i}(z)}{\sigma_{\text{obs},i}} \right]^2, \tag{2.5}$$

where $f_{\text{obs},i}$ is the observed flux in filter i with corresponding observational errors σ_i, and $f_{T,i}$ are the predicted fluxes of the template T. a_T is a scaling factor with which the model SED is multiplied to fit the data.

Many of the currently used template fitting photometric redshift codes employ a *Bayesian* approach, where the *Bayesian Theorem* is applied to the probability distribution:

$$P(\boldsymbol{\mu}|C,m) \propto \mathcal{L}(C,m|\boldsymbol{\mu}) \cdot P(\boldsymbol{\mu}), \tag{2.6}$$

where C and m denote the colors and magnitudes of the photometric data. $\boldsymbol{\mu}$ is the array of the model redshift z and luminosity (or absolute magnitude M) in a specific filter band. $\mathcal{L}(C,m|\boldsymbol{\mu})$ denotes the likelihood function which is proportional to $\exp\left(-\chi^2/2\right)$, with χ^2 from equation 2.5. The second factor in Equation 2.6 is the prior distribution of a template T

$$P(\boldsymbol{\mu}|T) = P_M(M|T) \cdot P_z(z|T). \tag{2.7}$$

These priors are specified *a priori* and enable us to assign specific probability distributions in redshift and absolute magnitude for every SED template. In general, any functional form can be used for the priors, but not all are sensible. We will explain the prior distributions used in this work later in this section.
By making use of priors one can assign additional probabilities to the individual templates, making sure that degeneracies in colors are lifted. These degeneracies can arise when several

SED types at various redshifts can produce the colors of an observed object equally well. Therefore, it is very important to optimize the redshift and the luminosity priors for each galaxy type, depending on the filter set and the data, i.e., photometric depth and galaxy properties (which, in turn, depend on the galaxy selection function). E.g., the Lyman-break of a galaxy shifted to high redshifts can look like the 4000 Å break of an older galaxy at lower redshifts, at least if one does not have photometry in the infrared (IR). We can assign a higher probability to the model SEDs of older (redder) galaxies at lower z, whereas young (bluer) galaxies can also be found at very high redshifts. We will make particular use of the z prior in Sections 5.3 and 6.3.1 when using SEDs specifically designed for different redshift ranges. Furthermore, the SEDs of luminous galaxies can differ from those that are fainter (cf. Sec. 5.5.2). This can also be accounted for by setting the luminosity prior accordingly.

There are several template fitting photometric redshift codes which differ mainly from one another through the template SEDs provided. The first published, and still widely used, photo-z codes are `LePhare`[2] (Arnouts et al., 1999; Ilbert et al., 2006) and `BPZ` [3] (*Bayesian Photometric Redshifts*; Benítez, 2000; Benítez et al., 2004; Coe et al., 2006). Since then many more photometric redshift codes have been devised. Here, we want to name a few popular template fitting algorithms used by the astronomical community:

- `HyperZ`[4] solely a maximum-likelihood method, by Bolzonella et al. (2000)
- `PhotoZ` Bayesian photo-z code used in this work and explained in Section 2.2.2 by Bender et al. (2001).
- `Z-PEG`[5] by Le Borgne & Rocca-Volmerange (2002), a combination with the stellar evolution code `PÉGASE`
- `IMPZ` a galaxy-quasar approach including extinction by Babbedge et al. (2004)
- `ZEBRA`[6] by Feldmann et al. (2006), Bayesian code that adjusts spectra to better match the photometry
- `EAZY`[7]Bayesian code that improves on templates through linear combinations of them, Brammer et al. (2008)

Template fitting methods are, in principle, very simple to implement and can be applied to any photometric data without the need of a training set and hence, spectroscopic redshifts. The fact that spectroscopic information is unnecessary is the greatest advantage of template fitting codes (compared to empirical approaches). However, one has to have a set of model SEDs which are able to represent the observed galaxy types. Then, parallel to the estimation of photo-zs, one is provided also with the rest frame properties of the best fit models such as the SED type, since it is known for the model SEDs, and the absolute magnitude which can easily be derived. We will look in more detail into available model SEDs and their performance in matching the SDSS luminous red galaxy sample, and the BOSS (Baryon

[2] `http://www.cfht.hawaii.edu/~arnouts/lephare.html`
[3] `http://www.stsci.edu/~dcoe/BPZ/`
[4] `http://webast.ast.obs-mip.fr/hyperz/`
[5] `http://imacdlb.iap.fr:8080/cgi-bin/zpeg/zpeg.pl`
[6] `http://www.astro.ethz.ch/research/Projects/ZEBRA/`
[7] `http://www.astro.yale.edu/eazy/?home`

Oscillation Spectroscopic Survey) CMASS sample in Sections 4.2.2 and 6.1.1.
Apart from the template set being representative, it is also important to have a detailed model of the photometric system of the telescope to predict the colors of the models accurately. Furthermore, the zero point offsets have to be well known. As mentioned earlier in this Chapter, the data are calibrated to a specific star that serves as a zero point. However, systematic offsets from that zero point are not unusual and have to be corrected for. If the data set is badly calibrated, one can recalibrate the zero points *a posteriori* if a catalog with spectroscopic redshift measurements is available. We can then fit the models to the data with known spec-zs and calculate the offset between the magnitudes predicted by the models and the observed magnitudes. These corrections can then be applied to the photometric data set. The same procedure can also be done using a data set with accurate photometric redshifts, as shown by Molino et al. (2014). They used the precise photometric redshifts of blue galaxies with prominent emission lines from the narrowband survey ALHAMBRA (Moles et al., 2008) to correct the zero points (and thereby the redshift estimation) of the whole galaxy sample. This is of course only possible where accurate photometric redshifts are available, which is feasible for a narrowband survey like ALHAMBRA where emission lines can be resolved, but not for broad-band surveys like SDSS, Pan-STARRS, or DES. The correction of the zero points in this manner is real if one has a representative model set for the underlying data, but can introduce magnitude shifts that are non-existent if the model set is not a good representation. We will perform this correction on the DES-SV data in Chapter 7, where we have a spectroscopic calibration set, and where the applied offsets greatly improve the quality of the photometric redshifts of the query set.

The `PhotoZ` Code

In this section we want to briefly detail out the properties of the Bayesian template fitting photometric redshift code used in this thesis, `PhotoZ` (Bender et al., 2001). We mentioned in the last section that `PhotoZ` uses a Bayesian approach and enables the user to set prior probabilities to the individual model SEDs. Parallel to the photo-z estimation, one can name a library of stellar SEDs and fit those to the photometry with various extinction values. This can be very useful as a means of separating a catalog into stars and galaxies. We will not need to use this feature since we will deal with preselected galaxy catalogs throughout this work.
Furthermore, we can assign additional errors $\sigma_{\mathrm{dat},i}$ to the data if we feel they are underestimated, or to change the weights of different filters in the maximum likelihood fit. We can also define additional uncertainties σ_T for the templates. Then the denominator of Equation 2.5 reads

$$a_T \cdot \sigma_T + (\sigma_{\mathrm{obs},i} + \sigma_{\mathrm{dat},i}). \tag{2.8}$$

In principle any template set can be introduced to the code, while it is recommendable to adjust the prior probabilities sensibly and according to the individual model SEDs (see Sec. 2.2.2). The priors of the `PhotoZ` code are parameterized as

$$P_x(x|T) \propto \exp\left[-\ln(2)\left(\frac{x-\mu_x}{\sigma_x}\right)^{p_x}\right], \tag{2.9}$$

where $x = z, M$. μ_x is the most probable, i.e., expected value of x, whereas σ_x controls the width of the prior distribution and p_x its steepness. We use only even values for p_x, as the

integral over the probability would diverge for odd values of p_x. The priors were chosen in this functional form to be able to vary the transitions between regions of different probabilities in such a way that they may be both smooth and have a box-like shape. Figure 2.2 demonstrates the prior probability distributions for different values of σ_x and p_x.

An example output of the `PhotoZ` code is shown in Figure 2.3 for the exemplary galaxy 7304 whose spectrum and photometry was already presented in Fig. 2.1. In the upper panel of Fig. 2.3 one can see the PDF as a function of redshift $P(z)$. It is plotted for the eight best fitting models, where the PDF of the best fitting model is highlighted by a light green line. The stacked PDF of all SEDs in the model set is drawn by a dashed line. The lower panel of Figure 2.3 shows the best fitting model SED at the best fitting redshift $z_{\rm best}$, the mode of its PDF from Equation 2.9. The underlying photometry in the SDSS filters is represented by red points with error bars that denote the photometric measurement uncertainties of SDSS. If a second solution exists with $|z_{\rm 2nd\,best} - z_{\rm best}| > 0.2 \cdot (1 + z_{\rm best})$, $z_{\rm 2nd\,best}$ is included in the output of the code (dark green line in Fig. 2.3). The second best fitting model can be useful if one has to deal with degeneracies of colors (cf. Section 2.2.2) and the second best solution is more close to the spec-z. The values in brackets after $z_{\rm best}$ and $z_{\rm 2nd\,best}$ are the photometric redshift errors δz_* estimated by the `PhotoZ` code. They are calculated as the root of the mean squared deviation from the best fitting value,

$$\delta z_* = \left(\sum_z (z - z_*)^2 \cdot P(z) \right)^{1/2}, \tag{2.10}$$

where z_* denotes $z_{\rm best}$, or $z_{\rm 2nd\,best}$ respectively, and z runs over discrete redshift steps. The best fitting SED of the fit to the spectroscopic redshift is plotted in black in Figure 2.3, and the best star is shown by a dashed gray line. The predicted photometric fluxes in the SDSS filters and four additional filters are presented by circles (or crosses in the case of the stellar model).

The `PhotoZ` code has been successfully applied in the past to a variety of photometric catalogs (Drory et al., 2001; Gabasch et al., 2004; Drory et al., 2005; Feulner et al., 2005; Brimioulle et al., 2008; Gabasch et al., 2008; Brimioulle et al., 2013; Gruen et al., 2013, 2014; Sánchez et al., 2014). It is also part of the PS1 Photometric Classification Server (Saglia et al., 2012) and the Astro-WISE package (Saglia et al., 2013).

Dropout Technique

Before concluding the chapter on photometric redshifts, we want to briefly describe the *dropout* technique to estimate redshifts from photometry. It can be applied only to star forming galaxies at high redshifts, but it should be mentioned for the sake of completeness. The neutral hydrogen (HI) in the intergalactic medium (IGM) absorbs photons with wavelengths shorter than the Lyman-α line, $\lambda_{\rm Ly\alpha} = 1216\,$Å. Even very small densities of HI ($X_{\rm HI} \sim 10^{-4}$) can result in the complete absorption of the UV part of the galaxy spectrum (Fan et al., 2006) and create the so-called *Gunn-Peterson Trough* (Gunn & Peterson, 1965). These galaxies are named *Lyman-break* galaxies (Steidel et al., 1996b; Dunlop, 2013). The dropout technique (Steidel et al., 1996a; Steidel, 1996) is then simply the search for objects that are detected by redder filters, which are not detected in bluer filters, so-called *dropouts*. The redshift is determined by the first (bluest) filter in which the galaxy can be detected, in which the

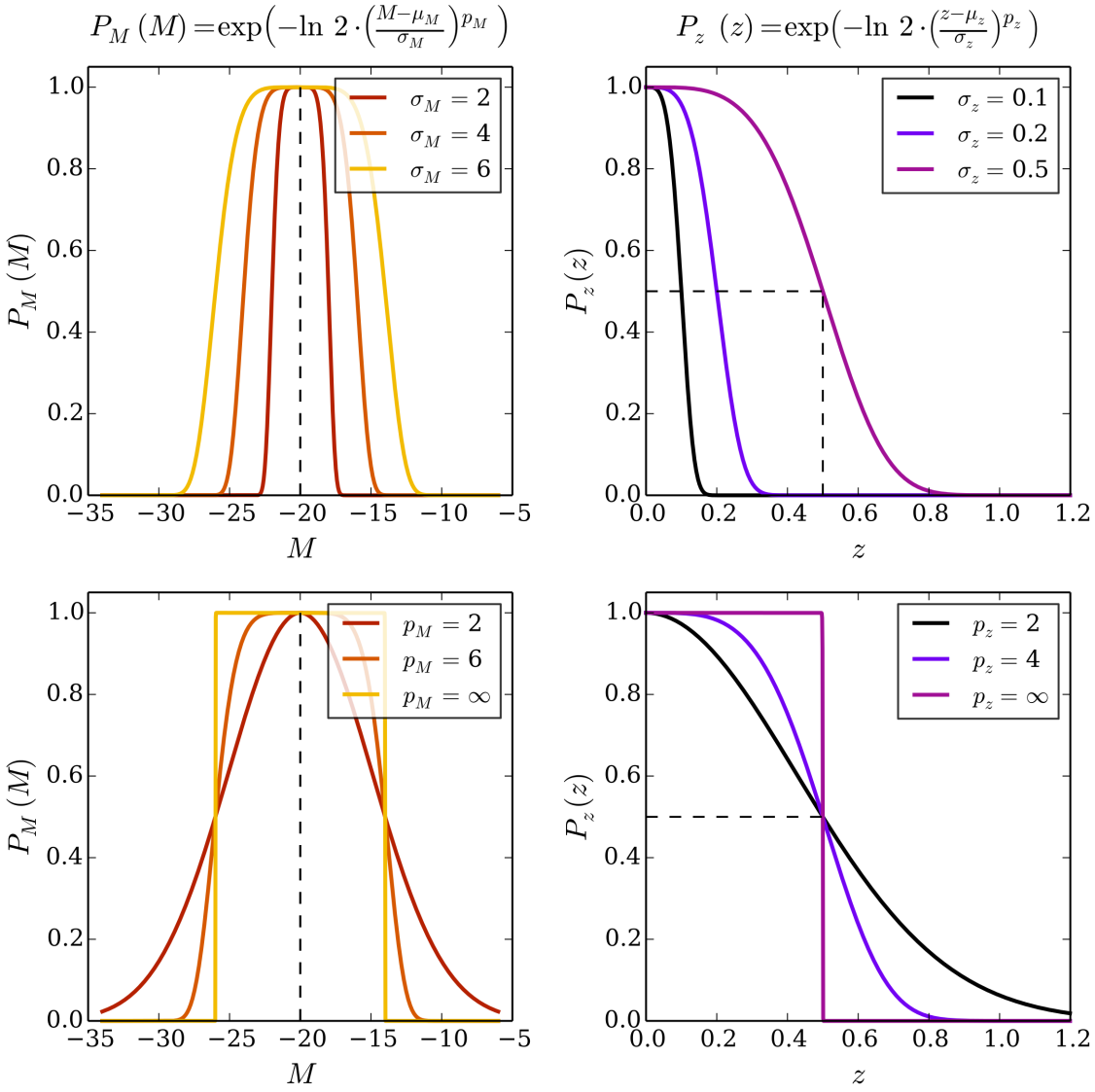

Figure 2.2: Prior probability distributions for absolute magnitude M and redshift z. The probabilities follow $P(x) \propto \exp\left(-\ln(2) \cdot \left(\frac{x-\mu_x}{\sigma_x}\right)^{p_x}\right)$, for $x = z, M$. The left panels show the luminosity prior P_M with $\mu_M = -20$ for different σ_M (upper panel) and p_M (lower panel). The upper right panel shows the redshift prior for varying σ_z, while p_z varies in the lower panel, both with $\mu_z = 0$. The dashed black lines in the left panels denote the peak of the distributions, and $P_z(x) = 0.5$ in the right panels.

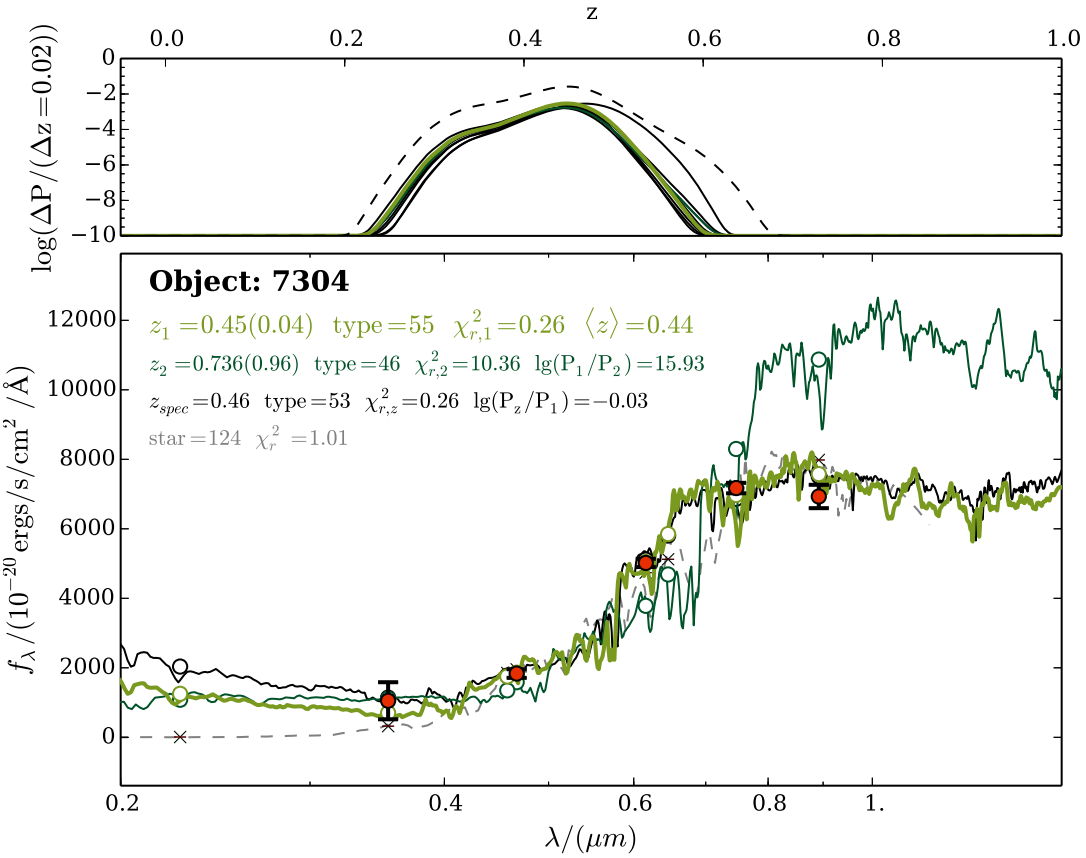

Figure 2.3: Output of the `PhotoZ` code on the example of object 7304 (SDSS object ID 587725505018986933, see also Fig. 2.1). *Upper panel:* The probability distributions $P(z)$ of the eight best fitting model SEDs are shown as functions of redshift by solid black lines, where the PDF of the best fitting model is highlighted in light green as is its SED in the lower panel. The stacked PDF $P_{\rm stack}(z)$ of all models in the model set is represented by a dashed line. *Lower panel:* Red points show the data with error bars in the SDSS filter system u, g, r, i, and z. The solid black line is the best fitting SED when models are fitted to a known spectroscopic redshift. The best fitting SED for a photo-z fit is highlighted in light green, as well as the fluxes predicted by the model which are shown by circles. The second best fitting SED is colored in dark green, whereas the best fitting stellar SED is represented by a dashed gray line. In the upper left corner we print the resulting photo-z values (with the estimated error in brackets), reduced χ^2, and model numbers (tagged *type*). $\langle z\rangle$ is the median redshift where $\int_{-\infty}^{\langle z\rangle} P_{stack}(z)\,dz = \frac{1}{2}$. Also, the relation between the probabilities of the fit and the best fitting model are given.

Lyman-α line is located.
As mentioned above, breaks in the spectra of galaxies can be degenerate in the sense that they can yield similar colors for different galaxy types and redshifts. The dropout technique can suffer from the same degeneracies, since the Lyman-α break can be confused with the 4000 Å break if the spectral flux bluewards of the 4000 Å break is below the detection limit (Kriek et al., 2006). The degeneracies can be lifted if colors redwards of the break feature are taken into account which give information about the precise nature of the break.

Chapter 3

The Sloan Digital Sky Survey Data

In the predominant part of this work we use photometric and spectroscopic data from the *Sloan Digital Sky Survey* (SDSS; York et al., 2000). SDSS has been carrying out a number of different projects, and we will concentrate on galaxy observations, namely the Legacy survey and the *Baryon Oscillation Spectroscopic Survey* (BOSS).

The Sloan Foundation telescope (Gunn et al., 2006) is situated at the Apache Point Observatory in New Mexico, located in the northern hemisphere at latitude 32° 46' 49.30" N, longitude 105° 49' 13.50" W, and elevation 2788 m. It has a 2.5 m hyperbolic primary mirror and 1.3 m secondary mirror in a Ritchey-Chrétien configuration and is mounted altazimuthally, typical for telescopes with attached heavy instruments. First light of the telescope was in 1998, it began main operations in 2000 and has been operating ever since. The 120-megapixel camera (Gunn et al., 1998) is made up by six columns and five rows, one for every filter, of CCD chips with a resolution of 2048×2048 pixels each. Observations are carried out in a drift-scan mode where the camera is moved in such a way that an object travels along a row of CCDs (and therefore the five filters) while the CCDs are being read out. The movement speed is such that every observed object has an effective exposure time of 54 seconds in each filter.

The SDSS filter curves are presented in Figure 3.1, which shows the filter throughputs with applied CCD sensitivity and atmospheric extinction. We also plot additional filters which we will use in this work in Figure 3.1. These are two filters from the *Galaxy Evolution Explorer* (GALEX, Martin et al., 2005) in the far UV (FUV) and the near UV (NUV), the Johnson-Morgan filter B (Johnson & Morgan, 1953), and the Cousins filter R. The SDSS system consists of five filters, u, g, r, i, z, with central wavelengths of 3551, 4686, 6166, 7480, and 8932 Å (Fukugita et al., 1996; Smith et al., 2002). They cover the visible part of the spectrum as well as near ultraviolet (NUV, i.e., u band) to near infrared (NIR, i.e., z band) wavelengths, with a total range of 2980 to 11,230 Å. We give an exemplary image of an elliptical galaxy observed through the SDSS filters in Chapter 4.1.1, Figure 4.3.

Additionally to the imaging camera, there are two spectrographs mounted on the telescope with a blue and a red channel each allowing to observe a wavelength range of 3800 to 9200 Å with a spectral resolution $\lambda/\Delta\lambda$ of 1850 to 2200. Both spectrographs are equipped with two CCDs with 2048×2048 pixels each, and objects are observed through 3" diameter fibers. Under good conditions typically three 15 minute exposures are taken from each object. Figure 3.2

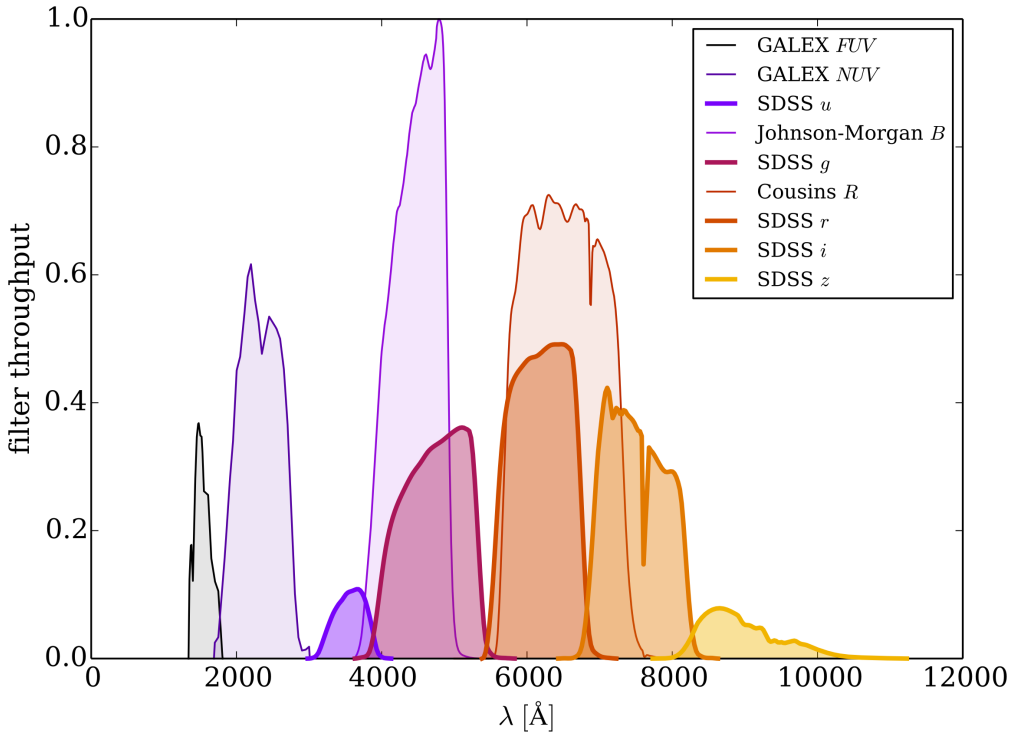

Figure 3.1: Throughput of filters used in this work. SDSS *ugriz* filters with applied atmospheric extinction and CCD sensitivity are highlighted by thick lines in the plot. Additionally, we often convolve our models with other filters to calculate the fluxes we would get if we observed the galaxy through them. These are shown by thinner lines. One is the Johnson-Morgan filter B (Johnson & Morgan, 1953), and the second the Cousins filter R. For the analyses of the higher energy part of model spectra we also include the GALEX NUV and FUV filters.

shows the spectrum of our exemplary galaxy 7304 (which we also plotted in Figure 2.1) as created by the SDSS Finding Chart[1]. Further information on SDSS is provided on the SDSS webpages[2] and in York et al. (2000). A detailed description of the camera can be found in Gunn et al. (1998).

We mentioned previously the multiplicity of projects carried out by SDSS. SDSS was initially planned for six years from 2000 to 2006 (SDSS-I), and was extended until 2008 (SDSS-II) due to the success in the earlier stage. The surveys were concluded with the final data release 7 (DR7, Abazajian et al., 2009) of SDSS-II (which is why we often use the term SDSS-II interchangeably for SDSS-I/II). At this time the Sloan Legacy survey had taken photometric and spectroscopic data of 930,000 galaxies, 120,000 quasars, and 225,000 stars, detected in 8,400 square degrees of the sky, which is about one quarter of the total area. An image of the imaging and spectroscopic footprint of the SDSS-II DR7 is provided in Figure 3.3.

[1]SDSS-I/II: `http://skyserver.sdss2.org/dr7/en/tools/explore/obj.asp`
SDSS-III: `http://skyserver.sdss3.org/dr10/en/tools/chart/chartinfo.aspx`
[2]SDSS-I/II: `http://www.sdss2.org/`, SDSS-III: `http://www.sdss3.org/`

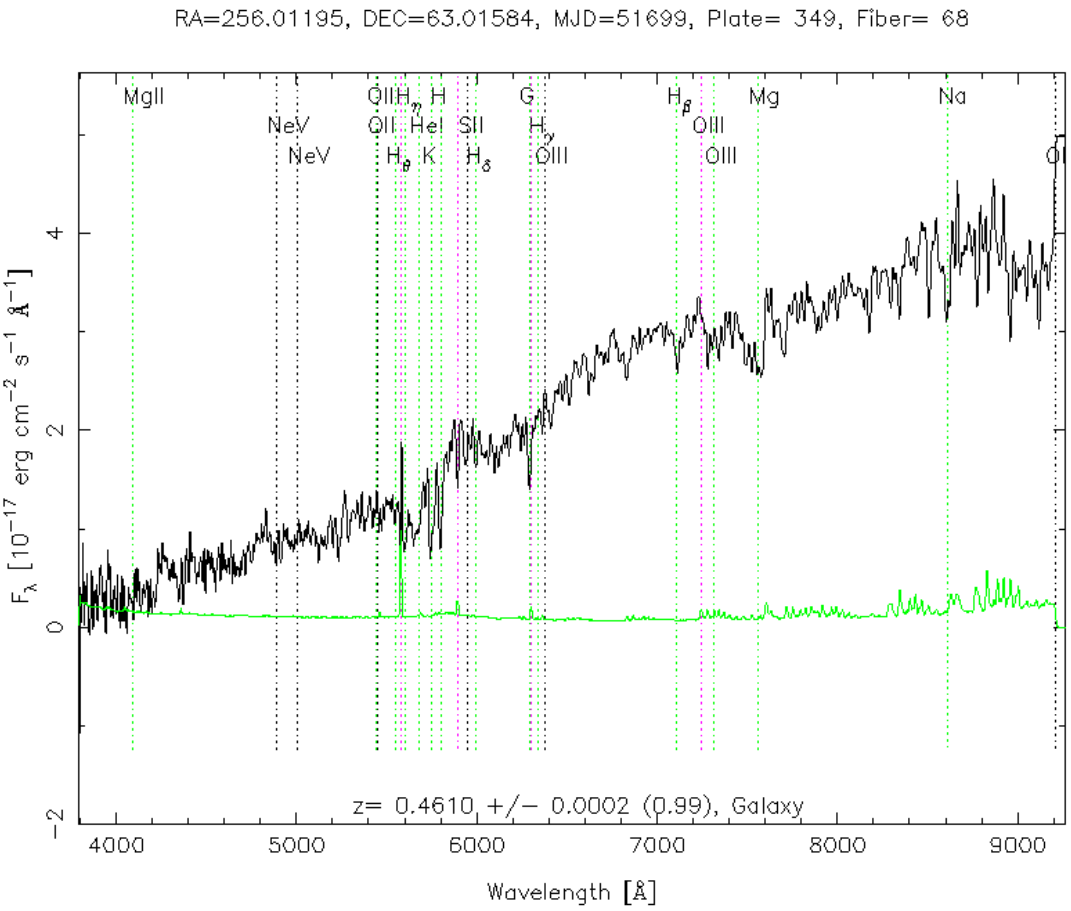

Figure 3.2: Spectrum of Object 7304 (SDSS `objID` 587725505018986933) as provided by the SDSS website. The spectral absorption and emission lines are highlighted in the spectrum by dashed vertical lines. The spectroscopic redshift is $z_{\mathrm{spec}} \approx 0.46$, determined with a confidence of 99 %. At ~ 5800 Å one can see the 4000 Å break feature (near the K, H lines) which is shifted to a higher wavelength by $1 + z_{\mathrm{spec}} \approx 1.46$. The green line at $f_\lambda \sim 0$ is the residual of the spectrum to the continuum and shows the spectral lines. The spectrum can be accessed via the SDSS-II SkyServer website's Finding Chart tool[3].

After 2008, funding for subsequent surveys was accepted, among them the *Baryon Oscillation Spectroscopic Survey* (BOSS, Eisenstein et al., 2011; Dawson et al., 2013) from which we will also use data in this work. The latest public data release of BOSS (DR10, Ahn et al., 2014) targets galaxies within 14,555 square degrees of the sky obtaining spectra for 1,848,851 luminous red galaxies. We show the SDSS-III DR10 footprint in Figure 3.4. The two spectrographs used in BOSS are rebuilt from the original SDSS spectrographs and cover a wavelength range of 3600 to 10,400 Å at a resolution of 1560 to 2650 Å (Smee et al., 2013). The fiber diameter is 2″ instead of 3″ in SDSS-II with 1000 fibers per plate instead of 640, CCDs are more sensitive in the blue and the red channels and have 4K×4K pixels.

[3]image credit: `http://skyserver.sdss2.org/dr7/en/tools/explore/obj.asp?id=587725505018986933`

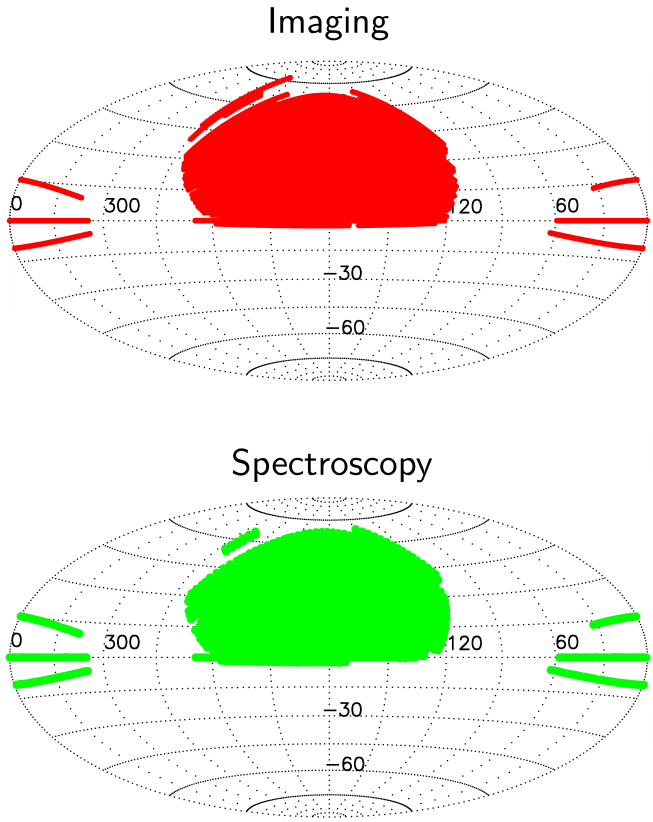

Figure 3.3: SDSS DR7 footprint for the imaging (red, upper panel) and spectroscopic (green, lower panel) part of the Legacy survey[4].

In the following (Sec. 3.1) we will introduce the selection criteria of the luminous red galaxy sample of SDSS-II. After that, we will discuss the BOSS LOWZ and CMASS samples in Section 3.2.

3.1 The SDSS-II LRG Sample

In Section 5 we will create model SEDs on the basis of the SDSS-II *luminous red galaxy* (LRG) sample to use with photometric redshift template fitting techniques. Therefore, we present here the selection criteria used to select LRGs from the SDSS galaxies. We already mentioned in Chapter 2 the prominent feature at 4000 Å (the 4000 Å break, cf. Fig. 2.1), below which fluxes have lower values than for wavelengths greater than 4000 Å (the origins of this feature will be explained in Section 4.1.1). The LRG selection criteria are defined in such a way that galaxies exhibit prominent 4000 Å breaks, and therefore red spectra. Spectra which have higher fluxes in longer wavelength ranges compared to shorter wavelengths are dubbed *red*, and *blue* for the opposite case. Eisenstein et al. (2001) developed cuts in magnitude and color separating LRGs from the other galaxy types taking their evolution with redshift into account. Objects passing these criteria are flagged accordingly and can be obtained via the

[4]image credit: `http://www.sdss2.org/dr7/coverage/index.html`, modified

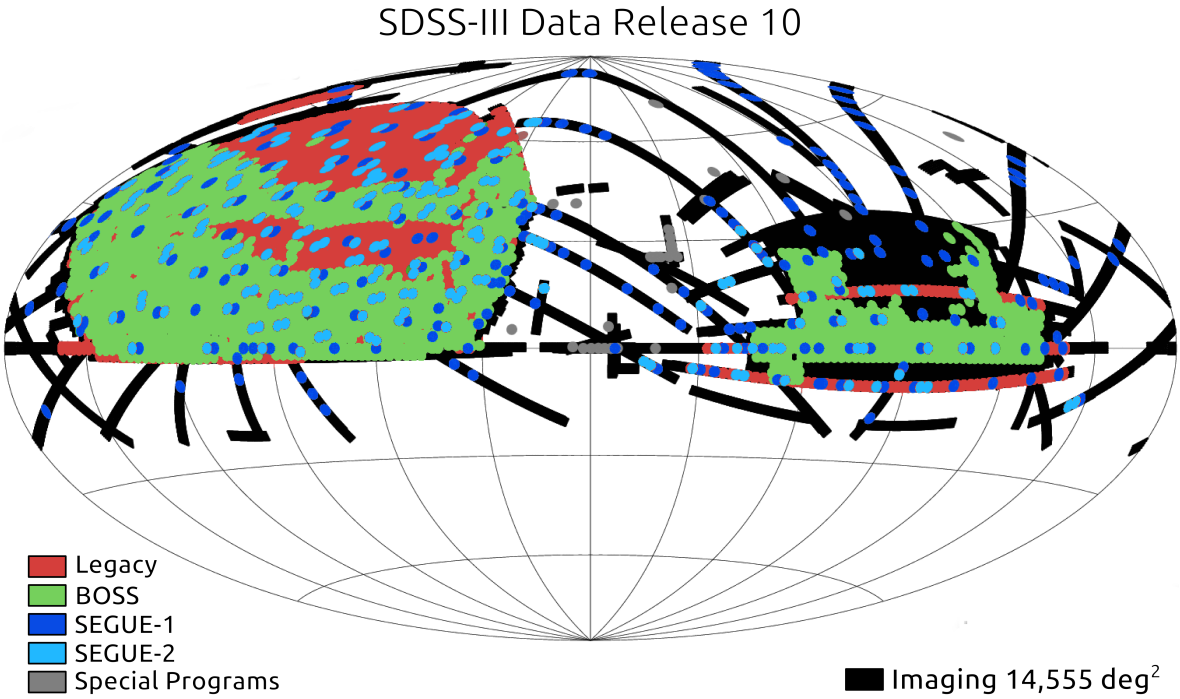

Figure 3.4: SDSS DR10 footprint[5]. The BOSS survey area is highlighted in green.

Catalog Archive Server[6] (CAS). In Appendix E.1 we list the SQL query used to create the LRG sample to enable the reader to download the same catalogs that are being used in this work.

From the resulting sample we remove objects that do not pass the cuts of Padmanabhan et al. (2005), which differ marginally from the selection criteria of Eisenstein et al. (2001). We will list them in the following, but we make three color definitions in advance (analogous to Padmanabhan et al., 2005):

$$\begin{aligned} c_\perp &\equiv (r-i) - (g-r)/4 - 0.18 \\ c_\parallel &\equiv 0.7 \cdot (g-r) + 1.2 \cdot [(r-i) - 0.18] \\ d_\perp &\equiv (r-i) + (g-r)/8 \approx r - i. \end{aligned} \tag{3.1}$$

The definitions are chosen in such a way that they track a passively evolving stellar population with redshift, i.e., solely an aging stellar population without star formation after the birth of the galaxy (cf. Section 4.2). Furthermore, the color definitions ensure that $c_\parallel = g - r$ if $c_\perp = 0$.

The colors $c_\parallel$, $c_\perp$, and $d_\perp$ are shown as functions of redshift in the lower panel of Figure 3.5, which is a reproduction of Figure 2 in Padmanabhan et al. (2005). In the upper panel we plot the colors of the LRG model by Maraston et al. (2009) when projected through the SDSS filters. We set the formation redshift to $z_f = 3.64$ (corresponding to an age of 12 Gyr at $z = 0$) and assume solely passive evolution of the stellar population, i.e., no further star

[5] image credit: `http://www.sdss3.org/dr10/`, modified
[6] `http://cas.sdss.org/astrodr7/en/`

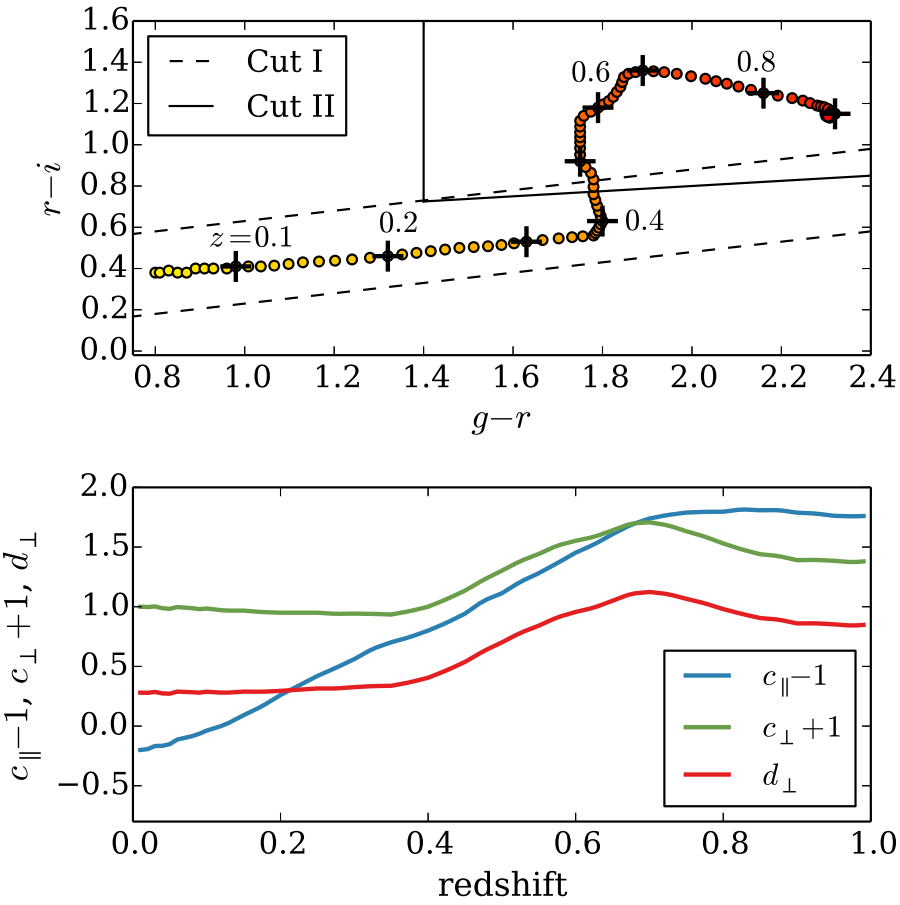

Figure 3.5: This figure is reproduced after Figure 2 in Padmanabhan et al. (2005) *Upper panel:* In yellow to red dots the $g-r$ versus $r-i$ colors of the passively evolving LRG model by Maraston et al. (2009) with formation redshift of $z_f = 3.64$ (corresponding to an age of 12 Gyr at $z = 0$) is shown as a function of redshift. Redshift steps of width 0.1 are indicated by crosses and labeled accordingly. The dashed line shows *CutI*, whereas the solid line is *CutII*. *Lower panel:* $c_\parallel$ (blue), $c_\perp$ (green), and $d_\perp$ (red) of the model are plotted as a function of redshift.

formation. We can observe a turning point in the color track at about $z \sim 0.4$ which is where the 4000 Å break moves from the g into the r filter. This is the reason why Eisenstein et al. (2001) developed two different color cuts, one for a lower redshift sample (*CutI* for $z \lesssim 0.4$) and one for higher z (*CutII* for $z \gtrsim 0.4$). Intuitively, one could suggest making a cut in the $u-g$ versus $g-r$ space to isolate early-type galaxies at lower z from late types at higher redshifts. Unfortunately, the u band magnitude is very close to the detection limit of SDSS ($u \approx 22$) for galaxies of interest ($r \approx 19$). Therefore, one would use the well-measured $r-i$ color to resolve the problem, but there is an accidental degeneracy in colors at $z \lesssim 0.4$ due to the SDSS filters (see Eisenstein et al., 2001, and discussions therein). However, one can bypass this issue by making use of the bivariate color-magnitude relation of galaxies and impose additional magnitude-color cuts (see Equation 3.3 in the following).

Padmanabhan et al. (2005) propose the following cuts for the lower redshift sample (*CutI*):

$$|c_\perp| < 0.2 \tag{3.2}$$
$$r_{\mathrm{Petro}} < 13.6 + c_\parallel/0.3 \tag{3.3}$$
$$r_{\mathrm{Petro}} < 19.7, \tag{3.4}$$

where r_{Petro} is the Petrosian magnitude (Petrosian, 1976) in the r band explained in Appendix E.1. It is used for consistency with the SDSS target selection. The cut in $c_\perp$ ensures that we are focusing on the low redshift regime (shown by dashed lines in the upper panel of Figure 3.5). The flux cut in Equation 3.3 is mainly to ensure good spectroscopic performance. For the higher redshift portion of the LRG sample Padmanabhan et al. (2005) use (*CutII*):

$$d_\perp < 0.55 \tag{3.5}$$
$$g - r > 0.4 \tag{3.6}$$
$$i < 18.3 + 2d_\perp \tag{3.7}$$
$$i < 20. \tag{3.8}$$

The first two cuts are shown in the upper panel of Figure 3.5 by solid lines, where the second of the two ensures a separation from the stellar locus. In this redshift regime, the 4000 Å break moves through the r band filter, which implies that the K-correction (the correction in magnitude of a galaxy in the rest frame) to the r band magnitude is sensitive to redshift. Since the i band magnitude is not sensible to redshift, a selection in i is more robust (Eqs. 3.7 and 3.8). All selection cuts are imposed on the SDSS extinction corrected (Schlegel et al., 1998) model magnitudes, extracted from the better fit between an exponential and a de Vaucouleurs profile to the surface brightness of the galaxy image (see Sec. 4.1 and App. E.1).
For a reasonable evaluation of the photometric redshift quality later on, we require that the spectroscopic redshift errors fulfill $\Delta z_{\mathrm{spec}} \leq 0.002$, the resolution of the photometric redshift estimations in Chapter 5.3, and that the spectroscopic redshifts were determined with high confidence ($> 99\%$). The resulting LRG sample contains approximately 140,000 objects.

Figure 3.6 presents the colors of the LRG sample and their median as a function of redshift. The lower panels of the four major panels also show the colors, but with the median colors as offsets. From the slope in the median $g - r$ and $r - i$ colors, we can observe how the 4000 Å break moves from the g into the r band filter at $z \sim 0.4$. Furthermore, $g - r$ can be considered as a bijective function of redshift up to $z \sim 0.4$. The same holds for the $r - i$ color throughout the considered redshift range with an increase in steepness for $z \gtrsim 0.4$. From this we can infer that these two colors are a good redshift indicator (at least for LRGs) and that photo-zs can be determined to a high precision. We will confirm that this is true in Section 5.3, and that the photo-z accuracy of red galaxies outperforms that of blue galaxies in Chapter 7.

3.2 The BOSS LOWZ and CMASS Samples

BOSS was designed to measure the spatial distribution of LRGs and quasars to investigate the baryon acoustic oscillations that are imprinted on the large scale structure of today's Universe (cf. Chap. 1). We already mentioned that BOSS is a spectroscopic survey that observes luminous red galaxies out to an approximate redshift of $z \lesssim 1.0$. Figure 3.7 shows the spectroscopic redshift distribution of the BOSS catalog used in this work by a gray histogram. In

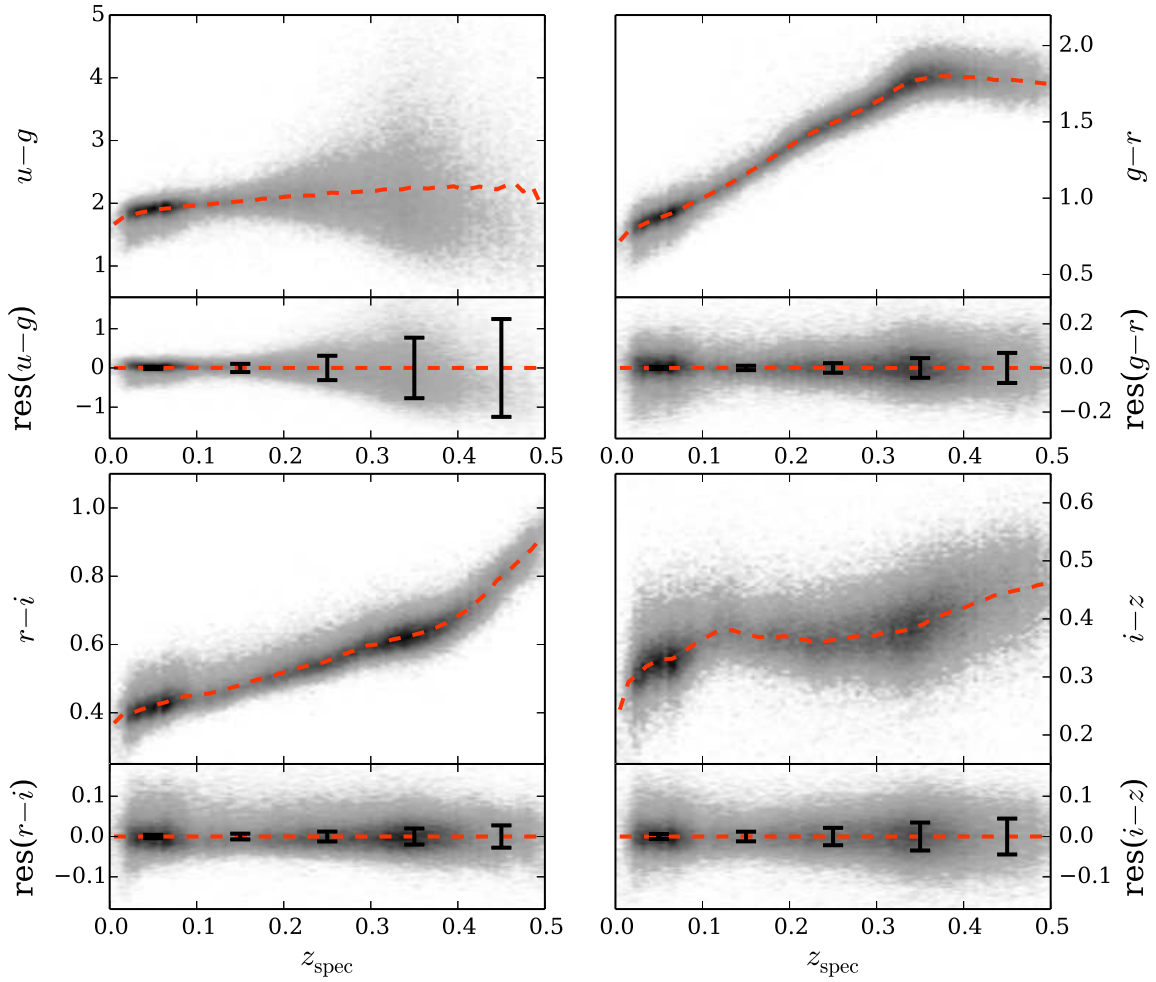

Figure 3.6: The upper parts of the four major panels show $z_{\rm spec}$ versus the $u-g$, $g-r$, $r-i$, and $i-z$ colors of SDSS-II LRG Sample. The median color as a function of redshift is highlighted by red dashed lines. In the lower panels we plot the residual color from its median (therefore the median color is the $y=0$ line). Error bars denote the median color errors in five redshift bins.

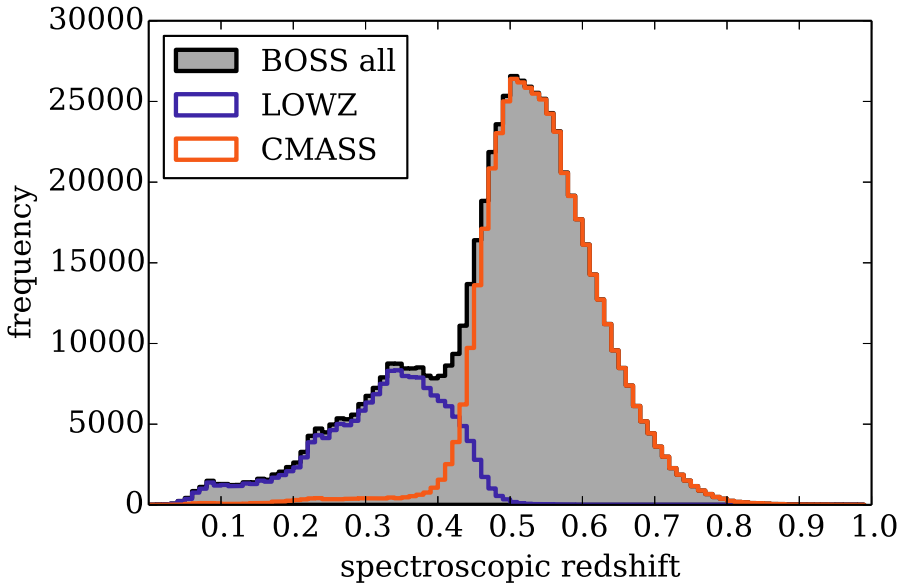

Figure 3.7: Spectroscopic redshift distribution of the BOSS catalog used in this work. The LOWZ sample is highlighted by a blue line, whereas the CMASS sample is plotted in orange. The total BOSS sample is plotted in gray.

Appendix E.2 we present the SQL query used to generate the catalog from the SDSS database.

BOSS target galaxies are selected from SDSS-II imaging data in such a way that they have high luminosities and masses. Furthermore, the BOSS target selection requires approximate uniform stellar masses throughout the redshift range of $0.2 \lesssim z \lesssim 0.6$ (Eisenstein et al., 2011). Above $z \sim 0.6$, the BOSS sample is magnitude limited. Similar to the cuts of the SDSS-II LRG sample (see Section 3.1, Eisenstein et al., 2001; Padmanabhan et al., 2005), a number of magnitude and color cuts are applied to ensure the above requirements are fulfilled. As with the SDSS-II sample, the BOSS sample is selected in two ways yielding two subsamples, a lower redshift sample tagged *LOWZ* for $z \lesssim 0.4$, and a higher redshift sample, dubbed *CMASS* for the constant mass requirement (Eisenstein et al., 2011).
For the selection cuts the colors $c_\perp$, $c_\parallel$, and $d_\perp$ are defined in the same way as in Section 3.1, Equation 3.1. The only difference is that Eisenstein et al. (2011) use composite model (*cmodel*, Abazajian et al., 2004, cf. Appendix E.2) magnitudes instead of model magnitudes for the calculation of colors, and model magnitudes in magnitude cuts. Where the SDSS model magnitudes are extracted from a better fit of a de Vaucouleurs and an exponential profile to the surface brightness of a galaxy, the cmodel magnitudes are derived using the best fitting combination of both. The selection cuts in colors for both samples are the same as for the LRGs (see Section 3.1, and Figure 3.5), only the magnitude cuts differ. The LOWZ sample is selected by:

$$|c_\perp| < 0.2 \tag{3.9}$$

$$r < 13.6 + c_\parallel/0.3 \tag{3.10}$$

$$16 < r < 19.5. \tag{3.11}$$

Note, that the magnitude cuts are not anymore made with Petrosian magnitudes (as in Sec. 3.1) but model magnitudes. The CMASS sample is defined by:

$$|d_\perp| > 0.55 \tag{3.12}$$

$$i < 19.86 + 1.6 \cdot (d_\perp - 0.8) \tag{3.13}$$

$$17.5 < i < 19.9. \tag{3.14}$$

Some of the galaxies passing these cuts were already spectroscopically observed by SDSS-II, and are not re-observed in BOSS if the spectroscopic redshifts were already measured to a high accuracy in SDSS-II. For further information about the BOSS target selection we refer to Eisenstein et al. (2011) and the DR10 target selection website[7].
Our BOSS catalog comprises 659, 532 galaxies (see Appendix E.2 for the SQL query and further requirements). The redshift distributions of the LOWZ and CMASS samples are shown in Figure 3.7 together with the distribution of the whole BOSS sample.

[7] `http://www.sdss3.org/dr10/algorithms/boss_galaxy_ts.php`

Chapter 4

Model Spectral Energy Distributions

In this chapter we investigate the properties of galaxies with emphasis on their spectra. In Section 4.1 we will give an introduction to galaxies in general and on elliptical galaxies in detail, since we will mainly focus on them in this work. We will particularly address their spectra which hold important information about the galaxy properties, e.g., star formation history and chemical abundances. But spectra also comprise clues for the estimation of photometric redshifts, specifically with template fitting methods. In Section 4.2 we will focus on the star formation history of galaxies and the creation of synthetic model spectra. We will also analyze a number of public empirical, semi-empirical, and synthetic model SEDs in terms of their colors, and analyze how well they can represent the SDSS-II LRGs.

4.1 Galaxies

It was in the middle of the 18th century when Kant and Wright first suggested that the Milky Way is a finite-sized accumulation of stars. Kant raised the notion that the observed diffuse objects, termed "elliptical nebulae" at that time, could also be limited conglomerations of stars very like the Milky Way. There was a long ongoing debate about the potential extragalactic nature of those "elliptical nebulae", which was finally resolved by Edwin Hubble in 1923 who measured the distance of the Andromeda galaxy (M31) using Cepheid variable stars. Shortly afterwards Hubble suggested a classification of galaxies in terms of their morphology which is still used today (the *Hubble Tuning Fork*, Hubble, 1926, see Figure 4.1). The left side of the fork comprises elliptical galaxies. They exhibit spheroidal shapes on the very left side to more skewed elliptical shapes on the right. Elliptical shaped galaxies have the designation En, where n denotes 10 times the ellipticity ϵ of their isophotes, i.e., $n = 10 \cdot \epsilon$ with $\epsilon \equiv 1 - \beta/\alpha$, α and β being the apparent major and minor axes of the ellipse when projected onto the sky. Spiral galaxies carry the designation S and SB for those exhibiting a bar in the galaxy center. Furthermore, there is a transitional class between the ellipticals and spirals, the *S0* or *SB0* lenticular galaxies, also having the B suffix when exhibiting a bar. Lastly, there are irregular galaxies (Irr), who show non-regular features in their morphology that do not fall into the other two classifications. The Hubble Tuning Fork is only a rough classification, and there are various galaxy subclasses which could enter the diagram and several other attempts to classify galaxies (e.g., Morgan, 1958, from which only the designation of the diffuse supergiant cD ellipticals is still used today) that differ from Hubble's Tuning Fork.

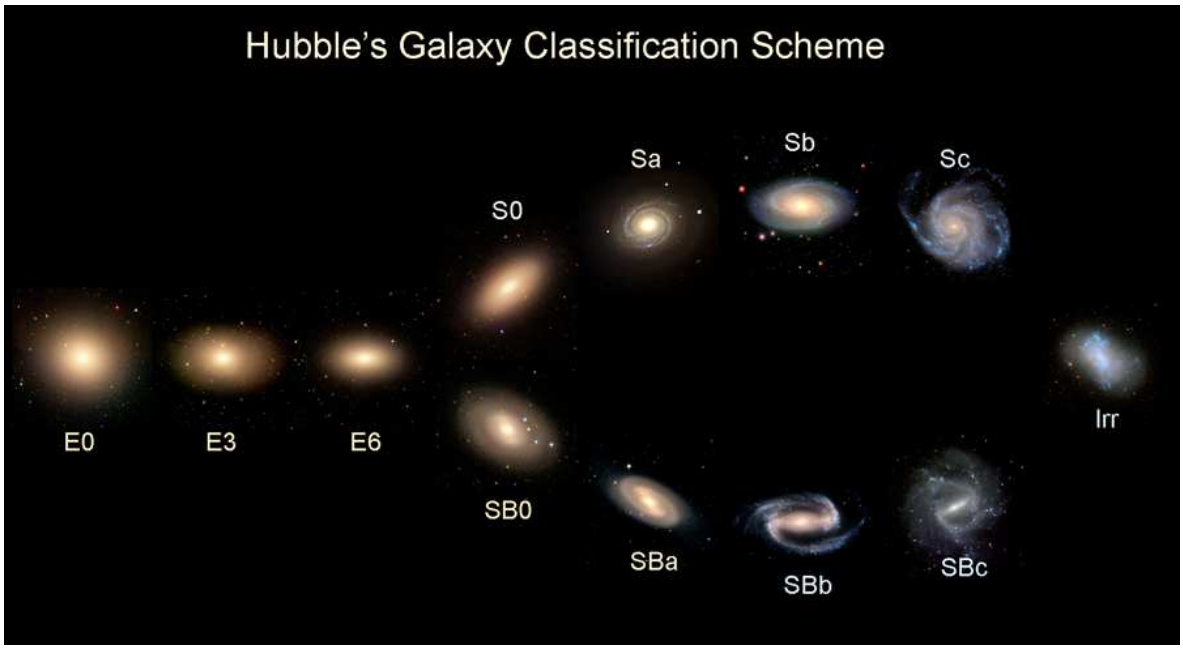

Figure 4.1: Galaxy classification in the Hubble Tuning Fork[1]. Galaxies on the left side of the fork are called *early-type*, whereas galaxies with a morphology from the right part of the fork are called *late-type* galaxies. These designations are a historical artifact and not descriptive of the age of the galaxy (see text for an explanation). Irregular galaxies (*Irr*) do not fit in the classification scheme, but are appended on the right side of the fork.

Hubble falsely thought of the Tuning Fork as an evolutionary path, wherefore elliptical and lenticular galaxies are often called *early-type* galaxies (ETGs), and spirals *late-type*. This has been proven wrong, and recent studies found that ellipticals are likely a result of major mergers of late-types and also early-type galaxies (e.g., Toomre & Toomre, 1972; Barnes & Hernquist, 1996; Naab & Burkert, 2003; Bell et al., 2004). There are further shortcomings of the Hubble diagram, such as the subjective classification: A galaxy viewed edge-on, can look very different from the face-on view. Furthermore, the morphology depends much on the wavelength range covered by the filter through which it is observed. Despite all this, the classes still correlate with physical properties of the galaxies such as star formation, luminosity, color, and masses (Roberts & Haynes, 1994), which is why the tuning fork is still being used today.
In the following, we will describe selected properties of elliptical galaxies which will be in the main focus of this thesis.

4.1.1 Elliptical Galaxies

Until the late 1970s elliptical galaxies were seen as simple systems, i.e., rotationally flattened ellipsoids without a disk, and with cores of constant density. They were also believed to be devoid of gas and dust. However, nowadays we know that ellipticals are in no way simple, but exhibit complicated stellar motion, core features, and can be divided into an array of subclasses that do not depend on their ellipticity.

[1]Image taken from the Galaxy Zoo Blog Website `http://blog.galaxyzoo.org/2010/05/12/types-of-galaxies/`. Image credit: Zookeeper Rob

The Surface Brightness Profile

The surface brightness I is a measure of luminosity per unit area and gives information about the distribution of stars in a galaxy. As a function of radius $I(r)$ the surface brightness of ellipticals can be described (when a more detailed model of $I(r)$ is not required) by the empirical $r^{1/4}$-law or *de Vaucouleurs Profile* (de Vaucouleurs, 1948, 1953):

$$I(r) = I_e \cdot \exp\left[-7.67\left((r/r_e)^{1/4} - 1\right)\right]. \tag{4.1}$$

r_e denotes the so-called *effective radius*, within which one half of the total light is enclosed under the assumption that the galaxy is circularly symmetric and the profile is valid at all radii (Caon et al., 1993). I_e is then the surface brightness at r_e. Although the de Vaucouleurs profile fits the majority of elliptical galaxies remarkably well, there are galaxies whose surface brightness profiles deviate significantly from the $r^{1/4}$ law (e.g., Capaccioli et al., 1988, Fig. 2). Today, the surface brightness profile of ellipticals is most often described by a generalized version of the de Vaucouleurs profile, the *Sérsic profile* (Sérsic, 1963, 1968):

$$I(r) = I_e \cdot \exp\left[b_n\left((r/r_e)^{1/n} - 1\right)\right]. \tag{4.2}$$

The constant b_n is defined in terms of the model parameter n, which describes the concentration of the stellar content, and therefore varies for different kinds of ellipticals. I.e., galaxies with low luminosities are better fitted by a Sérsic profile with a concentration index lower than 4, whereas $n > 4$ provides a better fit to high-luminosity ellipticals (e.g., Schombert, 1986; Caon et al., 1993; Binney & Merrifield, 1998; Kormendy et al., 2009; Bender et al., 2014).

Parameter Correlations

Elliptical galaxies show a number of parameter correlations which distinguishes them in comparison to later types. We want to discuss a few in the following.
There is a relation between the central radial velocity dispersions of stars in early-type galaxies and their luminosity L. We assume that the mass-to-light ratio M/L and the average surface brightness are constant. Furthermore, we assume a constant mass density as a function of radius. From the Virial Theorem, $\langle U\rangle + 2\langle K\rangle = 0$, we can then derive the *Faber-Jackson Relation* (empirically found by Faber & Jackson, 1976):

$$L \propto \sigma_0^4, \tag{4.3}$$

where σ_0 is the central value for the radial velocity dispersion.
Given the assumptions made in the derivation, it is not surprising that several authors found that the relation is not exactly true. I.e., for luminous ellipticals the exponent was derived to be ~ 5 (e.g., Schechter, 1980; Malumuth & Kirshner, 1981), whereas studies of their faint counterparts yield a value lower than 4 (e.g., Tonry, 1981).

Removing the assumption of the constant surface brightness from the above derivation of the Faber-Jackson relation, we get a correlation between the effective radius r_e, the central velocity dispersion σ_0 and the luminosity L (or effective surface brightness I_e) spanning a two-dimensional surface in the three-dimensional parameter space:

$$L \propto \sigma_0^\alpha\, r_e^\beta \tag{4.4}$$

$$r_e \propto \sigma_0^\gamma\, I_e^\delta, \tag{4.5}$$

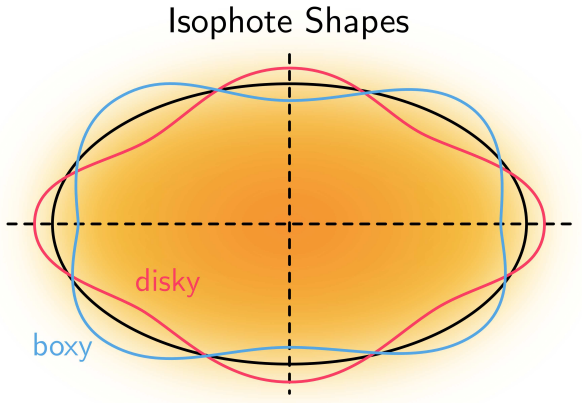

Figure 4.2: Schematic drawing of disky (red) and boxy (blue) isophotes of an elliptical galaxy (recreated after Bender et al., 1988, Fig. 5).

written in general form with variable exponents. This relation is called the *fundamental plane* of elliptical galaxies, with $\alpha, \beta \simeq 3, -0.7$ (Dressler et al., 1987; Djorgovski & Davis, 1987). Dressler et al. (1987) defined the photometric parameter D_n, the aperture which encloses a mean surface brightness of $20.75\,\mathrm{mag/arcsec^2}$ in the B band (cf. Fig. 3.1). In general, D_n will depend exponentially on r_e and I_e, and the relation of the fundamental plane can thereby be rewritten as $D_n \propto \sigma_0^\epsilon I_e^\zeta$, with $\epsilon, \zeta \simeq 1.4, 0.07$. Hence, the fundamental plane relation is sometimes also called the $D_n - \sigma$ *relation*.

Lastly we want to mention the tight correlations between the Mg_2 line strengths and $B-V$ colors (where the V band is centered in the visual at $\sim 5510\,\text{Å}$), and central velocity dispersions σ_0 reported by Bender et al. (1993). This is rather surprising since the Mg_2 indices were measured in the galaxy centers, whereas $B-V$ was observed to much larger radii. The data was taken from ellipticals spanning a large variety of masses and surface brightnesses.

The Shape of Elliptical Galaxies

Although nearly elliptical in shape, the isophotes of elliptical galaxies exhibit slight variations from true ellipses. Expanding the deviations of the observed isophotes from perfect ellipses $\delta(\vartheta)$ in a Fourier series (e.g., Bender & Moellenhoff, 1987) yields

$$\boldsymbol{\delta}(\vartheta) = \left(1 + \frac{a_4\,\cos(4\vartheta)}{r(\vartheta)}\right)\begin{pmatrix} a\,\cos\vartheta \\ b\,\sin\vartheta \end{pmatrix}, \tag{4.6}$$

where a and b are the (semi-)major axes of the ellipse, ϑ is the angle in $[0, 2\pi]$, and $r(\vartheta)$ is the radial distance of the isophote to the center of the ellipse. The majority of galaxies show small deviations from a perfect ellipse, i.e., $|a_4/a| \sim 0.01$. Depending on the value of $a_4 \lessgtr 0$ the isophote has a *boxy* ($a_4 < 0$) or *disky* ($a_4 > 0$) shape. We show a schematic image of disky and boxy isophotes in Figure 4.2 which is a reproduction of Fig. 5 in Bender et al. (1988).

We introduce the rotation parameter $(V/\sigma)^*$,

$$\left(\frac{V}{\sigma}\right)^* \equiv \frac{(V_\mathrm{rot}/\sigma)_\mathrm{observed}}{(V_\mathrm{rot}/\sigma)_\mathrm{isotropic}}, \tag{4.7}$$

with the rotational velocity V. Bender (1988) and Kormendy & Djorgovski (1989) found that $(V/\sigma)^*$ strongly correlates with the shape of the isophotes. Disky galaxies generally have $(V/\sigma)^*$ of order unity, and are therefore in part rotationally supported. Furthermore, their mass-to-light ratios are higher on average compared to the mean M/L of ellipticals with comparable L. Galaxies with disky isophotes are also less luminous in radio and X-ray wavelengths. Boxy galaxies on the other hand show smaller values of $(V/\sigma)^*$, and can therefore not have been flattened by rotation. In fact, most boxy galaxies are not rotationally symmetric, and their shape is due to random stellar motion. Often boxy galaxies exhibit counter-rotating cores which cannot be observed in disky galaxies. They are also more radio-loud and have higher X-ray luminosities. The structural differences between disky and boxy ellipticals may be a result of different progenitors in the merging of two galaxies from which the elliptical is formed (Khochfar & Burkert, 2003).

The Content and Spectra of Elliptical Galaxies

It was initially believed that elliptical galaxies are essentially devoid of gas and dust. However, several studies found that ellipticals contain gas and also layers of dust, but at a lower fraction than spiral galaxies (e.g., Trinchieri & Fabbiano, 1985; van Gorkom, 1992; Franx et al., 1994). Gas in ellipticals is predominantly present in a very hot $\sim 10^7$ K ionized state, presumably heated by supernovae and stellar winds, which can be detected by its X-ray emission (e.g. Trinchieri & Fabbiano, 1985; Schweizer, 1987). In this hot environment stars are not able to form. Furthermore, Hα emissions from warm ($\sim 10^4$ K) HII regions are being observed, as well as cold gas ~ 100 K from the HI 21 cm and CO lines. In many ellipticals, as well as in lenticular galaxies S0, also dust can be detected, which is partially located in a dust disk. Furthermore, the central regions of elliptical galaxies are redder and more metal-rich in comparison to the outer parts (Franx et al., 1989; Davies et al., 1993), where the term "metals" refers to elements heavier than He.
The colors of ellipticals appear yellow-red in the visual range of the spectrum which is an indication of an old stellar population. The composite SDSS image of our exemplary galaxy 7403 is given in Figure 4.3, where the galaxy appears even redder than in the rest frame due to the cosmological redshift of $z = 0.46$. The spectral energy distributions of elliptical galaxies are similar to those of relatively cool stars that resemble that of black body at ~ 5000 K, but with prominent absorption features. This suggests that major star formation has not occurred within the last billion years. Elliptical galaxies have basically completed star formation after an initial star forming phase or burst, and they shine mainly with their aging stars. The spectrum is dominated by the light of red giants which have moved off the main sequence in the *Hertzsprung-Russell diagram* (HRD, Fig. 4.4). Blue O and B stars have essentially died out leading to the notable lower fluxes shortwards of 4000 Å. The absence of young stars is the reason why ellipticals were (wrongly) believed to be devoid of gas. α elements (atomic number $Z \leq 22$, O, MG, Si, Ca, ...) are likely over-abundant with respect to iron (e.g., Faber et al., 1992; Worthey et al., 1992), especially in massive ellipticals. Supernovae type 2 produce predominantly α elements but also iron on relatively short timescales, while supernovae type 1A produce significantly more iron, but on longer timescales. This difference in timescales is the reason why an over-abundance of α elements is an indicator for short star formation timescales.

The spectra of ellipticals (cf. Figure 2.1 in Chapter 2.2.2) exhibit primarily absorption

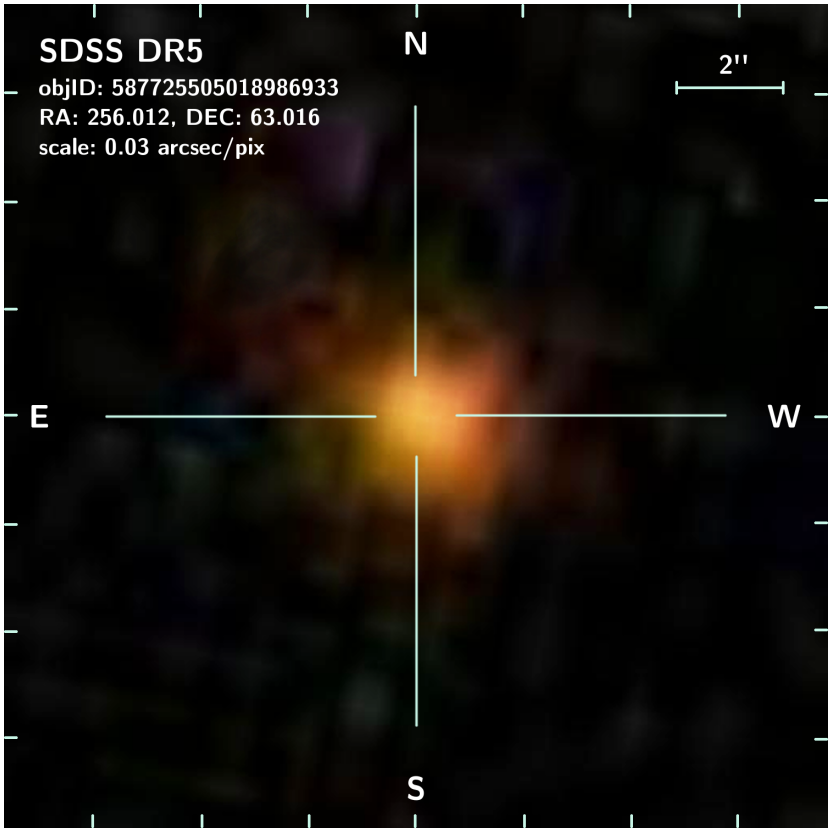

Figure 4.3: Composite image of the exemplary elliptical galaxy 7304 as observed by the SDSS in the g, r, and i filters. The image was created by the SDSS Finding Chart tool[2]. Colors were exaggerated to improve the clarity.

lines with no, or only sparse emission line features. The atoms and molecules in stellar atmospheres and the cold interstellar gas siphons off the stellar radiation at key frequencies resulting in the absorption features that are superimposed on the continuum of the black body spectrum. Conspicuous absorption lines in the spectra of elliptical galaxies are the lines of singly ionized Calcium (CaII, designated H and K by Fraunhofer) at wavelengths 3934 and 3969 Å respectively, as well as the Balmer lines of hydrogen, elicited by the changing opacity in stellar atmospheres. These, together with the lack of blue stars, are the cause for the 4000 Å break which is the most prominent feature in the spectra of elliptical galaxies and the reason why we can estimate photometric redshifts to a relatively high precision (cf. Sec. 2.2.2). The 4000 Å break tells us at a glance something about the underlying stellar population. If a galaxy spectrum shows an absence of metal lines and a strong blue component, the 4000 Å break vanishes. Furthermore, if an active galactic nucleus is present, its spectrum continues right across 4000 Å and the break is diluted by whatever fraction of light the nucleus contributes. In the presence of quasars the break can be very weak or even unmeasurable, because so little light comes from the surrounding galaxy. Elliptical galaxies show prominent 4000 Å breaks, and hence do not own a young stellar population, nor do they form stars in a significant amount. Nowadays, it is believed that elliptical galaxies form their stars early and fast.

[2] `http://skyserver.sdss2.org/dr7/en/tools/explore/obj.asp?id=587725505018986933`

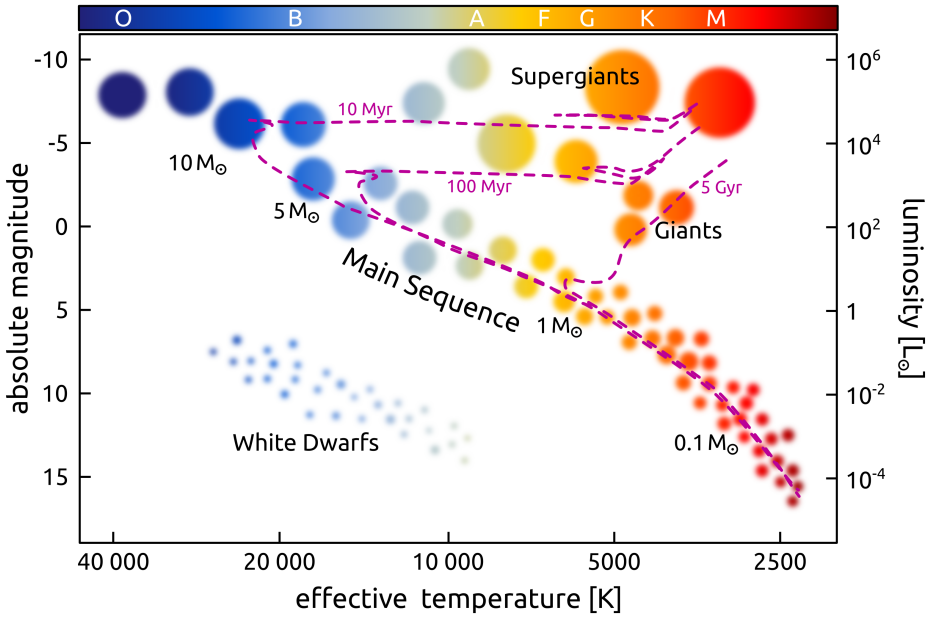

Figure 4.4: Schematic Hertzsprung-Russell Diagram. Isochrones are given by dashed lines. Own sketch.

4.2 Model Spectral Energy Distributions

In this chapter we want to present different sets of model spectral energy distributions that are publicly available. We present a number of model SEDs that are being used to estimate photo-zs. We compare their colors to those of the SDSS-II LRG data, and motivate why we need to construct new templates. Before that, we expand on Section 4.1.1 where we discussed the content and the spectra of elliptical galaxies, and give an introduction into stellar population evolution with focus on stellar population synthesis.

4.2.1 Galaxy Stellar Population Evolution

To interpret spectral data, knowledge about how the spectral features translate into the physical parameters of a galaxy and its star formation history (SFH) is required. The features emerge as result of evolution of the galaxy and its stellar content, wherefore we can derive spectra given a comprehensive library of SEDs with known properties and a theory about the evolution of stars in the HR diagram (cf. Fig. 4.4). The stellar population from any star formation history can be described by summing over various infinitesimal star bursts. Given an initial mass function (IMF) which describes the probability distribution of masses of newly formed stars, one can derive the galaxy SEDs with a stellar evolution prescription. This is a non-trivial task however as stars evolve off the main sequence. With a theoretical model about the stellar evolution as a function of time or age, the galaxy SED can then be constructed by adding spectra from an extensive stellar library in the predicted proportions. The relative hydrogen and Helium abundances of stars and galaxies can be described by their mass ratios $X = m_{\mathrm{H}}/M$ and $Y = m_{\mathrm{He}}/M$ respectively, where M is the total mass. The *metallicity*, i.e., the abundance of elements heavier than He (*"metals"*), is then $Z = 1 - X - Y$. The elemental contents of the sun are $X_\odot = 0.73$, $Y_\odot = 0.25$, and $Z_\odot = 0.02$.

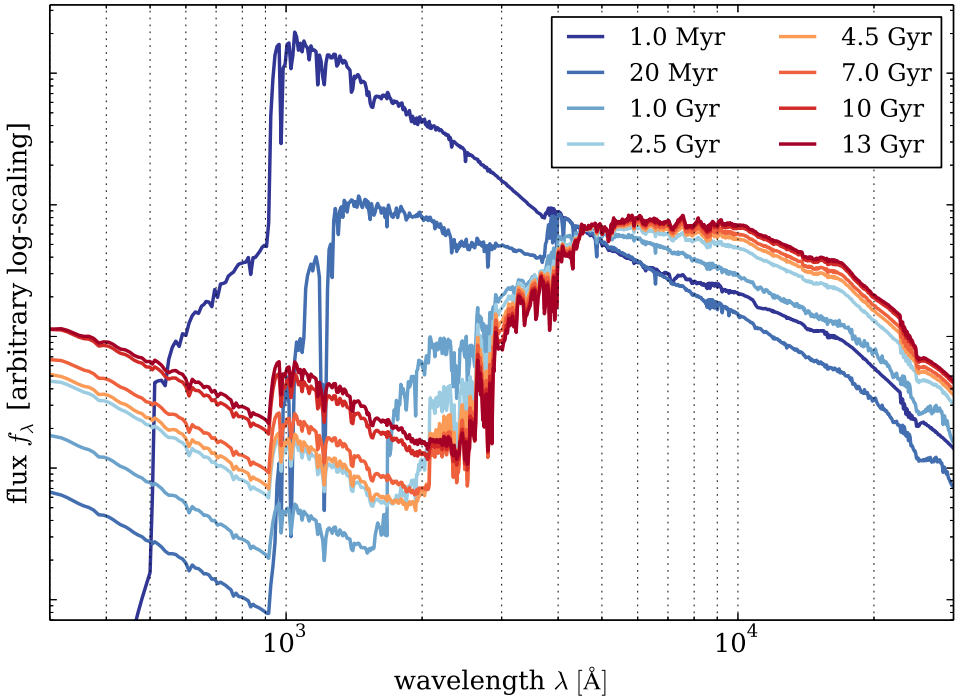

Figure 4.5: SEDs of BC03 Models with solar metallicity sampled at eight different ages. Note that the scaling is logarithmic on both axes. Ages run from 1 Myr to 13 Gyr and are color-coded from blue to red (see legend). Fluxes are normalized to the SDSS g band. Younger galaxies are dominated by light from the blue part of the spectrum, and vice versa for older galaxies.

Figure 4.5 shows the spectrum of a galaxy with solar metallicity sampled at different ages which was created with the evolutionary stellar population synthesis code `GALAXEV`[3] (Bruzual & Charlot, 2003, BC03). We will give a more detailed description of the code in Section 4.2.2 when we compare the predicted model SED colors with observational data and introduce the technique used by `GALAXEV` in more detail. For the SEDs in Figure 4.5 we assumed a single stellar population (SSP), meaning that all stars in the galaxy were formed at the same time (or age, $t = 0$) in a single δ-functional star burst. After that, the spectrum is only changed by the age evolution of the stars on and off the main sequence, without subsequent star formation (*passive evolution*). We mentioned already in Section 4.1.1 that we can observe the dying out of stars from the upper main sequence as a function of time. These massive stars dominate the spectrum shortwards of the 4000 Å break for young ages ~ 1 Myr, and we can see that the light is dimmed in this regime for increasing ages. As these stars are short lived, the most massive ones have left the main sequence at about 10 Myr, causing the flux in the UV to decline and to rise in the NIR. Until a few Gyrs the UV light is still continuously dimmed and the NIR flux is maintained by stars on the asymptotic giant branch (AGB). After several Gyrs the fluxes redwards of the 4000 Å break are dominated by red giants who have short lifetimes and the luminosity is determined by the rate at which

[3] `http://www.cida.ve/~bruzual/bc2003` or `http://www.iap.fr/~charlot/bc2003`

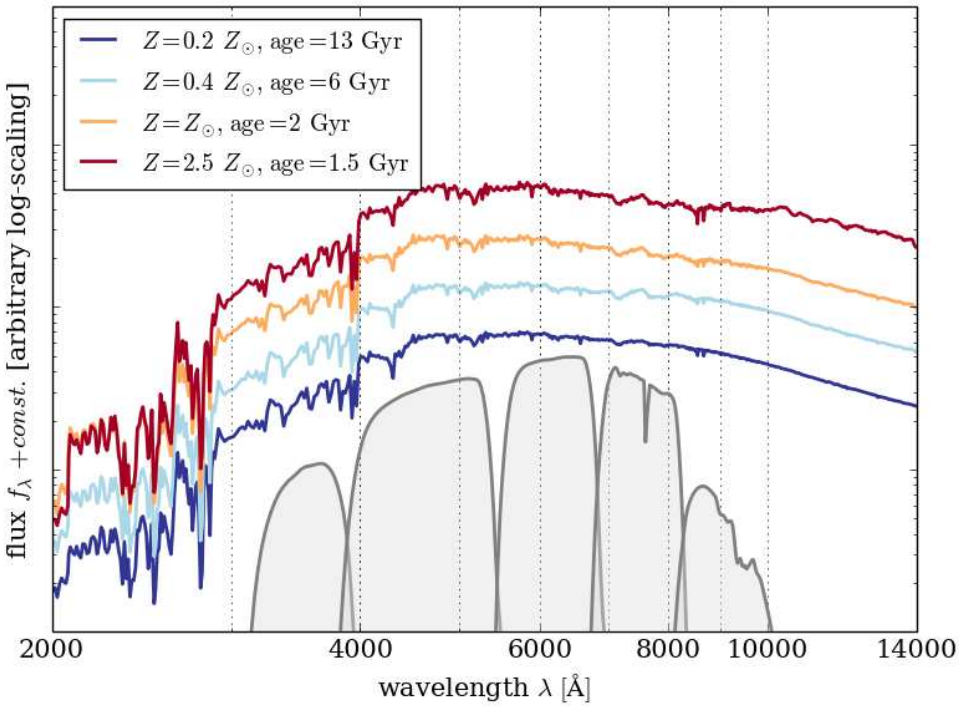

Figure 4.6: SEDs of BC03 SSP Models sampled at various metallicities and ages. The SEDs are normalized to the SDSS g band filter, and offsets to the SEDs have been applied for clarity of the figure. Both axes are scaled logarithmically. Furthermore, the SDSS *ugriz* filter curves are plotted in gray. Note, that the color-code is different from Fig. 4.5, it goes from blue to red for increasing values of Z (not age).

stars evolve off the main sequence. On the other hand, the light in the UV rises again due to post-AGB stars and hot dwarf stars which cause the upturn in the UV at $t \gtrsim 1$ Gyr. Although the changes in the redder part of the spectrum are not significant at ages of $\sim 4 - 13$ Gyr, we can observe an increase in the strengths of the absorption lines. This is due to the abundances of metals and because light is absorbed mostly through electron scattering in younger stars. While increases in line strengths are elicited by an aging stellar population they can also be caused by an increase in metallicity leading to the age-metallicity degeneracy discussed in the following.

In Figure 4.6 we show the SEDs of SSPs with different metallicities Z sampled at various ages, and the SDSS filter curves. We can observe that the SEDs look very much alike in the wavelength range covered by the SDSS filters, wherefore we cannot expect that the degeneracy between the age and the metallicity can be lifted through SDSS colors. However, we could resolve this issue by the inclusion of different colors outside the visible part of the spectrum. Furthermore, if a spectrum is available, one can determine the age and the metallicity of a galaxy without degeneracies, e.g., by means of the Hβ index, which is more sensitive to age than to metallicity.

In the precedent analyses we always assumed a single stellar population originating from a

single star burst at the birth of the galaxy. However, although (as mentioned above) elliptical galaxies are believed to form their stars at low ages and fast, assuming that their stellar content is created at a single moment in time is not realistic. The colors and SEDs of ellipticals may be reproduced well to some extent by an SSP, but we can expect that star formation occurred for a longer duration. Since the gas content in a galaxy is more and more consumed and bound in stars, it is generally assumed that the star formation rate (SFR) decreases over time. Usually, an exponentially decreasing star formation rate is used to describe the SFH of a galaxy (e.g., Shapley et al., 2005; Longhetti & Saracco, 2009):

$$SFR(t) \equiv \psi(t) \equiv M_{\odot} \cdot \tau^{-1} \cdot \exp(-t/\tau), \tag{4.8}$$

normalized to a mass of $1\,M_{\odot}$ at $t = 0$, while τ is the e-folding timescale which is a measure of how long stars are typically formed. These populations that do not originate from one single star burst are so-called *composite stellar populations* (CSPs). The functional form of Equation 4.8 is also some times called the τ *model* of star formation history.

4.2.2 Various Model SEDs in Comparison to the SDSS-II LRG Data

In this section we compare the colors of publicly available model SEDs, and potential template SEDs, with the colors of the SDSS-II LRG data, where we search for a good representation of the data. Most of the text was published already in Greisel et al. (2013). We extended it in some details.

The *PZstandard* Red Templates

The first set of models we investigate are three red models that were and are used with the `PhotoZ` code of Bender et al. (2001) in a great number of publications (Drory et al., 2001; Gabasch et al., 2004; Feulner et al., 2005; Brimioulle et al., 2008; Gabasch et al., 2008; Saglia et al., 2012; Brimioulle et al., 2013; Gruen et al., 2013, 2014) and which will be included in our model set when we estimate the photo-zs of the DES SV dataset in Sec. 7. We want to see whether they can represent the LRG data in terms of color, and therefore if they are a suitable representation of the LRGs for photo-z estimations. In the following they are tagged *PZstandard* as they are part of the usually used models with the `PhotoZ` code we work with in this thesis.
There are three red models currently in the *PZstandard* template set. One is an SED of an S0 galaxy (see Fig. 4.1) which was created by Mannucci et al. (2001). The other two are SEDs of elliptical galaxies created by Bender et al. (2001) in order to account for two slightly different subsets of red objects in the Hubble Deep Field North at $z \sim 0.2 - 0.4$. They are based on SSP models of Bruzual & Charlot (1993).
Figure 4.7 shows the SDSS $u-g$, $g-r$, $r-i$, and $i-z$ colors as functions of redshift, similar to Figure 3.6, but with the model colors plotted on top. Concerning the $u-g$ color all three models fit the data well at very low redshifts and lie well between the median data errors for redshifts greater than ~ 0.3. In between, all three models are bluer than the LRG data and fall outside the error bars as well as the densely populated region. The $g-r$ and $r-i$ colors have the greatest impact on photometric redshift estimation since those colors are varying as a function of redshift (at least in the z-range considered), and thus a relation between color and z can be established. These colors are represented best by the Mannucci et al. (2001) SED, although deviations from the LRG colors can be observed in $g-r$ for z lower than ~ 0.3

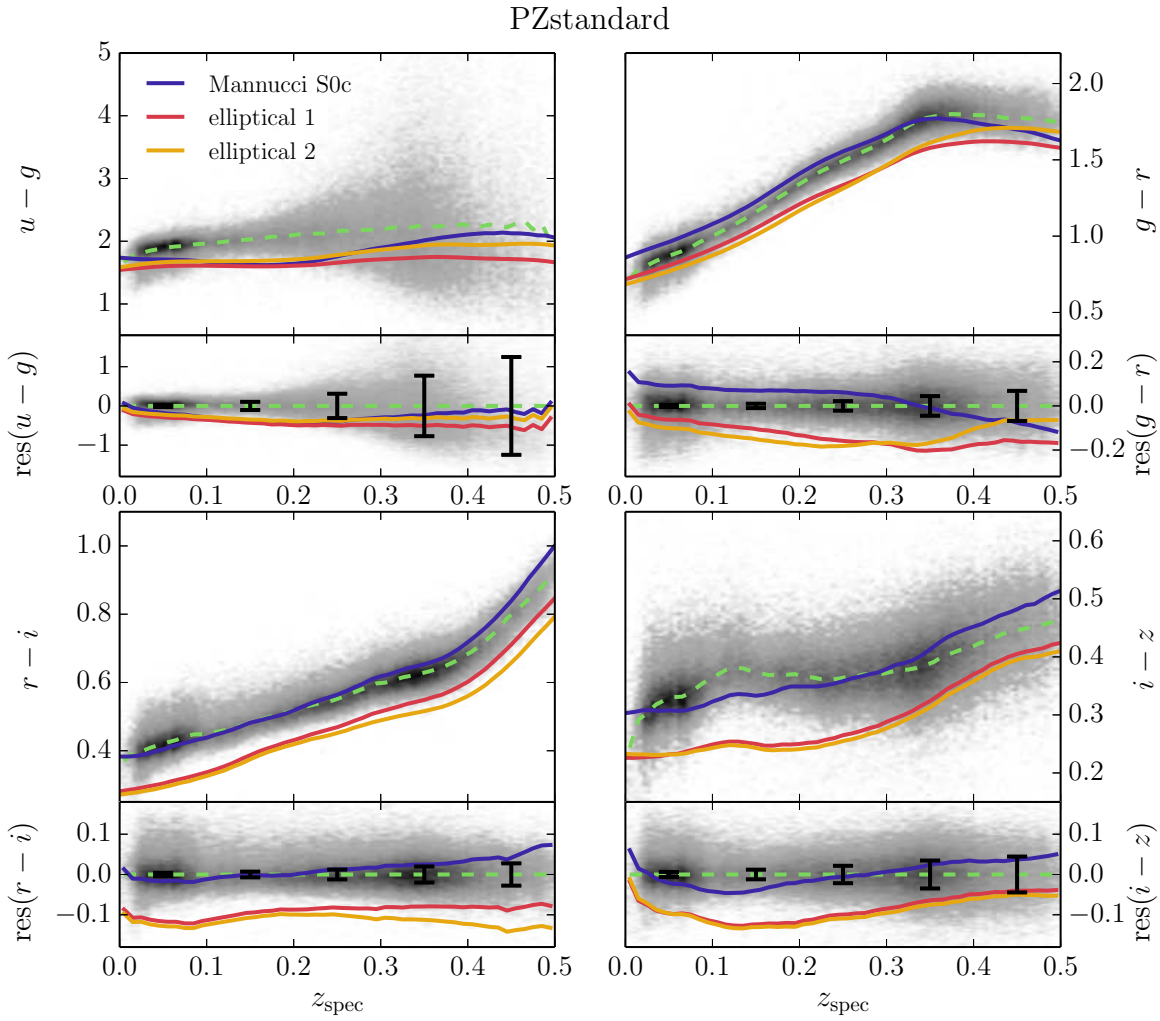

Figure 4.7: Redshift versus colors of the SDSS-II LRG sample (gray, see also Fig. 3.6) and of the *PZstandard* model set. The upper parts of the four major panels show $z_{\rm spec}$ versus the $u-g$, $g-r$, $r-i$, and $i-z$ colors. The median colors as functions of redshift are highlighted by green dashed lines. On top of that the colors of the *PZstandard* models are plotted. In the lower panels we show the residual color from the median color of the data (therefore the median color is the $y=0$ line). Also, the models are plotted on top of the data, where the data median was applied as an offset. Error bars denote the median data color errors in five redshift bins. The SEDs will be shown in the upper left panel of Fig. 7.1 in Sec. 7.

where the SED has a redder color and for z higher 0.45 where it is redder. Furthermore the $r-i$ color is redder for $z \gtrsim 0.3$. These deviations seem not to be of great significance, but $i-z$ is only matched at $z \sim 0.25$ and the form of the function differs strongly from the data median. The other two elliptical templates behave similar to the S0c SED in $u-g$ but are shifted greatly into bluer color values for the other three investigated colors.
We summarize that these models cannot represent the LRG data, although they are useful for the photometric redshift estimation of less luminous red galaxies in different data sets (e.g., Chapter 7 and Brimioulle et al., 2013).

CWW Templates

The templates used in almost every photometric redshift estimation with template fitting methods are the SEDs of Coleman, Wu, & Weedman (1980, hereafter CWW). We will concentrate on the elliptical CWW SED that is the mean spectrum of a number of local elliptical galaxies and should represent our data best. It was modified and extended by a number of authors and we will have a look at three variations. The `BPZ` code[4][5] (Benítez, 2000) provides a variation of the CWW elliptical spectrum which was extended to the UV through linear extrapolation, and to the near-IR with synthetic templates created with the `GISSEL` code. Extinction was accounted for following Madau (1995). The current public version of `BPZ` (Benítez et al., 2004) includes a repaired version of this elliptical SED, which was calibrated with an array of ground based data. The colors of both the extended and the repaired red CWW SED, are shown in Figure 4.8. Both match the data well in $u-g$ and $r-i$, with the exception of the former for higher z. There are deviations from the data in $g-r$ over the whole considered z range, except for $z \gtrsim 0.45$ for the original CWW SED, and for small to medium redshifts for the repaired version. Furthermore, they are able to match the data median in $i-z$ only for medium and high redshifts.
Another Bayesian photo-z software is `LePhare`[6] (Arnouts et al., 1999; Ilbert et al., 2006), which uses a wide range of template SEDs. One SED based on CWW spectra was generated by a linear interpolation of the four CWW SEDs and a star-forming galaxy which was constructed with the population synthesis code of Bruzual & Charlot. It was extended to the IR with the `GISSEL` code. The resulting colors of the interpolated red CWW SED are shown in Figure 4.8 together with the colors of the `BPZ` CWW models. Its colors are very similar to the original model and the repaired `BPZ` version, except in $u-g$, where it is bluer for higher z. The fact that these observationally created SEDs can reproduce the LRG colors well, and are therefore in principle well suited for photo-z estimations, was also pointed out by Benítez (2000).

BC03 Models

In this section we compare the colors of models from Bruzual & Charlot (2003, hereafter BC03) with the SDSS LRG colors. BC03 models can be created by the software `GALAXEV` provided by BC03 on the websites[7]. Since we will use BC03 models in Sections 5.1 and 6.2.1 as a basis to create model SEDs specifically designed for the SDSS-II LRGs and BOSS CMASS

[4] v1.98b: `http://acs.pha.jhu.edu/~txitxo/`
[5] v1.99.3: `http://www.its.caltech.edu/~coe/BPZ/`
[6] `http://www.cfht.hawaii.edu/~arnouts/lephare.html`
[7] `http://www2.iap.fr/users/charlot/bc2003/`, and `http://www.bruzual.org/~gbruzual/`

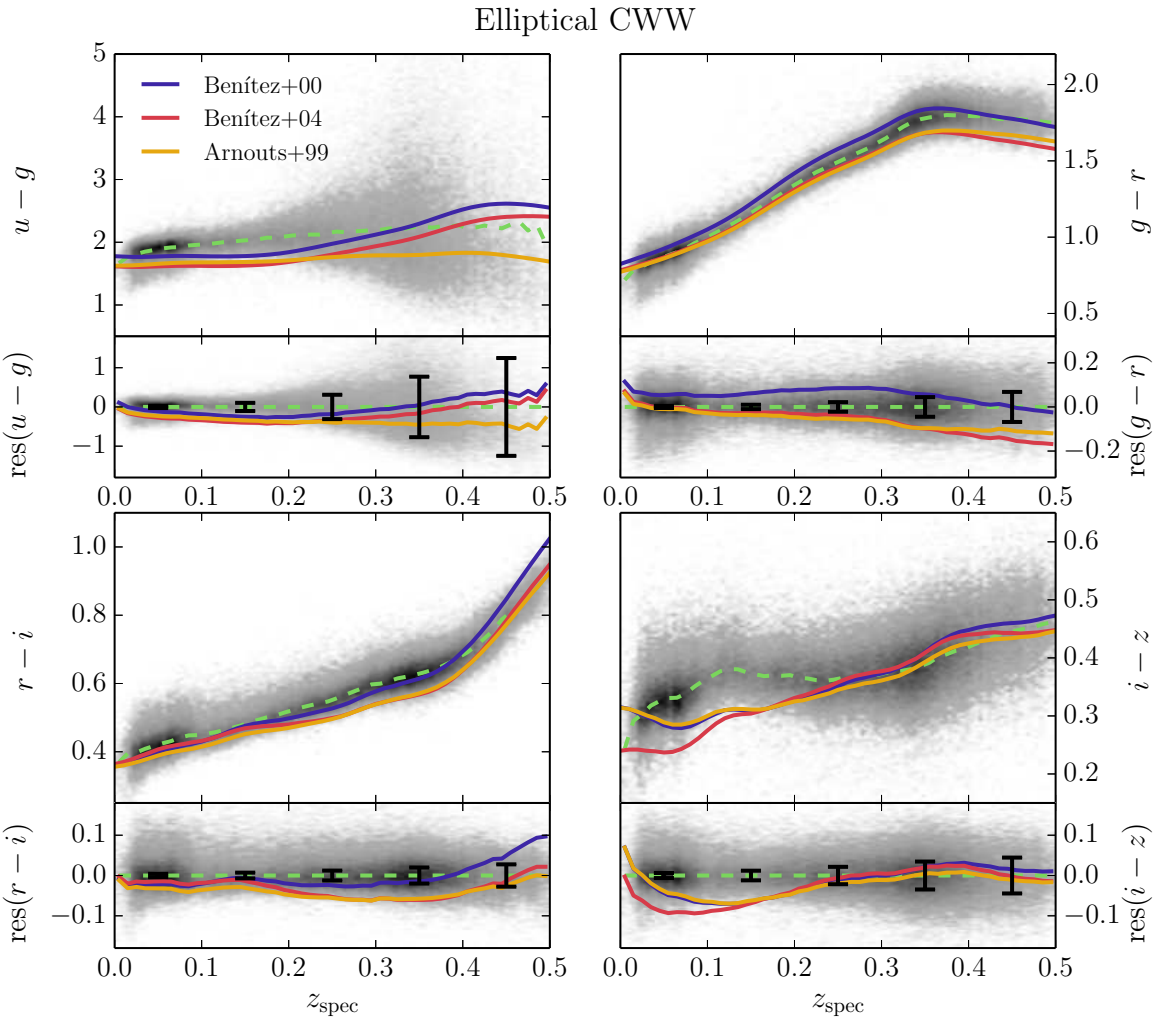

Figure 4.8: Redshift versus colors of the SDSS-II LRG sample and of several CWW based models. The blue line shows the colors of the original CWW model of an elliptical galaxy extended by Benítez (2000), whereas the other two SEDs are modified versions of it (a more detailed explanation is given in the text). A more detailed description is given in Figure 4.7.

galaxies respectively, we want to give a brief introduction into their creation. For further details we refer to Bruzual & Charlot (2003), as well as to the documentation of `GALAXEV`.
BC03 models can be created from a variety of huge stellar libraries, empirical as well as theoretical. Metallicities range from subsolar (i.e., $0.005\,Z_\odot$) to supersolar (i.e., $2.5\,Z_\odot$). Initial mass functions available are from Salpeter (1955) and Chabrier (2003). The Salpeter IMF follows a simple exponential law, $dN/dm \propto m^{-2.35}$, whereas the IMF of Chabrier (2003) varies depending on the stellar mass, similar to an IMF by Kroupa (2001). The Chabrier IMF follows the same function as a Salpeter IMF for masses above the solar mass, and has log-normal form for $m < M_\odot$:

$$\frac{dN}{d(\log m)} \propto \exp\left(-\frac{(\log m - \log m_c)^2}{2\sigma^2}\right), \tag{4.9}$$

with a characteristic mass $m_c = 0.079$ and $\sigma = 0.69$ for single stars, and $m_c = 0.22$, $\sigma = 0.57$ for binaries. The BC03 models are available for ages ranging from 0 to 20 Gyr, although we will extract SEDs only for ages lower than or equal to 13 Gyr, below the age of the Universe (13.8 Gyr, Planck Collaboration et al., 2014).
In the creation of model spectra BC03 use the so-called *isochrone synthesis technique*, cf. also Charlot & Bruzual (1991) and Bruzual & Charlot (1993), where a spectrum of an SSP is obtained by summing up the individual stellar spectra along an isochrone. A series of SSPs can then be chosen to recreate the stellar population of a CSP, with a given star formation history. Assuming the IMF does not depend on time (which may not be accurate, e.g, Weidner et al., 2013) the spectral flux per unit wavelength as a function of time can then be written as (see, e.g., BC03):

$$f_\lambda(t) = \int_0^t \psi(t - t')\; S_\lambda\left(t', \zeta(t - t')\right) dt', \tag{4.10}$$

where $\psi(t)$ is the SFR, and $\zeta(t)$ describes the metal enrichment (e.g., Tinsley, 1980). $S_\lambda\,(t', Z_i))$ is the spectrum of an SSP, hence the sum of the stellar spectra along an isochrone, per unit mass. An SSP can then also be described by Equation 4.10 with $\psi(t - t') = \delta(t - t')$. The stellar content of the galaxy is prescribed by the IMF, and stellar SEDs are assigned to the individual stars.

In the following analysis and in the model creation later on we use BC03 models created from the theoretical BaSeL 3.1 stellar library, which covers a wavelength range from 91 Å to 160 µm in 1221 wavelength steps. We use the Padova 1994 evolutionary tracks recommended by BC03, which provide a larger variety in metallicity than the Geneva tracks, and which are favored over the Padova 2000 tracks, since the integrated $V - K$ colors of ellipticals of the latter can only be matched at ages higher than the currently estimated age of the Universe. For further information about the stellar library and evolutionary tracks we refer the reader to BC03 and references therein. We extract model SEDs with a Chabrier IMF, and metallicities of $Z = 0.2, 0.4, 1, 2.5\,Z_\odot$. SFRs are either SSPs or CSPs with exponentially decreasing SFHs with e-folding timescales $\tau \in [0.5, 1.0, 2.0, 3.0, 5.0, 8.0, 20.0, 50.0]$ Gyr, where the latter approximate nearly constant SFRs. The BC03 models most frequently used in this work are sampled at ages from 0.01 to 13 Gyr. We will encounter deviations from this setup in Chapter 8. The code used in the creation and extraction of the BC03 models is presented in Appendix F.1.

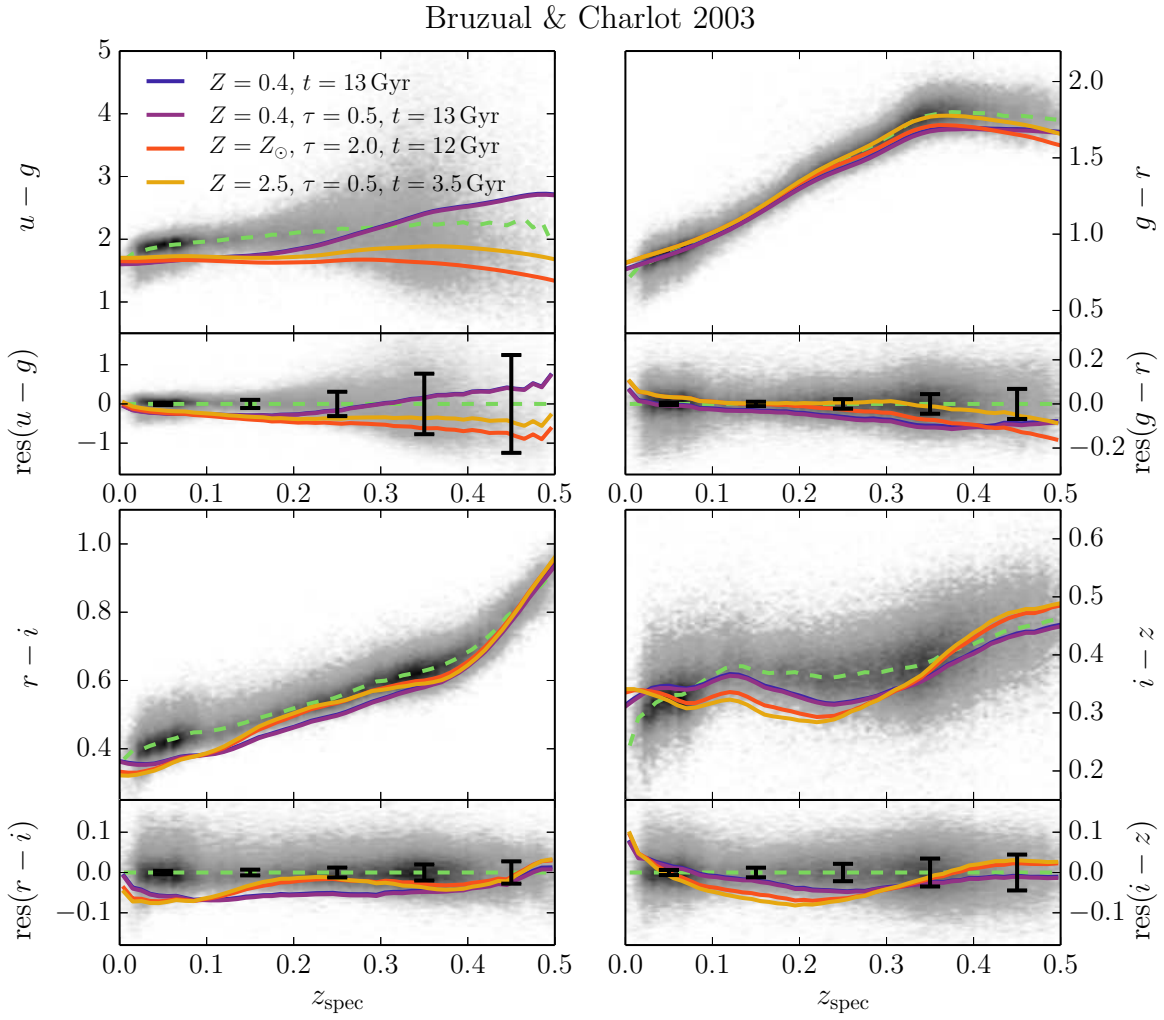

Figure 4.9: Redshift versus colors of the SDSS-II LRG sample and of several BC03 models, which are chosen for the plot in such a way that the error weighted residual between model and data colors is minimal. The solid lines are a selection of BC03 models that represent the data best, from which the blue line is an SSP and the other lines are CSPs. Parameters of the models are given in the figure and we refer to the text for detailed information. A more detailed description is given in Figure 4.7.

SEDs become redder in color with increasing metallicities, ages and decreasing e-folding timescales (cf. Figs. 4.5 and 4.6). We compare only those with the data that match their median best. Hence, we compute the residual between the data and the BC03 models for the $u-g$, $g-r$, $r-i$, and $i-z$ colors.

$$\mathrm{Res}^2(T) = \sum_{\text{objects}} \sum_{j\in\{u-g,\, g-r,\, r-i,\, i-z\}} \frac{(c_{j,\mathrm{data}} - c_{j,\mathrm{mod}})^2}{\sigma^2_{j,\mathrm{data}}}. \tag{4.11}$$

The model with the lowest residual is an old (13 Gyr) SSP with subsolar metallicity ($Z = 0.4\,Z_\odot$). The model that matches the colors of the data second best is a CSP with a comparably small e-folding timescale $\tau = 0.5$ Gyr. It has the same age as the SSP, and therefore it is expected that both models have very similar colors, since star formation has taken place in the CSP only for a short term in comparison to the age of the galaxy. The colors of both populations are presented in Figure 4.9. The two models are not easily distinguishable in the plot due to their similar colors. They show a reasonably good match in the $u-g$ color, but are too blue in $g-r$ at higher redshifts. Their $r-i$ color is bluer than the data throughout the whole z range, except for $z \gtrsim 0.45$. Finally, in $i-z$ they are too blue within $0.1 \lesssim z \lesssim 0.35$, and too red for small z.
In the plot we also show the third and forth best fitting models. The third has solar metallicity, $\tau = 2.0$ Gyr, and an age of 12 Gyr, whereas the forth best has a super-solar metallicity of $2.5\,Z_\odot$ an e-folding timescale of 0.5 Gyr, and a young age of 3.5 Gyr. Both are bluer than the best two models in $u-g$ for the upper half of the z-range. In their $g-r$ color the third and forth best model are similar to the first two, and the supersolar metallicity model provides a better match to the color median of the data. Furthermore, the third and forth best models better describe the data in the medium redshift range. However, the deviations in the $i-z$ color are much higher than for the two first best fitting models.
We conclude that we could not isolate a single or a small set of BC03 models that are able to recreate the SDSS-II LRG colors. As with the previously investigated model SEDs, the track in $i-z$ deviates strongly from the data.

`LePhare` Templates

As mentioned previously (Sec. 2.2.2), the `LePhare` code is one of the earliest published, and widely-used template fitting photometric redshift codes. It comes with a large variety of template SEDs which we want to analyze. We adopt the nomenclature of the models used within the `LePhare` code in the following. Again, we investigate only red SEDs that produce the smallest residuals from the data according to Eq. 4.11. The SED delivering the lowest residual originates from CWW models and was shown in Fig. 4.8. Figure 4.10 presents the colors of three other SEDs that yield the second, third and fourth smallest residuals from the data. The one termed *COSMOS* is from the `LePhare` set called *COSMOS_SED*. Models from this set were used in Ilbert et al. (2009) for photometric redshift estimates in the 2 deg^2 COSMOS field. The model tagged `GRASIL` is a 13 Gyr old elliptical galaxy that was created with the `GRASIL` code (Silva et al., 1998). The SED describing the data fourth best comes from the `LePhare` *42_GISSEL* set. It has an e-folding timescale of 1 Gyr, a metallicity of $Z = 0.02$ ($Z_\odot$), and an age of 7 Gyr. From Figure 4.10 one can infer that the $g-r$ colors of those three models are redder than the data for low z. In $r-i$, the three model SEDs are too blue for lower redshifts. Their $i-z$ color is redder than the data at $z \lesssim 0.07$ and $z \gtrsim 0.45$, and bluer at $0.07 \lesssim z \lesssim 0.3$

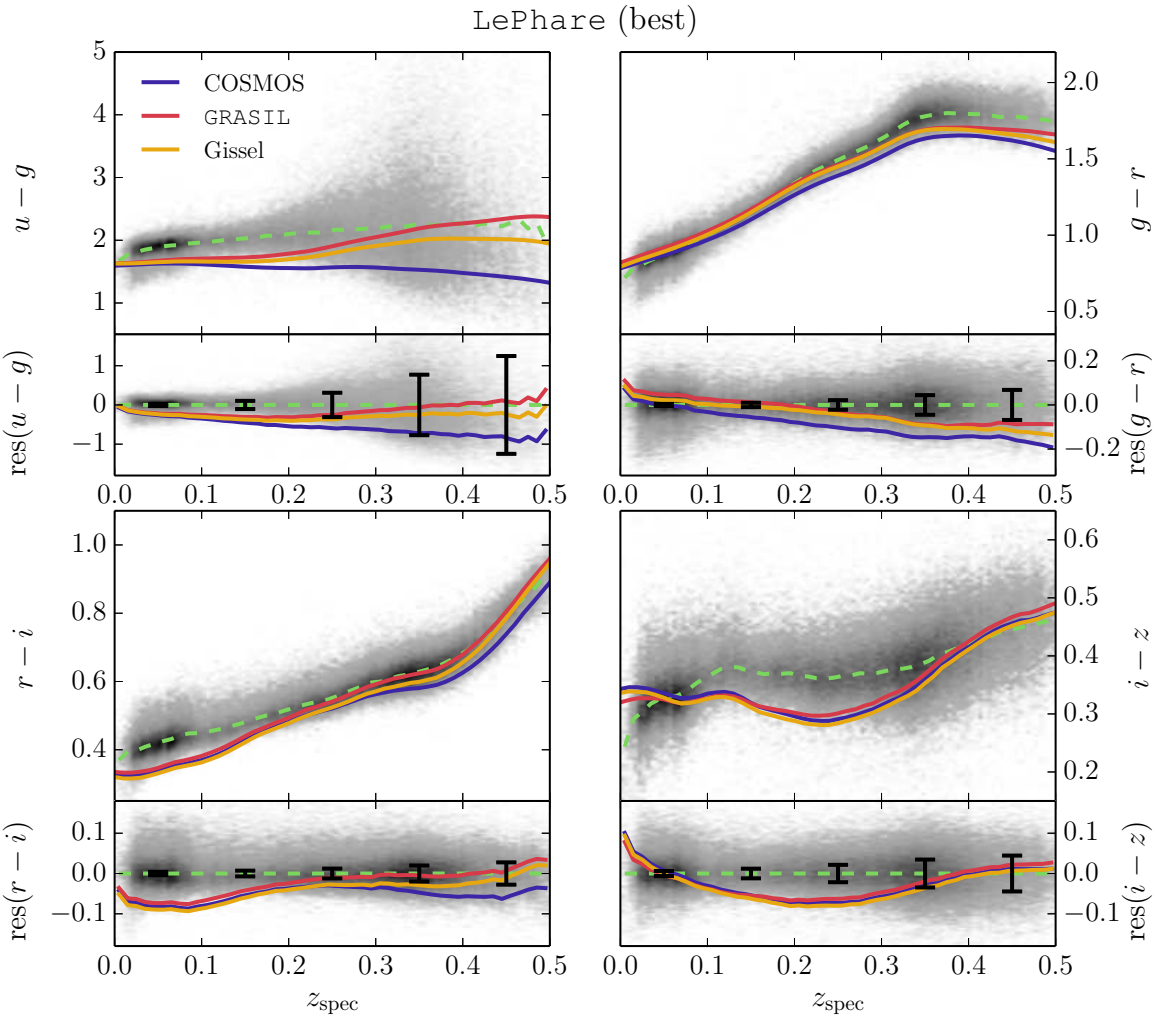

Figure 4.10: Redshift versus colors of the SDSS-II LRG sample and of some of the model SEDs from `LePhare`. The `LePhare` model SEDs were selected from the large variety of templates in such a way that the error weighted residual between model and data colors is minimal. A more detailed description is given in Figure 4.7.

Maraston 2009 Model

The problem of the color mismatch between the SDSS-II LRG data and available model SEDs was also addressed by Maraston et al. (2009, tagged M09 henceforth). In their paper they focused on the $g-r$ and $r-i$ colors, and could neither find a synthetical nor an empirical SED that could represent the data throughout the whole redshift range. M09 added blue, metal-poor stars to CSP models at a mass fraction of 3 % to create a model[8] that matches the LRG colors to a better extent. We present the model colors as a function of redshift for two different constant ages, 8 and 13 Gyrs in Figure 4.11 (blue and violet lines). The data seems to be represented best by SEDs with ages 14 and 15 Gyrs, but we omit showing these since in our cosmology the Universe is 13.7 Gyr old. All considered models show very similar $u-g$ colors. The 13 Gyr version of the M09 model produces a $g-r$ color which is redder than the data. It has $r-i$ colors that are too blue in comparison to the data at low redshifts, it fits the data well around $z \sim 0.38$, and is too red for higher z values. The $i-z$ values of the data are reproduced best by the 13 Gyr model, but still too blue in the lower redshift range. In contrast to that, the 8 Gyr model matches $g-r$ very well, but is too blue in $r-i$ until $z \sim 0.45$, as well as in $i-z$ for $z \lesssim 0.3$.
Additionally to the model at specific constant ages, we want to present the M09 model evolving with redshift. Therefore we interpolate the model in between the available ages with the galaxy evolution software `EzGal`(Mancone & Gonzalez, 2012). We use formation redshifts of $z_f = 3.0$ and 20, which correspond to an age of the Universe of 2.2 Gyr and 0.2 Gyr respectively and assume passive evolution afterwards. As mentioned earlier, the $u-g$ color is very similar for all versions of M09 considered. Both passively evolving models match the data very well in $g-r$, but show deviations to lower values in $r-i$ except for the lowest and highest z values, and in $i-z$ for $z \lesssim 0.3$ (similar to the 8 Gyr SED).

Summary

Altogether, we conclude that we could not find models that are able to represent the SDSS LRG data within their errors at all redshifts. The reason for this is that the colors of the SEDs differ from the data median as a function of z. The M09 models describe the SDSS LRGs best compared to the templates examined before, but they also suffer from notable deviations within some redshift intervals. The lower redshift range ($z \lesssim 0.1$) and the $u-g$ and $i-z$ colors were not considered in Maraston et al. (2009). But the M09 model is the only model considered that can recreate the shape of the $i-z$ redshift-color track. The two passively evolving versions of the M09 model match the data colors best, but they also suffer from notable deviations within some redshift intervals. Furthermore, if only the age of a SED is varied, one just obtains a one-dimensional sequence in color-color space. Instead, the colors of true SEDs are spread in two dimensions, which is a clear hint that not only the age of the LRGs, but also the SFH, and maybe metallicity, vary. A wider spread can thus only be produced if different SFHs and metallicities are considered, which are not publicly available for M09 models. We therefore conclude that in order to get SEDs which represent the data at all redshifts and in all colors, one has to consider different SED template sets for different redshift bins. We will do this in the next chapter.

[8] `http://www.icg.port.ac.uk/~maraston/Claudia's_Stellar_Population_Model.html`

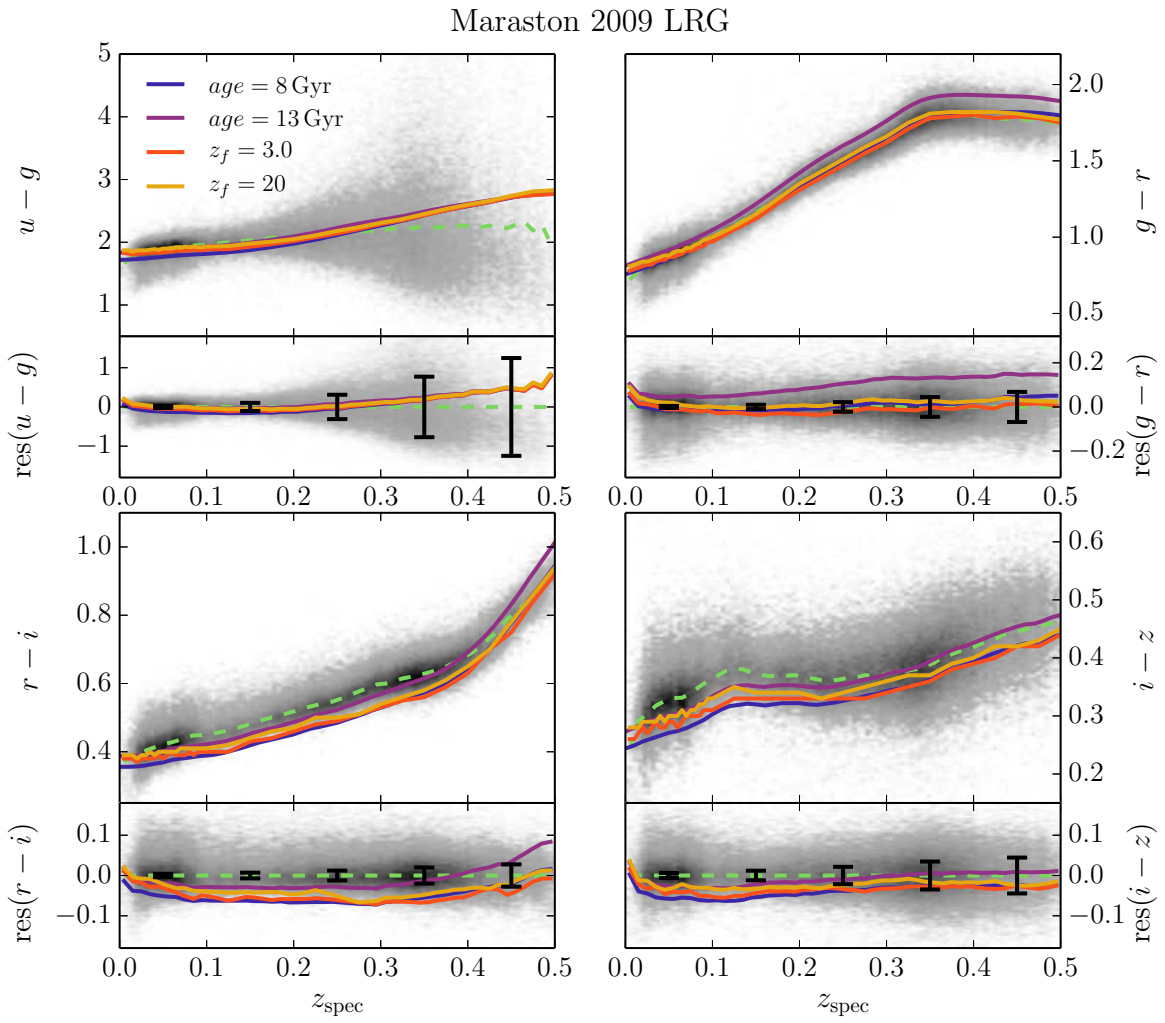

Figure 4.11: Redshift versus colors of the SDSS-II LRG sample and variations of the M09 LRG model. We plot the colors of the M09 model for fixed ages of 8 and 13 Gyrs. The other two lines are the same model evolving with redshift at formation redshifts of $z_f = 3$ and 20. A more detailed description is given in Figure 4.7.

Chapter 5

Novel SED Templates for SDSS-II LRGs

We describe the generation of SED templates for the SDSS-II LRG sample in the following. The creation and selection is equal to what I did in my master thesis, where we performed the same procedure but on a subsample of the SDSS-II LRGs with $u \leq 24.5$. Here, we create models for the whole sample without the u band magnitude cut, and additionally select SEDs that should account for galaxies with more peculiar colors than what is covered by the majority of LRGs. This chapter was published (together with a slightly different version of Sec. 4.2.2) in the *Astrophysical Journal* together with Stella Seitz, Niv Drory, Ralf Bender, Roberto P. Saglia, and Jan Snigula in 2013 (Greisel et al., 2013).

In the past, optimal template SEDs have been obtained by "repairing" individual SEDs empirically using the mismatch of observed and model colors (e.g., Csabai et al., 2003). In addition, or alternatively, a larger variety of template SEDs has been obtained by interpolating between a smaller number of template SEDs (e.g., Brammer et al., 2008). However, there are some caveats. If the models that one starts with are a bad match to the data, it is unlikely that a combination of them will be any better. Moreover, if one introduces SEDs that do not match the data, one deteriorates the photometric redshift quality. Furthermore, the question of how many SEDs are needed to match the data remains undecided.

5.1 Generating Best Fitting SEDs for Individual Objects in the LRG Catalog by SED Fitting

We start with SEDs that match the photometry of individual galaxies as well as possible, and subsequently select a small subset of these SEDs with the goal that they together describe the whole data set. To derive SEDs for individual galaxies, we make use of the largest appropriate freedom to describe a SED with five band data. I.e., we describe the SED of an LRG by a superposition of an SSP or CSP (main component) and a burst model. Furthermore, we allow for extinction of the main and the burst component. The burst and its extinction are physically motivated by the fact that LRGs can contain some young stars. The extinction of the main component is mainly there to increase the degree of freedom and allows to change the continuum slope (i.e., to introduce a "variance" of the main component) of the main stellar population.

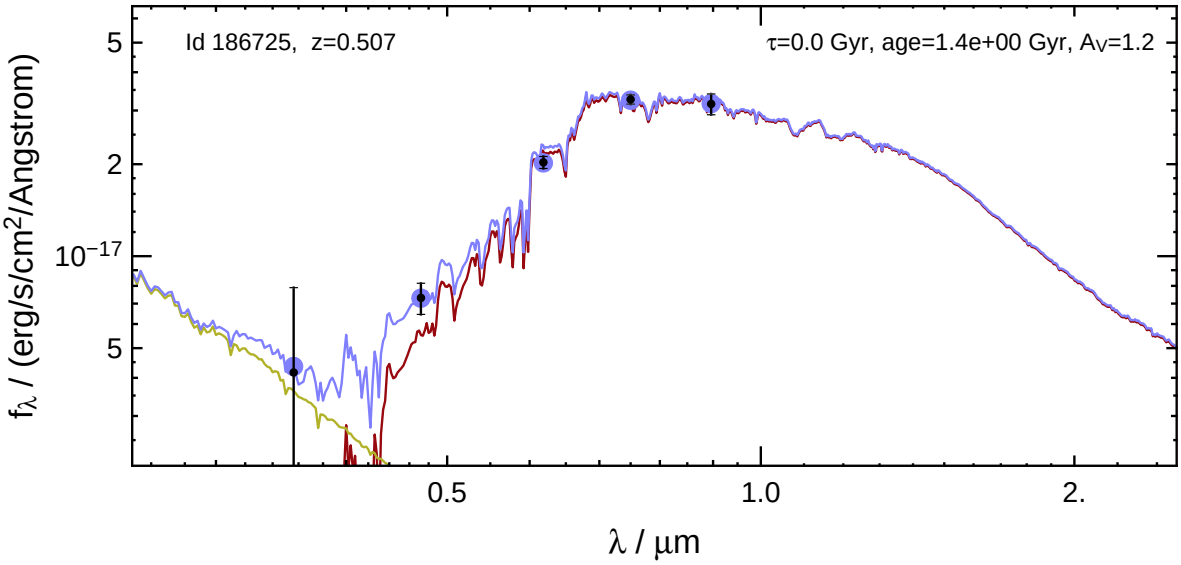

Figure 5.1: Example Output of `SEDfit`. The main component is plotted by a red line, whereas the burst is shown by a green line. The superposition of the two components plus the extinctions for both is shown by the blue SED. The blue points are fluxes of the superpositioned SED in the SDSS filters. Black dots with error bars show the observed photometry in the SDSS u, g, r, i, and z with measurement uncertainties. The ID and the redshift are given in the upper left corner of the plot. In the upper right corner we print the e-folding timescale ($\tau = 0.0$ depicts an SSP), the age ($t = 1.4$ Gyr), and the extinction in the V band ($A_V = 1.2$) of the main component.

We generate SEDs for LRGs with the SED fitting routine `SEDfit` (Drory et al., 2004). It chooses the best fitting SSP/CSP+burst combination through χ^2 minimization, concurrently allowing for dust reddening of both components following the extinction law of Calzetti et al. (2000). We present in Figure 5.1 an exemplary result of an SED fit with `SEDfit`.

In Section 4.2.2 we pointed out that for each model there is always a discrepancy between the model and data colors at some z, and we are not able to pick a set of models which describe the data equally well at all z. Thus, we create templates specifically designed to represent different redshift regions. We do so by splitting the spectroscopic SDSS LRG sample into five subcatalogs, all of them containing objects within redshift intervals of width 0.04, centered on $z = 0.02$, 0.1, 0.2, 0.3 and 0.4. These catalogs contain 9660, 38,583, 9292, 15,093, and 11,667 objects. They are then divided into two equally sized subcatalogs; one serving for template generation (derivation half) and the other one serving as reference when estimating photometric redshift accuracies (validation half). Subsequently, we run `SEDfit` with a model set consisting of 864 BC03 models from the BaSeL library (Bruzual & Charlot, 2003) with Padova 1994 evolutionary tracks and an IMF from Chabrier (2003). They have ages between 0.01 and 13 Gyr, e-folding timescales of $\tau = 0.0$, and $\tau = 0.5$ to 50 Gyr, and four different metallicities, $Z = 0.004$, 0.008, 0.02 ($Z_\odot$) and 0.05 (see also Sec. 4.2.2). In order to fit the photometry, these models can be superposed by a fraction of the burst model, with an essentially constant SFR (i.e., $\tau = 20$ Gyr), solar metallicity, and an age of 50 Myr. The burst mass-fraction is constrained to $[0.0, 0.01]$ with steps of 0.002. Both the CSP and the burst component can be extincted separately. This fitting procedure results in a SED that reproduces the data as well as possible. We end up with one "best fitting SED" for each object in the template generation catalog.

The relative performance of the SED fitting process with models sets of BC03, M09, and Maraston & Strömbäck (2011, hereafter M11) is demonstrated in Appendix B.1, where we also justify why we prefer BC03 models over M09/11 for our SED construction procedure. From Figures B.1 to B.3 in Appendix B.1 we conclude that BC03 models allow a better coverage of the colors of the LRGs than M09/11 models do, and fit the data better (Figure B.4).

5.2 Selection of Best Fitting SEDs for the Novel Template Set

We want to select a set of models from all the SEDs we produced for the template generation subcatalogs such that this set represents the five redshift subcatalogs in color space. We explain our procedure for the $z \approx 0$ redshift interval using Figure 5.2.

First we plot the distributions of the measured colors of the galaxies in the $u-g$ versus $r-i$ plane. Objects which have a u band error smaller (larger) than the median are displayed with light (dark) blue points. Second, we plot the colors of the best fitting SED for each object with red points. If the best fitting SEDs are a good match to the galaxy data, then the distribution of them in color space should be similar to the observed distribution, once the spreading by photometric errors is taken into account. The mean photometric errors for the two subsamples (with high and low u band errors) are shown in dark and light blue in the lower right corner of the panels in Figures 5.2, and B.5–B.9. It shows that galaxies with larger photometric errors in u in general also have larger photometric errors in $r-i$ and $i-z$, which implies that these objects are fainter in all observed bands, rather than just in the u band.
We conclude from the shown sizes of the mean photometric errors on one hand that the spread in colors is of physical origin, and not due to photometric errors. This also holds for higher redshifts where the error bars are of significant size. In order to confirm that, we simulate galaxy colors by assuming that the spread in color space is a result of the photometric errors only. We simulate the data by assuming that each object in fact has a color value of the average color. Then we assign a random deviation from this point, based on a Gaussian probability distribution with the measured error as standard deviation. Since the spread in the simulated data is smaller than that of the data, we confirm that the spread within the colors is of physical origin. On the other hand, we infer that the color distribution should hardly be broadened by photometric errors, and that the best fit SED color reproduce the variety of 'true' galaxy colors well. Only for the reddest colors ($u-g > 1.5$ and $r-i > 0.45$) we find mostly objects with larger than median photometric errors, such that some of these data points will be scattered out from the true distribution by large photometric errors.

We now select model SEDs for the "main galaxy" population. Therefore, we define an area in the *ugri*-plane which contains the majority of objects (black box), and then set a grid within this area (black grid). The chosen area is selected by eye and was confirmed to include at least 70% of the objects from that redshift bin. We want to select one object per cell which shall represent the other objects within that cell. We carry out this procedure only for cells that contain at least 20 galaxies. From each of those cells we pick the SEDs best fitting the five objects from that cell which have the lowest u band error.
As mentioned above, our objective is to produce photometric redshifts with biases as small as

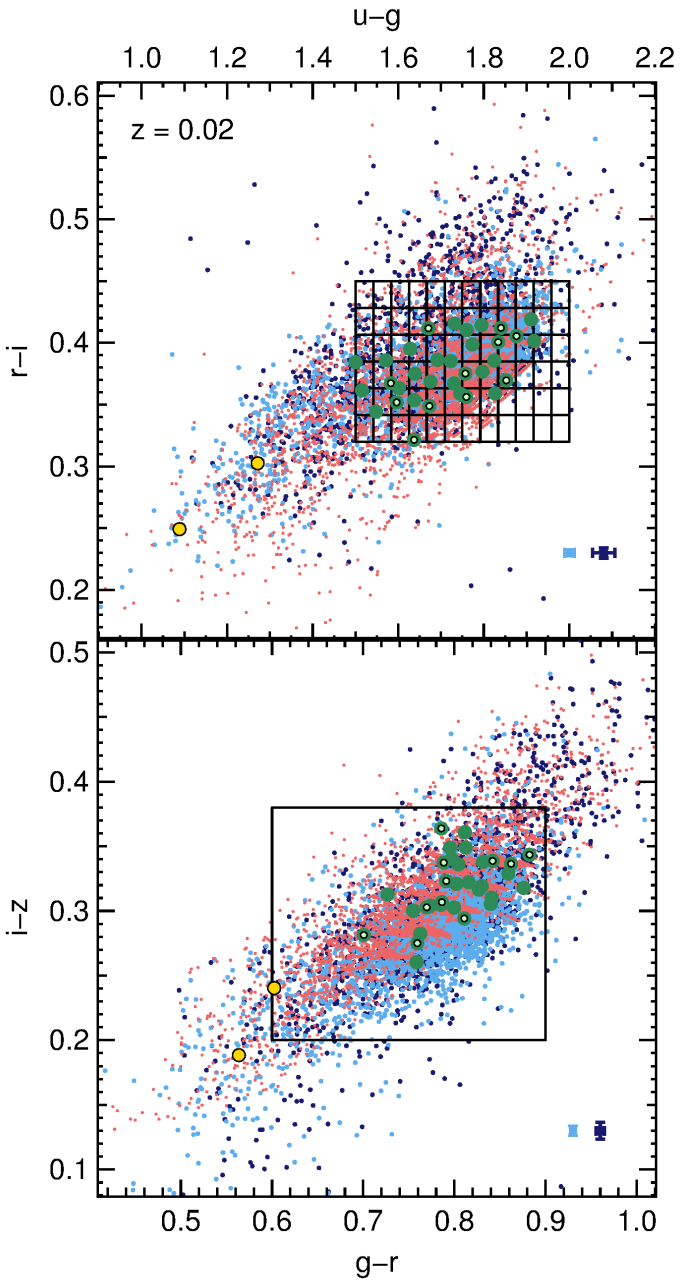

Figure 5.2: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.0 \leq z \leq 0.04$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The grid (upper panel) and the boundary (lower panel) within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs. The yellow dots are models that shall account for objects outside the selected boundaries in *ugri* and *griz*.

possible, and we therefore try to single out SEDs that produce small $\langle \Delta z \rangle$ on the catalogs of the regarded z range. Hence, we perform `PhotoZ` runs on the derivation half and on the validation half of the $z \approx 0.02$ subcatalog, the latter to confirm that we did not unintentionally divide the catalog into halves with different object properties. We perform the photometric redshift estimation with each of the preselected models separately, thus fitting only one model per run. Afterward, we pick the SED that renders the smallest bias on both subcatalog halves out of the five originating from the same cell. The selected model shall represent the objects within its neighborhood in color space in the following `PhotoZ` runs. Repeating the same procedure for all cells and redshift bins, we select roughly 20 to 35 model SEDs per z bin.

In order to save computation time for a `PhotoZ` run, we want to keep the number of templates to a minimum. Therefore, we calculate the root mean squared flux difference of each SED combination at every defined wavelength. The most similar SEDs are then removed from the final set, leaving typically about 10 models per redshift bin with a total number of 49 templates.

The colors of the models is plotted versus their redshift in Figure 5.3. Color–color relations of the new models in comparison to the LRGs are displayed in Figure 5.2, and Appendix B.2, Figures B.5–B.9. The templates represent the LRG colors and their spread for most values of z. The exceptions at $z \sim 0.4$ are discussed in Appendix B.1, Figure B.9. We test the performance of the new models with respect to photometric redshifts in the next section.
As an alternative we also selected template SEDs with a kd-tree. We generated one hyperplane for each of the four colors $u-g$, $g-r$, $r-i$, and $i-z$, thus a four-dimensional tree with 16 template SEDs for each redshift bin. One model was selected from each branch of the tree by the same criteria as already described. As this technique yielded SEDs not too much different, we kept our hand-selected set.

5.3 Photometric Redshift Precision with the Novel Template SEDs

In this section we use the newly created model SEDs to estimate photometric redshifts. Before analyzing the results however, we introduce some quality parameters. The photo-z error is defined as $\Delta z = z_{\mathrm{phot}} - z_{\mathrm{spec}}$ and $\Delta z_{rf} = \Delta z/(1 + z_{\mathrm{spec}})$ in the rest frame respectively. To define badly estimated photometric redshifts we use the rest frame error threshold $|\Delta z_{rf}| > 0.15$ introduced by Ilbert et al. (2006). Objects that fulfill this inequality are tagged *"catastrophic outliers"*, and their number fraction will be called η. $\sigma(\Delta z)$ shall denote the root of the sample variance of the photometric redshift error Δz. A parameter that quantifies the distribution perpendicular to the $z_{\mathrm{spec}} = z_{\mathrm{phot}}$ line without the inclusion of catastrophic outliers is $\sigma_{\mathrm{NMAD}} = 1.48 \cdot \mathrm{median}\,(\Delta z_{rf})_{\mathrm{non-outliers}}$ (Ilbert et al., 2006). With σ_{NMAD} one can make conclusions about the redshift uncertainty excluding the tails of the distribution. A detailed summary of the photo-z quality metrics used in this work is provided in Appendix A.

After these definitions we turn to the integration of the newly created model SEDs into `PhotoZ`. In Section 2.2.2 we mentioned the redshift priors we can impose on each template (cf. Eqs. 2.6 and 2.7). The total probability for a model-redshift combination is proportional to the prior redshift and luminosity (or absolute magnitude) probabilities $P(z)$ and $P(M)$.

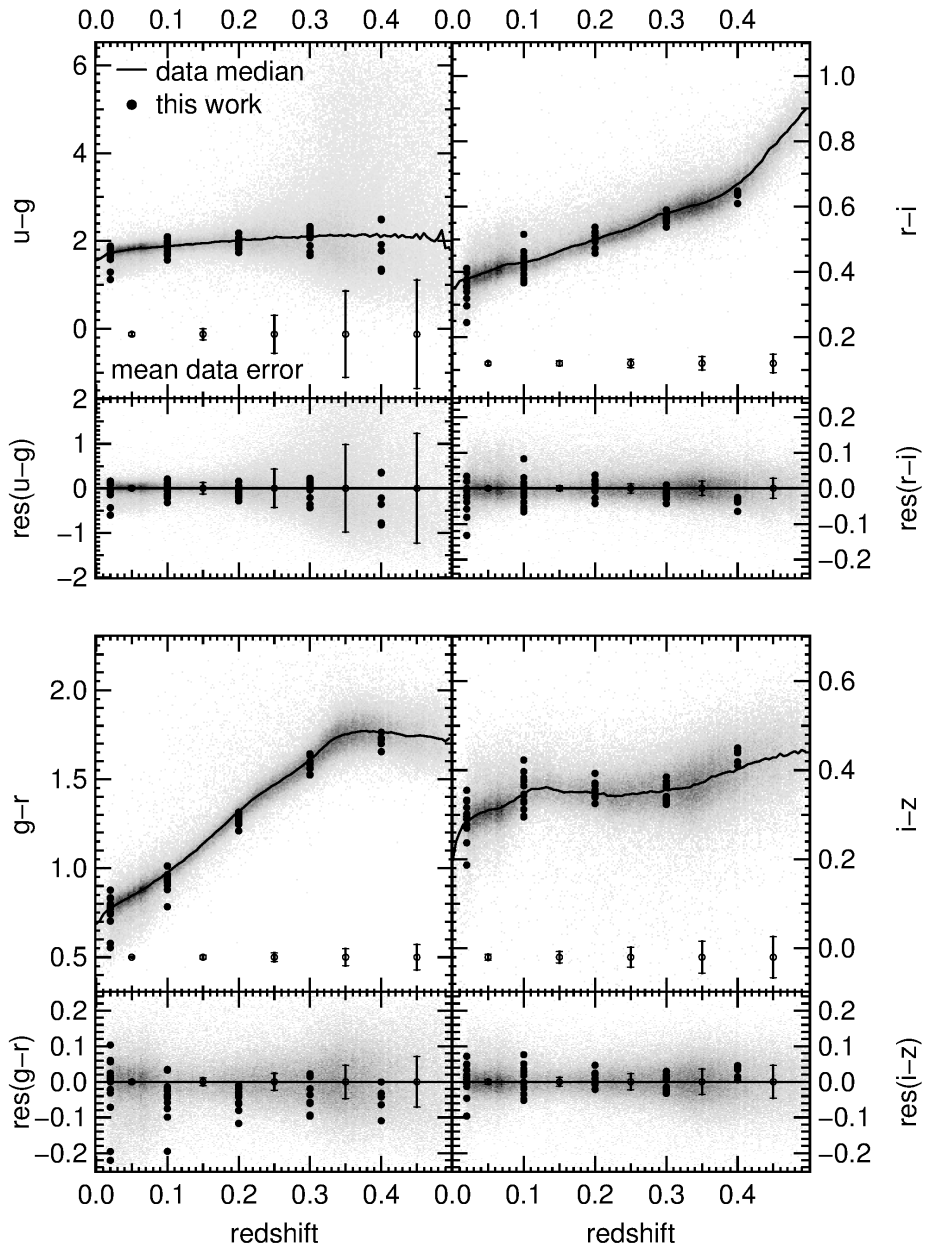

Figure 5.3: *Upper panels:* Colors of the novel template SEDs as a function of redshift. *Lower panels:* Redshift vs. color residuals between models and data. LRGs are represented by the density plot, whereas the solid line is their median. The error bars show the mean data errors in five redshift bins. The colors and the color residuals of the new template SEDs are represented by points. They are plotted at the mean values of the redshift bins from which they were created.

For the luminosity prior we set $\mu_M = -21$ in the Cousins R band (cf. Fig. 3.1), $\sigma_M = 3$, and $p_M = 6$, as luminosities $M_R \notin [-24, -18]$ are very rare among LRGs (see also Section 5.5.2, Figure 5.9) and are thereby excluded. Within this interval however, the luminosity probabilities are virtually equally distributed. We set the z prior to have a Gaussian shape ($p_z = 2$), in order to create a shallow transition from high to low probabilities. Although the LRG sample contains only objects with $z \lesssim 0.5$, the allowed photometric redshift range for all `PhotoZ` runs is $[0, 5.0]$. In this way we can assure that the photometric redshift results truly are below $z \approx 0.5$, although a larger redshift range is allowed. This is important for future work with deeper surveys where the redshift range is not known a priori. The expected values of μ_z are set to the centers of the according redshift interval for each model, thus $\mu_z = 0.02, 0.1, 0.2, 0.3, 0.4$. The redshift prior probability of the model is equal to 1/2 at the corresponding $\mu_z \pm 0.2$.

The photometric redshift results are shown in the upper panel of Figure 5.4. With the new red templates, the described priors, and a resolution of 0.002 we indeed get very accurate redshifts. The catastrophic outlier rate is very small at $\eta = 0.12\,\%$. The dispersions are $\sigma(\Delta z) = 0.027$ and $\sigma_{\mathrm{NMAD}} = 0.017$ respectively. The mean absolute deviation of z_{phot} from z_{spec} in the rest frame reads $\langle |\Delta z_{rf}| \rangle = 0.015$. We show the bias as a function of redshift in the middle panel of Figure 5.4. It exceeds 0.01 only for the highest redshifts ($z \geq 0.47$). The reason for the small positive bias at $z \lesssim 0.02$ is that there are no negative photometric redshifts that can cancel out overestimation. The total mean bias reads -0.0004. At $z_{\mathrm{spec}} \gtrsim 0.36$, the 4000 Å break position in wavelength cannot be determined to high precision. This is because at $z \sim 0.36$ the 4000 Å break is observed at $\lambda \sim 5500$ Å and thus lies in the gap between the SDSS g and r filter. We therefore expect that the error distribution widens at this redshift. The reason for the error becoming bimodal is likely because at this redshift we do not have "perfect templates" and thus the photometric redshifts are either over- or underestimated. Luckily the errors nearly average out for every redshift bin.
The bottom panel of Figure 5.4 shows the dependence of the scatter $\sigma(\Delta z)$ and σ_{NMAD} on z. At low z ($\lesssim 0.01$) it is up to 0.04, which is likely due to contamination of the LRG sample with non LRGs. This is supported by the fact that the χ^2 values in this region are significantly higher than for $z \gtrsim 0.01$. Apart from very small redshifts, the scatter stays more or less the same for all redshifts.

5.4 Comparison to SDSS Database Results

We now compare our photometric redshift results with those in the SDSS database. There are three kinds of photometric redshifts available in the database: One was obtained through template fitting (i.e., a hybrid technique of template fitting and a reparation of templates, see Csabai et al. 2003), and the other two are results of an artificial neural network (ANN) approach (Oyaizu et al., 2008). The results of the template fitting are shown in Figure 5.5. On average, the $\sigma(\Delta z)$ and σ_{NMAD} of these redshifts are smaller than for our method, but the bias is more unevenly spread in redshift space. Also for this method the median bias is smaller than 0.01 only in the redshift range of $0.07 \leq z \leq 0.32$. For other redshifts the bias exceeds that of our redshifts considerably.
The SDSS database values for photometric redshifts obtained with the ANN are shown in Figures 5.6 and 5.7, where we display the results of two techniques, *D1* and *CC2*. The

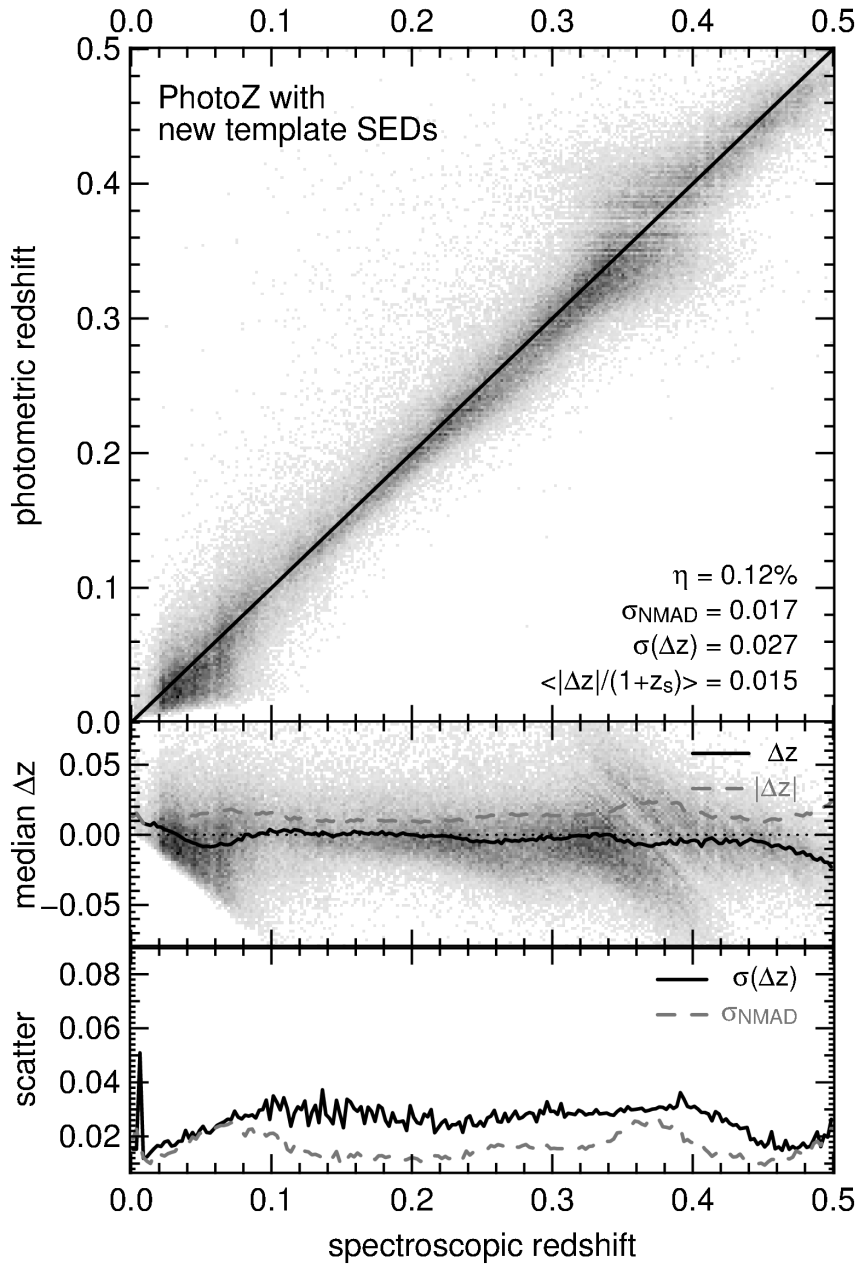

Figure 5.4: Photometric redshift results of a `PhotoZ` run with the newly created templates and 140,331 SDSS LRGs. We removed 20 objects that had undefined photometric redshifts. $\sigma(\Delta z)$, $\sigma_{\rm NMAD}$, and η are explained in Section 5.3 and Appendix A.

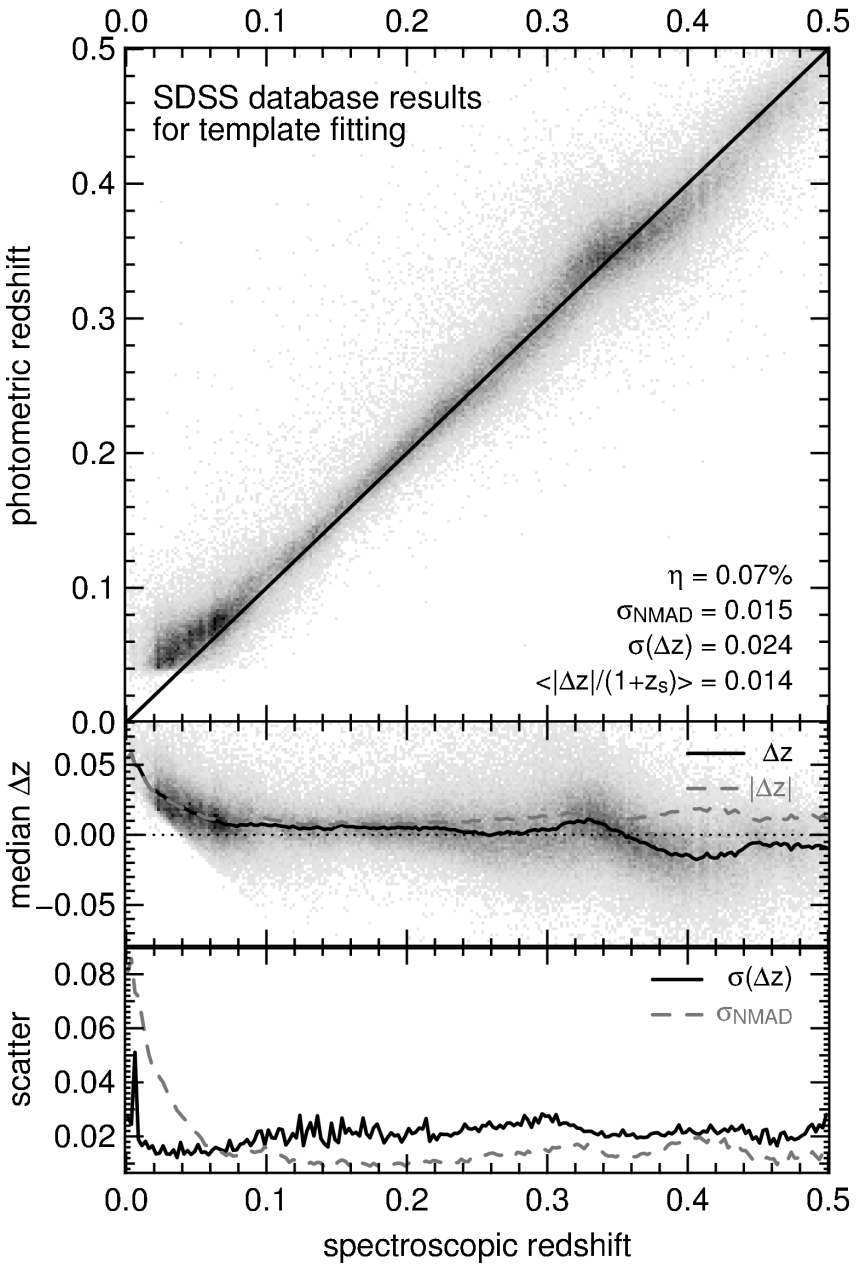

Figure 5.5: Photometric redshift results of the SDSS template fitting method. We removed objects with undefined photometric redshifts. η, $\sigma_{\rm NMAD}$, and $\sigma(\Delta z)$ are explained in Section 5.3 and Appendix A.

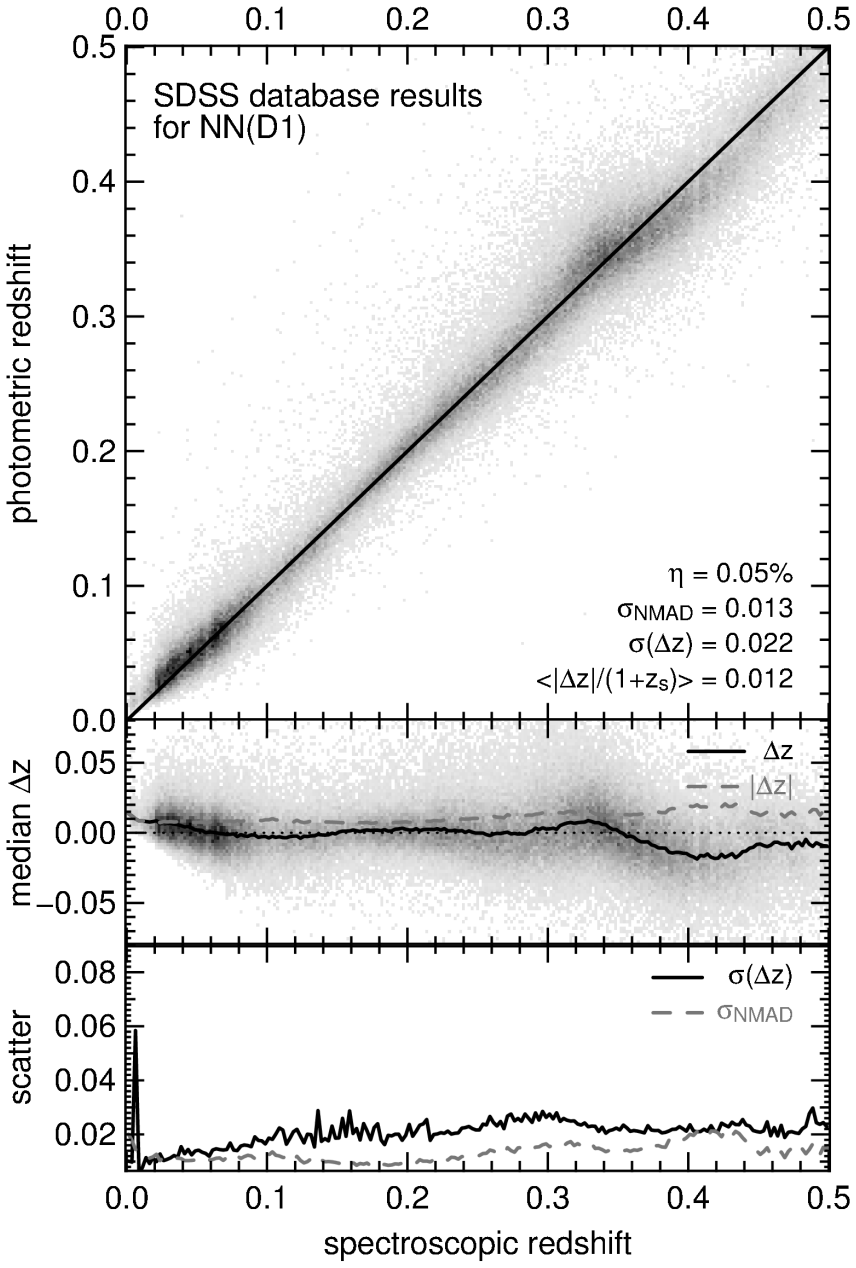

Figure 5.6: Photometric redshift results of SDSS ANN(D1). We removed objects with undefined photometric redshifts. η, $\sigma_{\rm NMAD}$, and $\sigma(\Delta z)$ are explained in Section 5.3 and Appendix A.

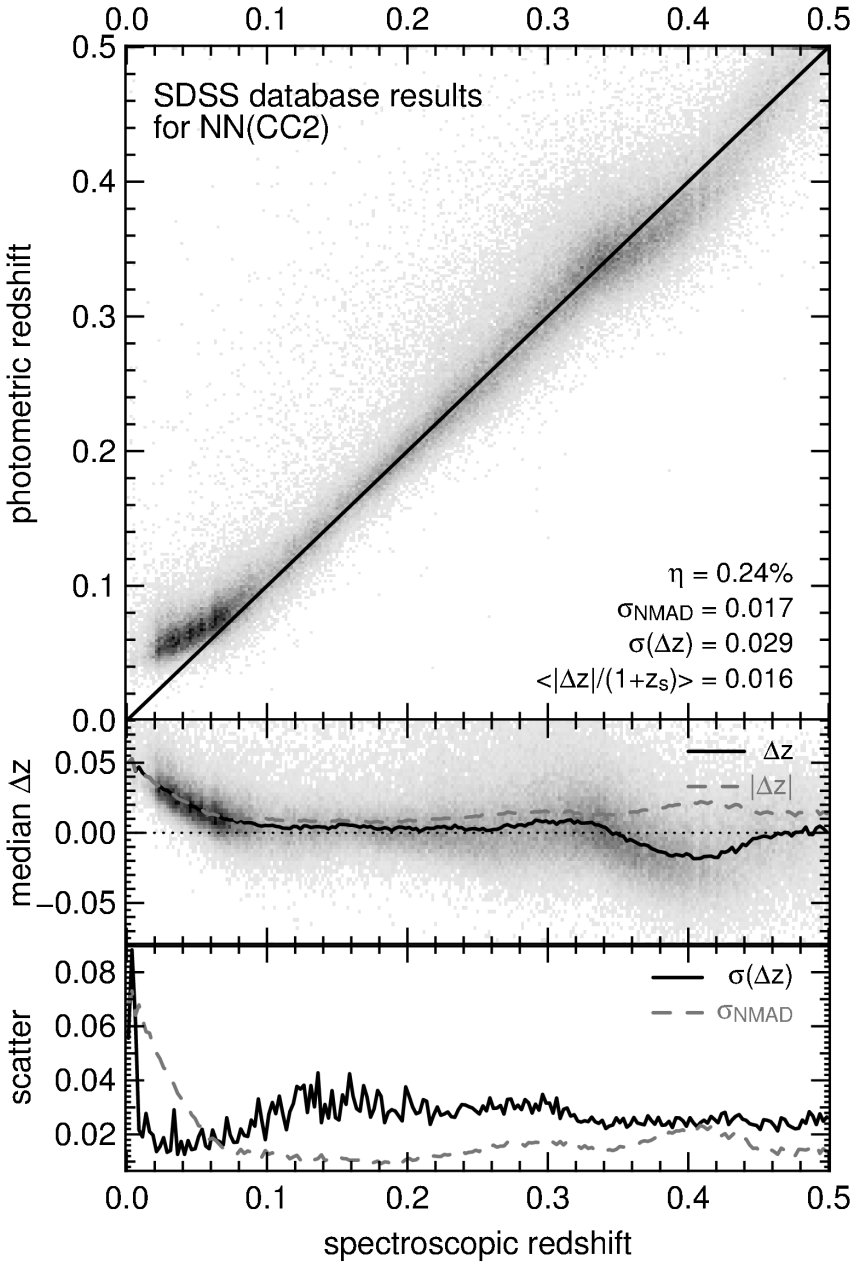

Figure 5.7: Photometric redshift results of SDSS ANN(CC2). η, $\sigma_{\rm NMAD}$, and $\sigma(\Delta z)$ are explained in Section 5.3 and Appendix A.

ANN(D1) method uses u, g, r, i, and z magnitudes and the concentration indices of all five bands for the determination of photometric redshifts. A concentration index of a passband i is defined as the value of the ratio between the radii encircling 50% and 90% of the Petrosian flux. The *CC2* approach on the other hand is based on the $u-g$, $g-r$, $r-i$, and $i-z$ colors and the concentration indices in the g, r and i bands. Photometric redshift results are shown in Figures 5.6 (*D1*) and 5.7 (*CC2*), respectively.
The *ANN(D1)* performs excellently in terms of scatter and bias from $z = 0.01$ to $z = 0.3$, but it yields a considerably larger bias than our redshifts for $z > 0.3$. Photometric redshifts from the *CC2* method are significantly overestimated for $z \leq 0.08$, where the bias exceeds 0.01. Additionally, the σ_{NMAD} is particularly high at small spectroscopic redshifts ($z \leq 0.06$), which is probably due to the overestimation in this range. At $z \approx 0.3$, Δz surpasses 0.01, whereas at intermediate redshifts, the photometric redshifts from the *ANN(CC2)* are slightly overestimated.
We conclude that `PhotoZ` with the new set of LRG templates delivers results competitive with empirical approaches.

5.5 Properties of the Novel Template Set

In this section we want to discuss the properties of our new template SEDs in the UV wavelength range and analyze how well they represent objects with UV data. Being able to match the UV colors of observed objects justifies that we discuss the novel SEDs from a physical point of view and with regard to the underlying galaxy population in Section 5.5.2.

5.5.1 UV colors of the Novel Templates and M09 Models

We show color–magnitude and the color–color relations of our final template SED set down to UV wavelengths within $0.05 \leq z_{\mathrm{spec}} \leq 0.12$ and compare our findings to those of Ree et al. (2012, hereafter R12). We want to evaluate if the novel models are able to represent the *Galaxy Evolution Explorer* (GALEX, Martin et al., 2005) data, since we used no information about the UV in the creation of our SEDs, as even the GALEX NUV filter central rest frame wavelength is $\lambda_{\mathrm{NUV,\,central}} = 2544\,\text{Å}$ at a redshift of 0.12, whereas the SDSS u band filter has a much larger central wavelength of $\lambda_{u,\,\mathrm{central}} = 3546\,\text{Å}$.

R12 selected ETGs in the SDSS data on the basis of morphology, requiring `fracDeV_r` ≈ 1 (see App. E), and removed contamination by late-types through visual inspection. The sample galaxies were then classified into quiescent, star-forming, composite, and active galactic nucleus (AGN) categories on the basis of their emission line characteristics. Quiescent ETGs were selected in such a way that they do not exhibit emission lines.
Figure 5.8 shows the data medians and their spread of the four early type galaxy (ETG) categories of the galaxy sample of R12 in color–magnitude ($u-r$ vs. M_R and FUV $-$ r vs. M_R) and color–color diagrams ($u-r$ vs. $u-g$ and FUV $-$ r vs, FUV $-$ NUV) in analogy to Figure 1 in R12. Additionally, we plot the colors and magnitudes of our final templates after fitting them to the spectroscopic redshift of the LRGs within $0.05 \leq z \leq 0.12$, as well as their intrinsic colors within that z range. The median values of $u-g$ and $u-r$ coincide with those of the quiescent and AGN ETGs of R12 (lower left panel), and we are able reproduce the observed UV-colors. The median of (M_R) on the other hand is about 0.4 mag lower in this work when compared to the quiescent and AGN ETGs of R12 (upper panels), visualizing

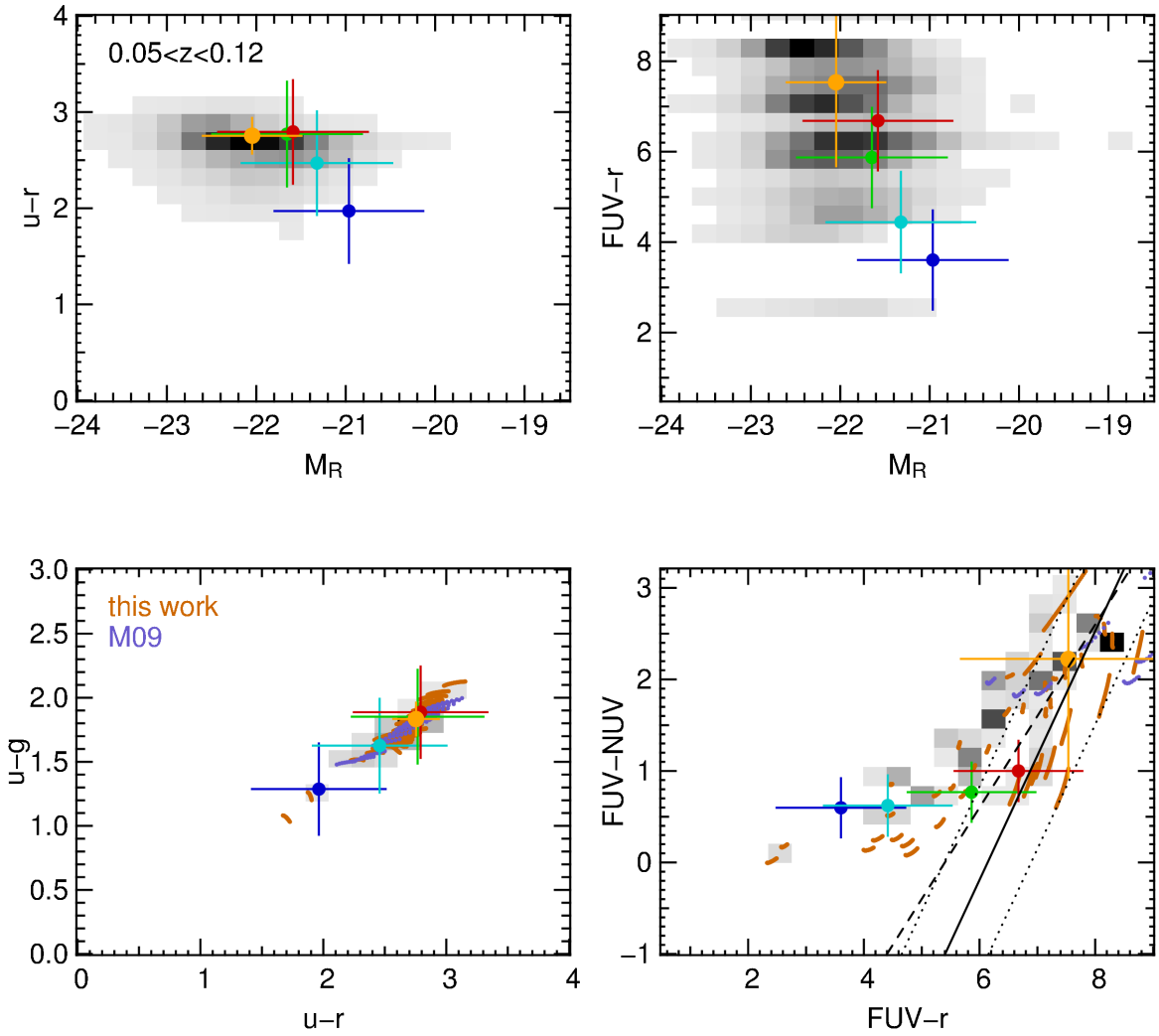

Figure 5.8: UV colors of the new template SEDs from a fit to the spectroscopic redshift of the LRGs within $0.05 \leq z_{\rm spec} \leq 0.12$ (gray/black). Yellow filled circles with error bars indicate the median colors and magnitudes with standard deviations. Medians and spreads of R12 are displayed in red (quiescent), green (AGN), cyan (composite) and blue (star-forming) ETGs. The orange sequences of dots joining a line segment in the lower two panels indicate the new models from SED fitting, whereas those in purple illustrate the colors of M09 models, both within the same redshift range. The solid line represents the UV red sequence of R12, the dotted lines are $3 \times$ MAD. The criterion of RSF galaxies (NUV $- r < 5.4$) found by Schawinski et al. (2007) is displayed by a dashed line. Plotting ranges correspond to those of R12 (Fig. 1).

that the objects in our sample are definitively LRGs. Furthermore, the median UV excess (FUV $- r$) of SEDs of this work is significantly smaller than that of R12 (~ 1 mag compared to the quiescent galaxies; upper right panel). The lower right panel displays the UV-excess versus FUV $-$ NUV, as well as the UV red sequence found by R12 (solid line) which is derived empirically through a linear fit to the observed colors with width $3 \times$ MAD (dotted lines). A fraction of our models however lie above the UV red sequence. Also the FUV $-$ NUV color is higher than in R12, due to the fact that our models have higher fluxes in the NUV band and lower fluxes in FUV. According to Schawinski et al. (2007), recently star forming (RSF) galaxies fulfill the semi-empirical criterion NUV $- r < 5.4$, which is plotted in Figure 5.8 (dashed line). A considerable part of our models is thus classified as RSF.
The lower panels of Figure 5.8 show M09 models on top of the SEDs of this work. In comparison to R12, M09 models neither cover the region of higher UV-excess (FUV $- r \lesssim 6$), nor those where FUV $-$ NUV $\lesssim 1.7$, and nearly all of them exhibit no RSF following the criterion of Schawinski et al. (2007). M09 models account for passive evolution only, which is why they do not populate these RSF regions. However, the M09 SEDs match the $u - r$ versus $u - g$ colors of quiescent, AGN and composite ETGs from R12. Nevertheless, at low redshifts the LRG sample consists not only of passively evolving galaxies, but shows signs of recent star formation. M11, which have a large variety of SFHs, ages, and metallicities, should be able to cover these regions.

5.5.2 Differences in the SEDs within $z - M_R$ Bins

By construction our new template SEDs do not necessarily consist only of old CSPs (which are necessary to match the 4000 Å break of the LRGs), but can have imprints of young stellar populations (i.e., more UV-excess relative to old populations) and dust extinction, changing the slope of the CSP and the burst. Instead of examining what kind of combination the LRGs are fitted by, we present how different the SEDs of galaxies are when matched by our templates. Therefore, we determine the best fitting SED for every LRG by fitting our new SED set at the spectroscopic redshifts of the objects and calculating the absolute magnitude in the R band filter thereafter. We then bin objects within the z versus M_R plane (Figure 5.9) and calculate the mean SED of that bin, by adding up all best fitting SED templates of objects from that bin with weights given by their flux in R. Doing this for different redshift and luminosity bins, we obtain the luminosity weighted light distribution of all objects within a given redshift and luminosity range, under the assumption that they are well described by one of the SEDs from our template set.

The results for luminosity bins with $z \leq 0.1$ are shown in Figure 5.10. The SEDs which match the SDSS data best exhibit a strong dependence on the brightness of the object: The fainter galaxies are, the more flux they have at $\lambda \lesssim 3500$ Å relative to the brightest $M_R \sim -24$ bin, and the steeper the slope of the SED gets for $\lambda \gtrsim 5000$ Å on average. Both observations can be explained by star formation in these galaxies, i.e., that low mass LRGs still build stars at low z in contrast to the more massive LRGs. Hence, the diversification in the $u - g$ color is not reached by simply selecting smoothly varying SFHs, but rather by adding to these single starburst events of different amounts. We will confirm this result in Section 5.5.2. Figure 5.11 shows the superposition of the models for $z \geq 0.1$ and $-24.5 \leq M_R \leq -22.7$ in six consecutive redshift bins. At blue wavelengths ($\lambda \leq 3000$ Å) the SEDs from objects at higher redshifts exhibit more flux than those at lower redshifts, which indicates increasing star formation turning on to higher z. These effects are in accordance to what one would expect

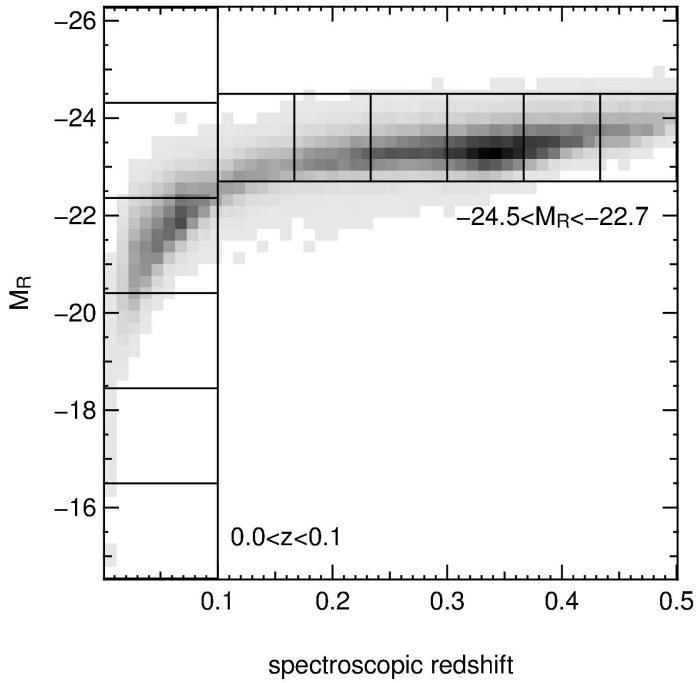

Figure 5.9: Density plot of spectroscopic z vs. absolute R band magnitude. M_R is obtained from the template SED which matches a galaxy best at its spectroscopic redshift. The solid lines show bins within z or M_R intervals. The luminosity weighted means of the template SEDs that match each object in a given $z - M_R$ bin best are shown in Figures 5.10 and 5.11.

for older stellar populations at low z and younger objects at higher z. Older red populations have (nearly) completed star formation and hence show low fluxes in bluer bands, whereas their fluxes in the red ($\lambda \geq 6000\,\text{Å}$) exhibit a smaller continuum slope compared to younger galaxies. The SED from the first redshift bin ($0.1 < z_{\text{spec}} < 0.17$) is slightly bluer than the following ($0.17 < z_{\text{spec}} < 0.23$), because there are still star forming objects within that bin. This can be seen in Figure 5.9: there are still fainter galaxies in this bin, wherefore not all of them are passively evolving. Additionally, Figure 5.12 in Section 5.5.2 confirms that there are still star forming galaxies in this z/M_R-range.

To support the latter discussion quantitatively, we perform a χ^2 fit to the variety of BC03 models that are the basis of our template generation in Section 5.1. We keep the metallicity fixed at $Z = Z_\odot$ as well as $\tau = 0.0\,\text{Gyr}$ to avoid degeneracies between Z, τ, and age. (We discussed these degeneracies in Section 4.2.2.) The best fitting ages for the redshift–luminosity bins at $z < 0.1$ are 9, 9, 7, 4.5, 4.5, and 3.5 Gyr from brighter to fainter bins. Keeping only the metallicity fixed at $Z = Z_\odot$, we observe that the best fitting models nearly all have $\tau = 1.0\,\text{Gyr}$, which supports the claim that they all exhibit ongoing star formation. The first (most luminous) bin has a best fitting τ of 0.0 Gyr, which is due to the very red SED we calculated for it. Anyway, since this bin contains only one object we cannot assume the SED to be representative, although it would be in good agreement with the sequence the superpositioned SEDs form, and it would not contradict the observation we make in Section 5.5.2 where we deduce from the relations between emission lines that no star-forming galaxies lie in this particular $z - M_R$ range.

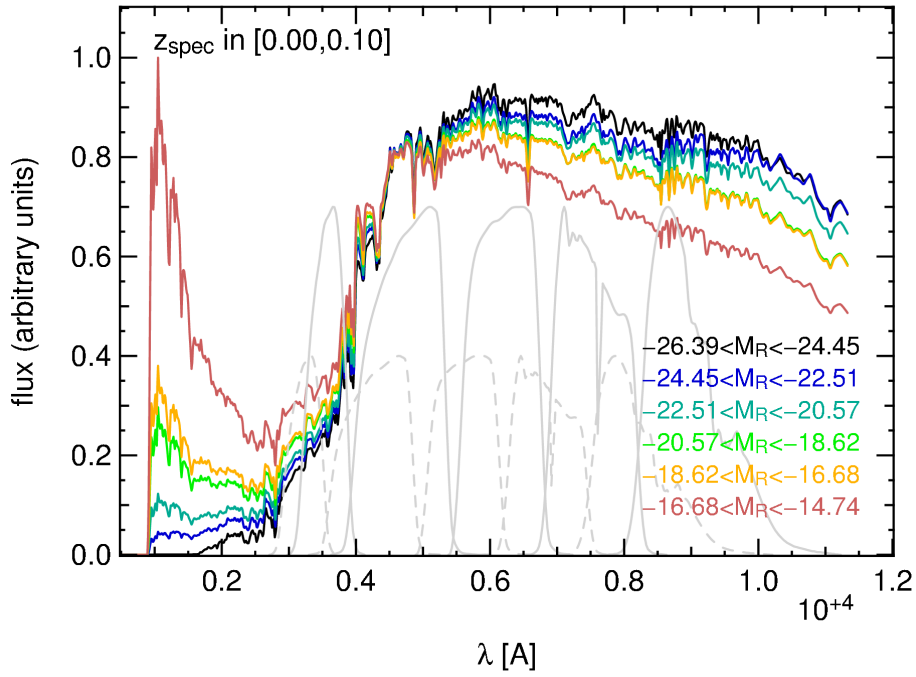

Figure 5.10: Flux weighted superposition of model SEDs (normalized to f_g) for objects with $0.0 \leq z_{\rm spec} \leq 0.1$ within consecutive luminosity bins. The solid gray lines are the normalized SDSS filter curves, whereas the dashed gray lines represent the filter curves at $z = 0.1$. The bins contain 1, 6943, 23,501, 1757, 123, and 33 objects (from higher to lower luminosities).

We also performed the χ^2 fitting of BC03 models to the superpositioned SEDs for $z > 0.1$ and $-24.5 \leq M_R \leq -22.7$ keeping Z at $Z_\odot$ and $\tau = 0.0\,$Gyr fixed. The best fitting ages are then: 12, 13, 11, 9, 9, and 9 Gyr with increasing z. As was mentioned above, the first redshift bin ($0.1 < z_{\rm spec} < 0.17$) still contains objects with ongoing star formation, which is why the best fitting age is lower than in the second bin ($0.17 < z_{\rm spec} < 0.23$). Nevertheless, the ages from the other five redshift slices are monotonically decreasing. By keeping only the metallicity fixed, the τ values increase from lower to higher redshift, which also yields to bluer spectra. Furthermore, the differences in the light travel time between the centers of the first and the last redshift bins is $\sim 3.1\,$Gyr, which is in agreement with the χ^2 fit results, and also supports what we expect when we assume that the underlying population of our sample at $z > 0.1$ predominantly consists of passively evolving galaxies, and that the variations in the SEDs are a result of aging.

Comparison to SDSS Star Forming LRGs

We want to compare our findings of the previous chapter to the observed spectra of SDSS LRGs. Therefore, we download the SDSS spectral subclassifications of our LRG sample that were produced by the `SpecBS` pipeline, where galaxies are further categorized by their emission lines. Galaxies that fulfill

$$\log_{10}(\mathrm{OIII/H}\beta) < 0.7 - 1.2(\log_{10}(\mathrm{NII/H}\alpha) - 0.4)$$

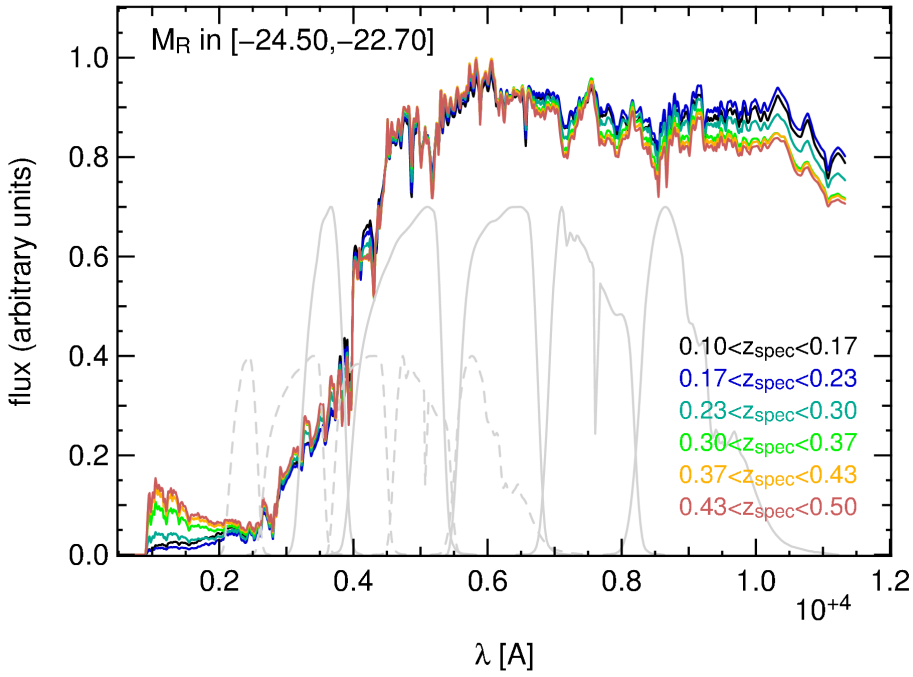

Figure 5.11: Flux weighted superposition of model SEDs (normalized to f_g) for objects with $-24.5 \leq M_R \leq -22.7$ within consecutive redshift bins. The solid gray lines are the normalized SDSS filter curves at $z = 0.1$, whereas the dashed gray lines represent the filter curves at $z = 0.5$. The bins contain 7927, 14,750, 20,328, 30,793, 19,648, and 8702 objects (from lower to higher redshifts).

are tagged as *"star forming"*. If they furthermore have an Hα equivalent width of equal or greater than 500 Å, they are classified as *"starburst"*. In our LRG sample only 10 objects fulfill the "starburst" criteria. In the following, we will refer to both tags ("star forming and "starburst") as *star forming*. Since the spectral wavelength coverage of SDSS is [3800 Å, 9200 Å], the Hα line drops out of the sample at $z > 0.4$. From the previously investigated $\sim$ 140,000 objects, $\sim$ 120,000 remain in the sample when we cut off LRGs that have $z > 0.4$. From these, 4000 show indications of star formation. The upper panel of Figure 5.12 shows the density of LRGs in the $z - M_R$ plane, and the lower panel displays the relative abundances of LRGs exhibiting star formation. Here, one can observe that the fraction of local star forming LRGs increases toward smaller luminosities. Additionally, we show in Figure 5.13 the percental fractions of star-forming LRGs for the same redshift-absolute R band magnitude bins that we investigated in Section 5.5.2, Figures 5.9 and 5.10. The errors in the intermediate bins are Poissonian errors. The errors of the outer two luminosity bins, which do not contain as many objects, are Binomial with a confidence level of 84 %. In principle, the highest and lowest M_R bins could be neglected, since they contain only 1 and 33 objects, respectively. Except for the region of faint galaxies (the lowest luminosity bin), the frequencies of star-forming LRGs shows an increase toward fainter magnitudes. This is in agreement with what we would expect from Figure 5.10, where we saw that the mean model SEDs for faint LRGs are bluer and show more star-forming activity, as well as with the results of the χ^2 fit we performed in the last section.

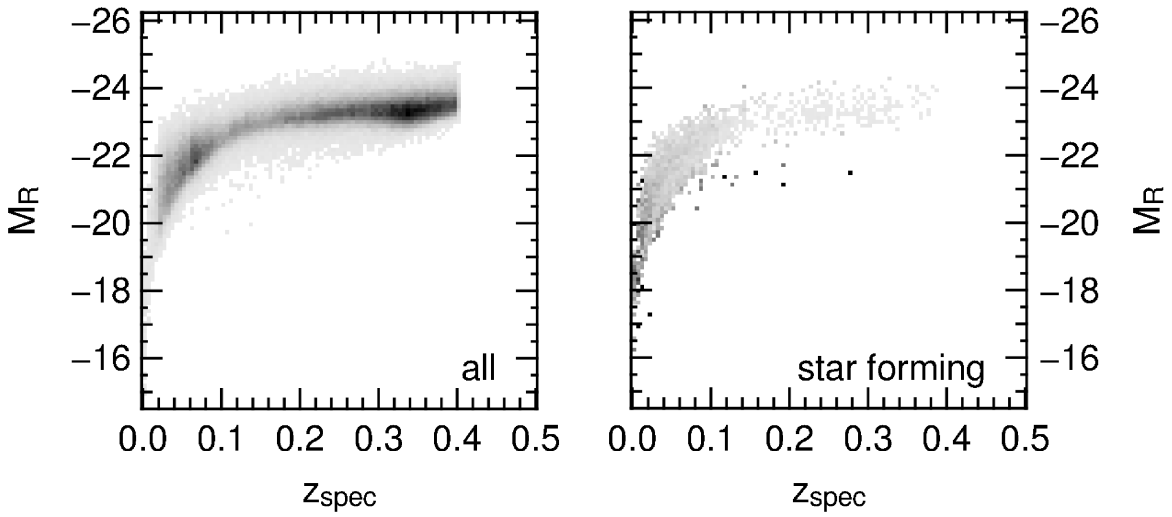

Figure 5.12: Density plots of z vs. absolute R band magnitude of LRGs. *Upper panel:* all LRGs where Hα lies within the wavelength coverage of SDSS. *Lower panel:* Relative abundance of LRGs fulfilling $\log_{10}(\mathrm{OIII/H}\beta) < 0.7 - 1.2(\log_{10}(\mathrm{NII/H}\alpha) - 0.4)$ and hence showing star formation activity.

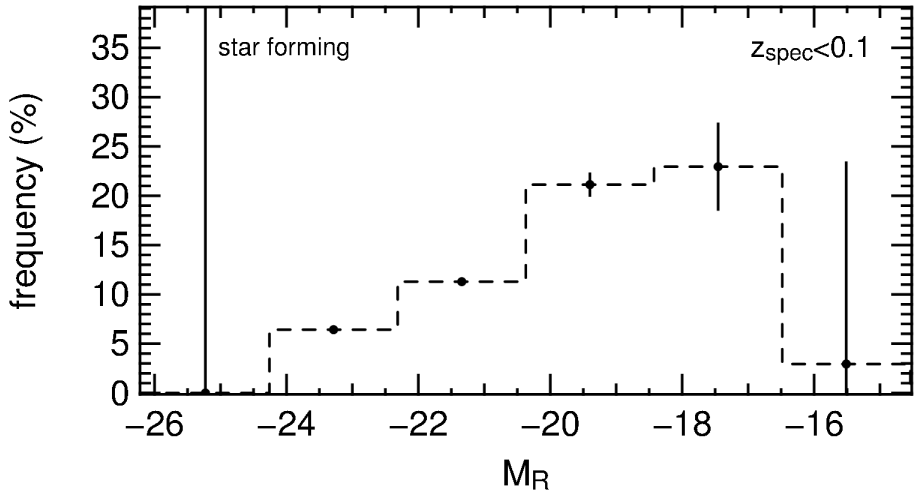

Figure 5.13: Percental fraction of $z < 0.1$ star forming LRGs as a function of luminosity. Bins are the same as in Figure 5.9 and 5.10. For a definition of "star forming" see the text. Error bars indicate the Poissonian errors for the intermediate bins, and Binomial errors (with 84 % confidence) for the outer bins which are not as highly populated. The bins contain (from higher to lower luminosity) 1, 6917, 23,429, 1751, 122, and 34 objects in total (star-forming and non-star-forming).

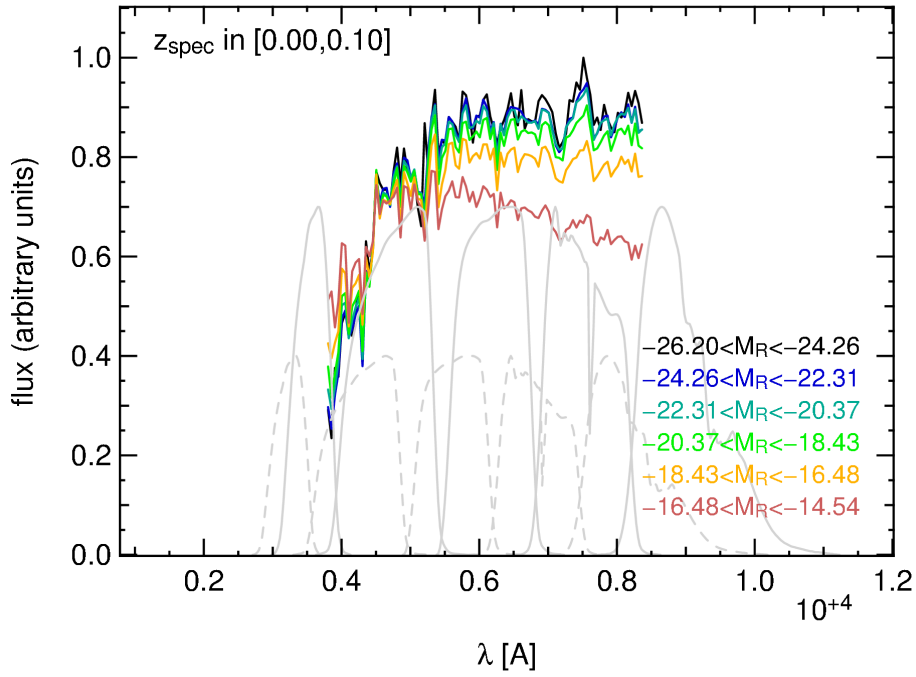

Figure 5.14: Stacked rest frame spectra for SDSS LRGs at $z < 0.1$ and within the same $z - M_R$ bins as in Figures 5.9, 5.10, and 5.13. The SEDs were normalized to f_g, interpolated to a resolution of 50 Å, and weighted by the calculated f_R (likewise as those in Figure 5.10) in the stacking. Different colors represent different M_R bins. The spectra are shown within the combined SDSS wavelength coverage for objects between $z = 0.0$ and 0.1, whereas plotting limits are the same as in Figure 5.10. SDSS filter transmissions are displayed by solid ($z = 0.0$) and dashed ($z = 0.1$) gray lines.

In addition to that we downloaded spectra for random subsamples of the objects with $z < 0.1$ and stacked them. Thereby we weighted them by the best fit SED flux in the r band as we did with the superposition of our models in Section 5.5.2. In each luminosity bin the spectra of a subsample containing 100 objects (less, if the bin contains less objects) were superpositioned. We show the results in Figure 5.14, where we consider the added spectra in the rest frame wavelength range which is covered by SDSS for objects with $z < 0.1$. Comparing Figure 5.14 to Figure 5.10, we observe that the stacked SDSS spectra show a very similar behavior as the average best fitting SEDs that were inspected previously, supporting that our novel SEDs are able to describe the data of this population.
At greater redshifts ($z \gtrsim 0.1$), no faint galaxies are observable, and only 370 ($\sim 0.5\,\%$) LRGs are classified as star forming. This region is therefore strongly dominated by objects that show no recent star formation. The variations we saw in Figure 5.11 are not caused by different fractions of star forming galaxies (which are anyhow of very low abundance), but result from aging only, which was also supported by the χ^2 fit to BC03 models we performed in Section 5.5.2.
We also stacked spectra of objects from these $z - M_R$ bins in the same manner as above. Unfortunately, one cannot see the effect of a flatter NIR and declining UV flux at low z compared to a declining NIR and increasing UV flux at high z as clearly as in Figure 5.11, because the NIR part is only observable for objects with low z, and the UV only for objects with high z. However, the trends at low z (i.e., flat NIR flux) and at high z (increasing UV

flux) are qualitatively there.

5.6 Summary

In Chapter 4.2.2 we analyzed the colors as a function of redshift of a number of publicly available red model SEDs in the SDSS filter system and compared them to the colors of spectroscopic SDSS LRGs at redshifts from $z = 0$ to 0.5. We found that no investigated model set can describe the $u-g$, $g-r$, $r-i$, and $i-z$ colors of the data at all redshifts. Thus we created new red model SEDs for five redshift bins on the basis of the LRG photometry by fitting a superposition of a BC03 CSP model and a burst SED, while allowing for dust extinction of both components. We estimated photometric redshifts with the Bayesian template fitting code `PhotoZ` and a template set containing the created SEDs with according redshift and luminosity priors. The resulting catastrophic outlier rate is $\eta = 0.12\%$, whereas the scatter reads $\sigma(\Delta z) = 0.027$ and $\sigma_{\mathrm{NMAD}} = 0.017$ for non-outliers. The overall absolute rest frame error reaches $\langle|\Delta z_{rf}\rangle = 0.015$. We compared our results with photometric redshift estimates available in the SDSS database, both for the template fitting approach and for the neural networks. Our $\langle|\Delta z_{rf}\rangle$, $\sigma(\Delta z)$, and σ_{NMAD} are of the same order as the SDSS database template fitting values, although with a slightly higher catastrophic outlier rate. Nevertheless, the medium bias varies significantly less from $\Delta z = 0$ with our template SEDs. The total values of our scatter, bias and outlier rates marginally differ from the SDSS ANN(CC2) photometric redshifts and are worse than the results of ANN(D1). However, the bias of the ANN redshifts reaches higher values within some redshift ranges.
We investigated the properties of the novel SEDs and demonstrated that they also describe the variety of GALEX NUV and FUV colors well. Performing an SED fit at the spectroscopic redshift of the LRGs, we observed that the flux weighted mean SEDs of local galaxies ($z \leq 0.1$) vary notably as a function of luminosity, meaning that faint galaxies show increasing indications of ongoing star formation. The stacked spectra from the SDSS show a very similar behavior. Furthermore, we find that the fraction of star forming SDSS LRGs at $z < 0.1$ increases with decreasing R band luminosity. This is in agreement with our finding that the mean SED is bluer for faint local galaxies. We also observe a difference in the flux weighted mean SEDs as a function of z, which is an effect of aging only and not due to recent star formation. In Appendices B.1 and B.2 we showed that BC03 models cover the data better than M09 and M11 in color space, which is why we favored the BC03 relative to the M09 and M11 model sets for the creation of our new SEDs. Finally, we also showed how our new template SED colors compare to the data in the $u - g$ vs. $r - i$, $g - r$ vs. $i - z$, $g - r$ vs. $r - i$, and $r - i$ vs. $i - z$ planes.

Chapter 6

Novel SED Templates and Precise Photometric Redshifts for BOSS CMASS Galaxies

In this chapter we present the efforts made to create Model SEDs that match the BOSS CMASS sample galaxies described in Section 3.2. The content is part of a paper which was recently submitted to the *Monthly Notices of the Royal Astronomical Society* together with Stella Seitz, Niv Drory, Ralf Bender, Roberto P. Saglia, and Jan Snigula . The introduction of the article is severely shortened, since it was in most part already covered in previous chapters.

We generate again a model set which is designed to match the data in question. In Chapter 5 (and Greisel et al., 2013, hereafter G13) we created model SEDs for spectroscopically observed luminous red galaxies (LRGs) from the SDSS-II DR7 (York et al., 2000; Eisenstein et al., 2001; Abazajian et al., 2009). The LRG sample of SDSS-II included LRGs only up to redshift $z \approx 0.5$. In this chapter we extend the work done in Chapter 5 and generate a set of model SEDs on the basis of the CMASS sample of the *Baryon Oscillation Spectroscopic Survey* (BOSS, Eisenstein et al., 2011; Dawson et al., 2013; Ahn et al., 2014) presented in Section 3.2. These templates are selected on the basis of the colors of the CMASS galaxies. They are specifically designed to match the colors of the CMASS galaxies and selected to produce accurate photometric redshifts with template fitting techniques.
In the next section we briefly expand the description of the properties of the CMASS sample of Section 3.2 and explain the selection of the CMASS catalog used in the following. We will then go into detail about the generation of model SEDs and how we modify them in order to fit the data colors to a better extent. Our goal is to derive a model set which is to be used with template fitting photometric redshift codes, whose selection we will focus on at the end of Section 6.2. In Section 6.3 we present the photometric redshift results of our new models and compare them with the two different photo-zs available on the SDSS database. We also predict the colors of our models when fitted to the more accurate one of the SDSS photo-zs in comparison to their prediction with the redshifts of this work. Finally, we give a summary and conclusions in Section 6.4.

6.1 The BOSS CMASS Sample

We use the spectroscopic BOSS data of the CMASS galaxies presented in Section 3.2 which can be acquired the SDSS `CasJobs` website[1] For the selection of our sample we employ the SDSS `clean` photometry flag. This flag ensures that we do not have duplicates in our sample by removing multiple detections on different frames. Also, objects with deblending problems are removed, as well as the ones where more than 20 % of the point spread function (PSF[2]) flux is interpolated over (that is, only 80 % of the flux is actually detected). Additionally to that, we demand that resulting objects were detected in the first pass (unbinned image), are not saturated, and that a radial profile could be constructed.
We require furthermore that objects in our sample have spectroscopic redshifts determined to a high accuracy. Therefore, we chose only objects for our catalog whose spec-z warning flag is equal to zero. We want to create models specifically designed for galaxies at higher redshifts than in G13. CMASS galaxies are furthermore very sparse at redshifts $z \gtrsim 0.8$, which is why we reduce the sample to galaxies at $0.45 \leq z \leq 0.9$. The resulting catalog then contains 486,934 CMASS galaxies. The `CasJobs` SQL query used to acquire this catalog is given in Appendix E.2.
Figure 6.1 shows the redshift distribution of the such selected CMASS galaxies. In the upper panel we plot the density of the sample in z_{spec} versus absolute magnitude in the r band. M_r is derived by fitting the LRG models of G13 to the data at known spectroscopic redshifts. The lower panel of Figure 6.1 presents the normalized frequency in spec-z (cf. Fig. 3.7).

6.1.1 Colors of M09 Models and BOSS CMASS Galaxies

In this section we want to compare the predicted colors of model SEDs to the colors of the BOSS CMASS sample. In Figure 6.2 we show the case of the M09 LRG model which is used in the definition of the color cuts of the BOSS galaxy sample. The M09 model is sampled at ages of 1.0 and 8.0 Gyr and we predict the colors while redshifting the SED (keeping the age fixed). Furthermore, we consider two passively evolving variations of the M09 model with formation redshifts of $z_f = 1.0$ and 3.0. Since the M09 model is available only at distinct ages, we have to interpolate between them to determine the SEDs as a function of redshift. The passive evolution was computed with the `EzGal` software (Mancone & Gonzalez, 2012). M09 specifically created this model to match the colors of the SDSS-II LRGs, and it fits their median colors well (cf. Fig. 4.11 in Sec. 4.2.2, or Fig. 1 and 2 in M09). This is also true for the colors of the CMASS sample, where we find that the model with $z_f = 1.0$ fits the $u - g$, $g - r$, and $i - z$ colors best in the observed cases. In the case of the $g - r$ and $i - z$ colors there are deviations from the data for $z \gtrsim 0.8$, but also the sample size of the CMASS galaxies in this region decreases significantly. However, the predicted colors in $r - i$ are too blue up to a redshift of $z \sim 0.72$ and lie outside the median uncertainties of the data. Furthermore, the color errors are too small to be responsible for the spread in color, wherefore the color values of the data is due to the galaxy population not the photometric uncertainties.
Therefore, the data cannot be matched by the M09 model simultaneously in all colors, also not when taking evolution effects into account. The spread in color can be explained by differences in ages of the galaxy, but could be also due to different stellar populations, i.e., the metallicity and distribution of stellar ages.

[1] `http://skyserver.sdss3.org/CasJobs/`.
[2] The point spread function describes the blurring of an image of a point source on a CCD.

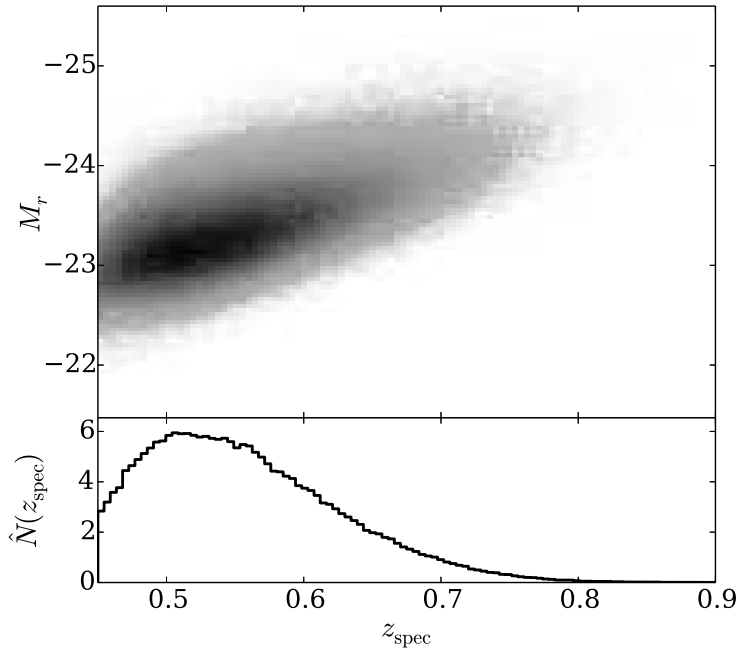

Figure 6.1: *Upper panel:* Spectroscopic redshift versus absolute magnitude in the r band derived from a fit of G13 LRG models to the CMASS catalog of this work. *Lower panel:* Normalized frequency in spectroscopic redshift of the catalog.

6.2 New SED Templates

We aim to create a set of templates that can be used for photometric redshift estimation of galaxies with similar properties as the CMASS sample. To create models specifically designed to match the colors of the BOSS data, we fit a number of model SEDs to the data at their known spectroscopic redshift and select from the best fitting models a subset that should represent the data in terms of colors, while yielding accurate photometric redshifts.
We expect the galaxy population to vary as a function of redshift. In the process of photo-z estimation with template fitting we can account for that by assigning different redshift prior probabilities to individual model SEDs. One could be tempted to use a huge amount of SEDs with different properties that can match the data in question and yield reliable photo-z results (while heavily increasing the computation time). However, this is not the case since one has to deal with degeneracies in colors resulting from different galaxy properties, e.g., age and metallicity. Also, introducing peculiar templates can deteriorate the $P(z)$ of a galaxy. Therefore, one has to carefully select a small set of templates able to match the galaxy catalog in question. In order to create a set of models that match the data at different redshifts we fit a variety of model SEDs to the CMASS galaxies within four redshift bins. The bins are centered on $z = 0.5, 0.6, 0.7, 0.8$ (continuing the sequence from G13) with interval widths of $\Delta z = 0.04$.
In the following we first give a short introduction of the SED fitting procedure and turn afterwards to the description of the models we used as a basis to construct new model SEDs. After that we explain how we select models to represent the data that should then serve as a

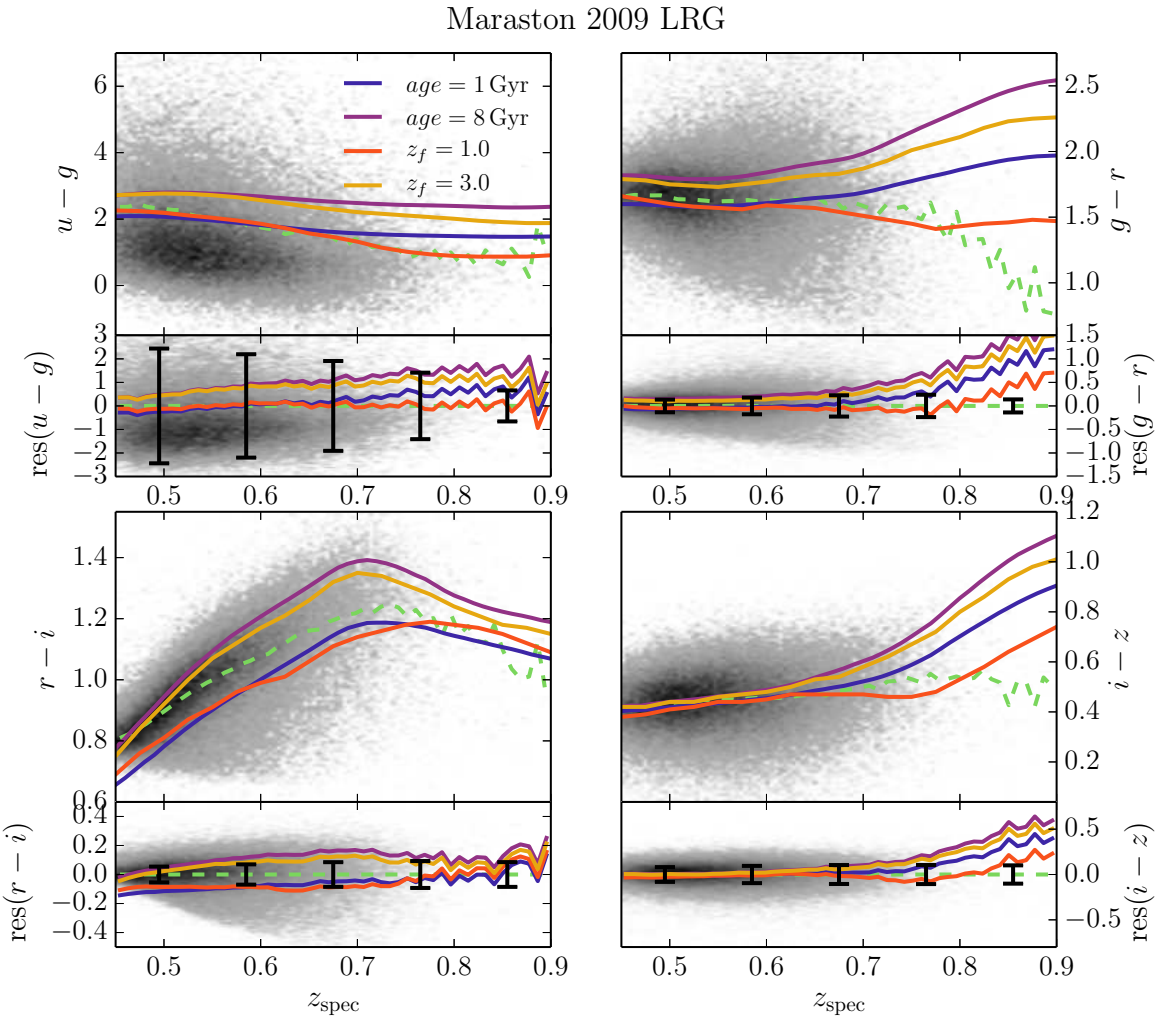

Figure 6.2: In the upper panels of the four major panels the CMASS colors are plotted as a function of redshift in gray, where the median is shown by a green dashed line. The lower panels are normalized to the median color and show the residuals. Error bars show the median data errors in five redshift bins. On top of that the predicted colors of the Maraston M09 are plotted for several ages and formation redshifts. The M09 model variations with constant ages of 1 and 8 Gyr are plotted by a blue and a violet line respectively. Furthermore, the red and yellow lines show passively evolving versions with formation redshifts of $z_f = 1.0$ and $z_f = 3.0$.

template set for photometric redshift estimation.

6.2.1 Generating Model SEDs for CMASS Galaxies by SED Fitting

To fit model SEDs to the data we use the SED fitting software `SEDfit` (Drory et al., 2004) explained in Section 5.1. The setup of the `SEDfit` code is the same as in Section 5.1. We perform the SED fitting procedure in such a way that signal-to-noise ratios smaller than 1 are considered upper limits.
The 4000 Å break is the most significant feature in the spectra of red galaxies and the reason why we can estimate photo-zs from broad band photometry to a high accuracy. At redshifts populated by the CMASS galaxies the 4000 Å break lies within the r band, wherefore the g band is the bluest band needed to derive the position of the break in wavelength. The u band is very shallow and it may deteriorate the quality of the SED fitting results and the photometric redshifts if the errors are not determined accurately. For these reasons we decide to omit it in the following SED fitting and photometric redshift estimation. Furthermore, the SED fits are not performed on the whole data of a z bin because of the high computation time, but we randomly select subsamples from within each z bin containing ~ 2000 objects.

SED Fitting Results with BC03 Models

We tried a variety of available model SEDs in the SED fitting procedure. The most extensive public libraries originate from BC03, Maraston (1998, 2005, hereafter M05), and M11. We fit all of these models in separate trial runs in order to select a basic model set to proceed further, since large sets of models require enormous computation times.
We create synthetical BC03 models from the BaSeL 3.1 library, using the recommended Padova 1994 evolutionary tracks and the IMF by Chabrier (2003). We generate models with four different metallicities, $Z = 0.2$, 0.4, 1, and $2.5\,Z_\odot$ (see also Sec. 4.2.2). From these we produce SSP, as well as CSP models. SSPs assume that all stars are formed instantaneously in a delta-functional star burst at the birth of the galaxy and are passively evolving afterwards (cf. Secs. 4.2.1 and 4.2.2). Usually, the star formation rate of a CSP is assumed that behaves proportional to $\exp(-t/\tau)$ (at least for galaxies at lower redshifts $z \lesssim 1$), where t is the age of the galaxy and τ is the e-folding timescale (e.g., Shapley et al., 2005; Longhetti & Saracco, 2009). We produce models with varying τ values, $\tau = 0.5, 1.0, 3.0, 5.0, 8.0, 20, 50$ Gyr, where the latter simulate an almost constant star formation rate (SFR). We sample the models at ages ranging from 10 Myr to 13 Gyr. In order to exploit the maximum freedom available, we also create model SEDs with increasing SFR, hence $\tau < 0$. This kind of SFH is mostly important for high(er) redshifts ($z \gtrsim 2$) and yields more physical results (e.g., Maraston et al., 2010; Monna et al., 2014, and Chap. 8), but was used for SED fitting also at lower redshifts (e.g., Pforr et al., 2012). In our case the increasing SFR models do not significantly increase the range in color space needed to match the CMASS galaxies which is why we omit them in the following to save computation time.
We create the same variety of CSPs for the M05 and M11 models (at least where possible) using the stellar population synthesis code `EzGal` (Mancone & Gonzalez, 2012), and fit them to the data as well. We show the distribution in likelihoods in Figure C.1 in Appendix C for the SED fits of BC03, M05, and M11 models to the CMASS data in the four redshift bins. The BC03 models outperform the M05 and M11 models in terms of their χ^2 values marginally at most. Figure C.1 shows that the fits with BC03 models have a slightly higher frequency of

lower χ^2 values than M05/11. Furthermore, considering Figure 20 in G13, the BC03 models are a better match to the SDSS-II LRG data in terms of χ^2 values (this is not so evident here, as seen in Figure C.1). Because of the versatility in the model creation with `GALAXEV`, and the variety of provided ages and metallicities, we choose to use the BC03 models in the remainder of this work.

In Figures 6.3 and 6.4 we present the SED fitting results with BC03 SSP and CSP models with increasing and decreasing SFRs. The data color is plotted in gray scales with according contours, and the median color errors are shown by error bars in each panel. The predicted colors of the best fitting models are plotted over the data distribution in colored contours. Both sets of contours are plotted at the same frequency levels.
The bluer $i-z$ color range populated by the CMASS galaxies in the redshift bins at $z \sim 0.5$. $z \sim 0.6$ and $z \sim 0.7$ is not populated by the best fitting BC03 models. Additionally, the redder $g-r$ colors at $z \sim 0.7$ are also not matched by the BC03 models. All these offsets in color cannot be accounted for by the photometric errors which are also given in Figures 6.3 and 6.4.
In Appendix B.1 we show in Figure B.3 that the predicted $i-z$ color of the SDSS-II LRG sample cannot be matched by the models for the highest redshift bin ~ 0.4, at least not while the other three colors are fitted simultaneously. The model colors were too red in comparison to the data, which means that the decrease in the continuum slope redwards of the 4000 Å break of the model SEDs is not strong enough. The mismatch in $i-z$ is also present for the CMASS sample which is located at even higher redshifts than the SDSS-II LRGs. The slope of the continuum is changed by variations in the model parameters, metallicity, SFH, and age, as well as by the amount of dust, but we have exploited all these degrees of freedom. We will explain in the following section how we modify the continuum slope to produce SEDs which match the colors of the CMASS galaxies to a better extent.

Modification of the Red Continuum Slope

The red continuum slope of SEDs follows a function proportional to $\lambda^{-\beta'}$, heavily modified by absorption lines. Therefore, we can also change the steepness of the continuum by varying β'. This is done in such a way that we multiply the SED redwards of a wavelength λ_i with $a \cdot \lambda^{-\beta}$, where a is chosen such that the fluxes of the underlying model SED and the modified SED coincide at λ_i. In this way we can change the continuum slope for $\lambda \geq \lambda_i$, wherefore the created SED on average follows $\lambda^{-(\beta+\beta')}$. We will term the such modified models "$\lambda^{-\beta}$ models" in the remainder of the text.

We exemplary show the resulting SEDs of an LRG model from G13 when modified by $a \cdot \lambda^{-\beta}$ for $\lambda \geq \lambda_i = 4000$ Å and $\beta = 0.5, 1.0, 1.5$ in Figure 6.5. We also plot the SDSS filters as positioned in wavelength in the observed frame of $z = 0.5$ and $z = 0.9$, the approximate extreme redshift values of the CMASS sample. We can see that for $\lambda \geq 4000$ Å the change in slope mostly affects the i and z band at lower redshifts. For higher redshifts $\lambda_i = 4000$ Å has moved in the r band. Therefore, we create variations of the BC03 SSP and the CSP models with decreasing SFR explained above, with $\lambda_i \in [3000, 3500, 4000, 4500, 5000, 5500, 6000]$ Å such that the variation does affect the fluxes in different filters for similar redshifts. Furthermore, the β values span a range from 0.5 to 2.0 with a step size of 0.5. Larger ranges for values of λ_i and/or β do not improve the results further, since the location in color spaces they

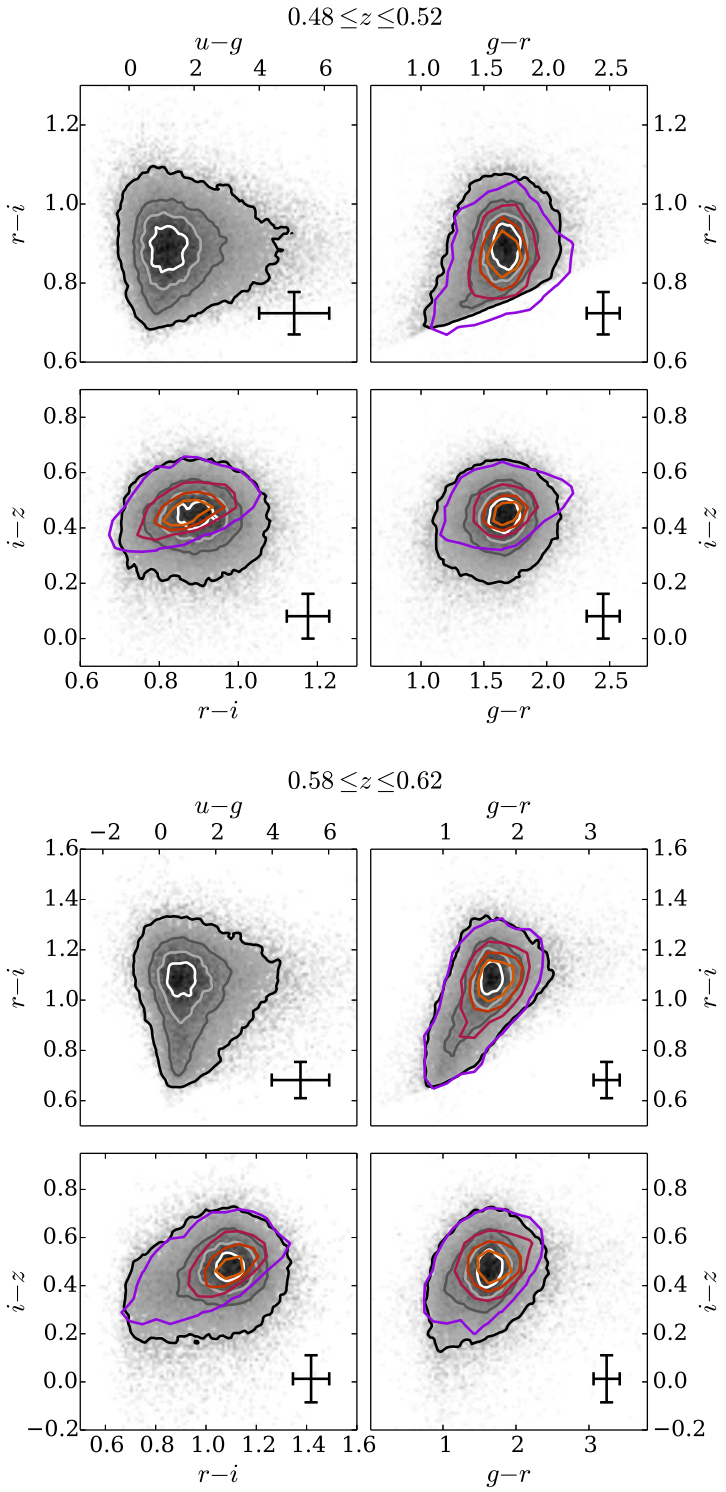

Figure 6.3: Color-color plots for CMASS galaxies at $0.48 \leq z \leq 0.52$ (upper panel) and $0.58 \leq z \leq 0.62$ (lower panel). The BOSS data is shown by gray shades and accompanying gray contours. Error bars denote the median error in color in this redshift bin. The resulting colors of the SED fitted unmodified BC03 models are shown by colored contours. Both sets of contours are drawn at the same frequency levels.

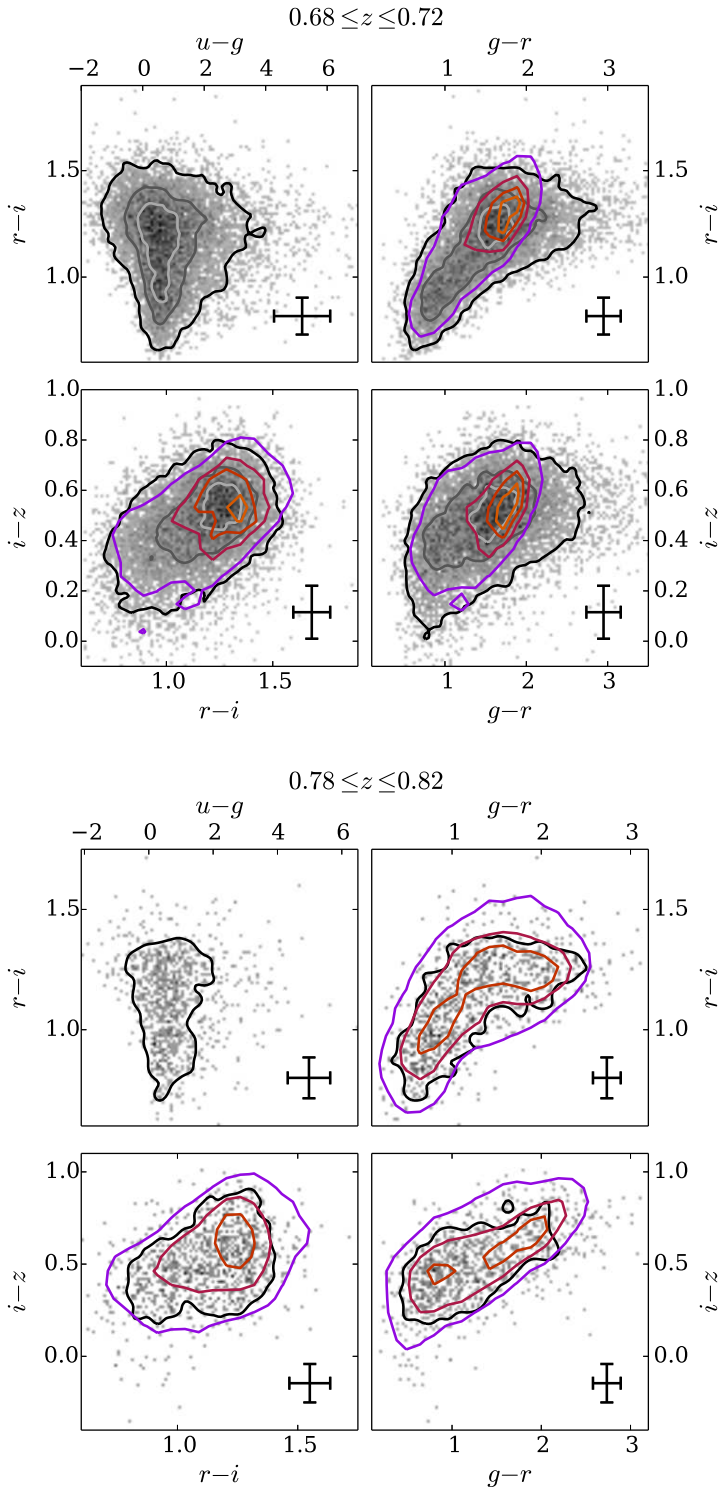

Figure 6.4: Color-color plots for CMASS galaxies and best fitting BC03 models at $0.68 \leq z \leq 0.72$ (upper panel) and $0.78 \leq z \leq 0.82$ (lower panel). See Figure 6.3 for a detailed description.

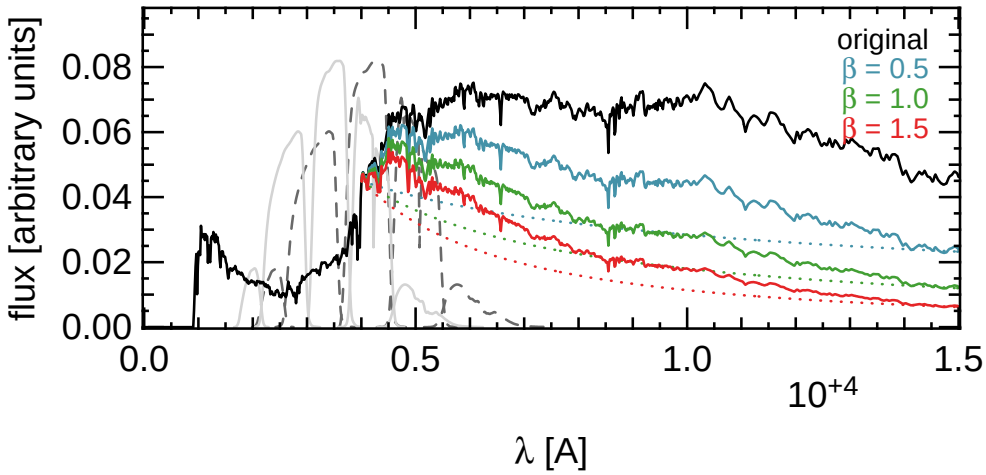

Figure 6.5: The black line is an LRG model SED from G13. The dotted blue, green, and red lines show the function $\lambda^{-\beta}$ with $\beta = 0.5, 1.0, 1.5$ for $\lambda \geq 4000$Å. The solid lines are the above SED multiplied by $\lambda^{-\beta}$, also for $\lambda \geq 4000$Å. In gray (dark gray dashed line, and light gray solid line) the SDSS *ugriz* filter curves are shown for a galaxy at $z = 0.5$, and $z = 0.8$ respectively.

would inhabit is already covered by the other models or not populated by the CMASS galaxies.

In Figure 6.6 we show the colors of the CMASS data (gray). We furthermore plot the colors of an exemplary SED from G13 (red), and the colors of the same model when modified by $\lambda^{-\beta}$, for $\beta = 0.5$ and 1 (green and blue). The data in Figure 6.6 is split into four redshift bins, and we plot the mean color and corresponding root mean square values of each bin. The points are color-coded in such a way that they darken with increasing redshift. This is shown in the gray color bar on the right of Figure 6.6, where white lines denote the mean redshifts within the bins. On top of the data, the color tracks as functions of redshift are plotted for the models. The model colors at the z values of the data are highlighted by crosses, and the tracks in the color spaces are again color-coded as a function of redshift. The track of the original model is plotted by varying intensities of red, whereas those of the corresponding $\lambda^{-\beta}$ models are plotted in green shades for $\beta = 0.5$, and blue shades for $\beta = 1$ respectively. All plotted models have $\lambda_i = 3000\,$Å. While the data colors (gray) are well matched by the original model (red track) at lower redshifts, the deviations from the mean color increase with increasing z. The modification by $\lambda^{-\beta}$ produces a better agreement with the mean data colors. The value of β necessary to fit the data best is somewhat ambiguous in this plot, but has to be decided individually for each galaxy together with the best fitting model.

SED Fitting Results with the Modified BC03 Models

We now introduce the modified versions of the BC03 models together with the original SSPs and CSPs with decreasing SFHs from Section 6.2.1 in the `SEDfit` code with the same fitting parameters as in Section 6.2.1. Figures 6.7 and 6.8 show the SED fitting results in color spaces of the CMASS data in the four z bins. We see that the contours of the data and the best fitting models (which are located at the same steps in frequency) almost coincide in Figures 6.7 and the upper panel of 6.8.

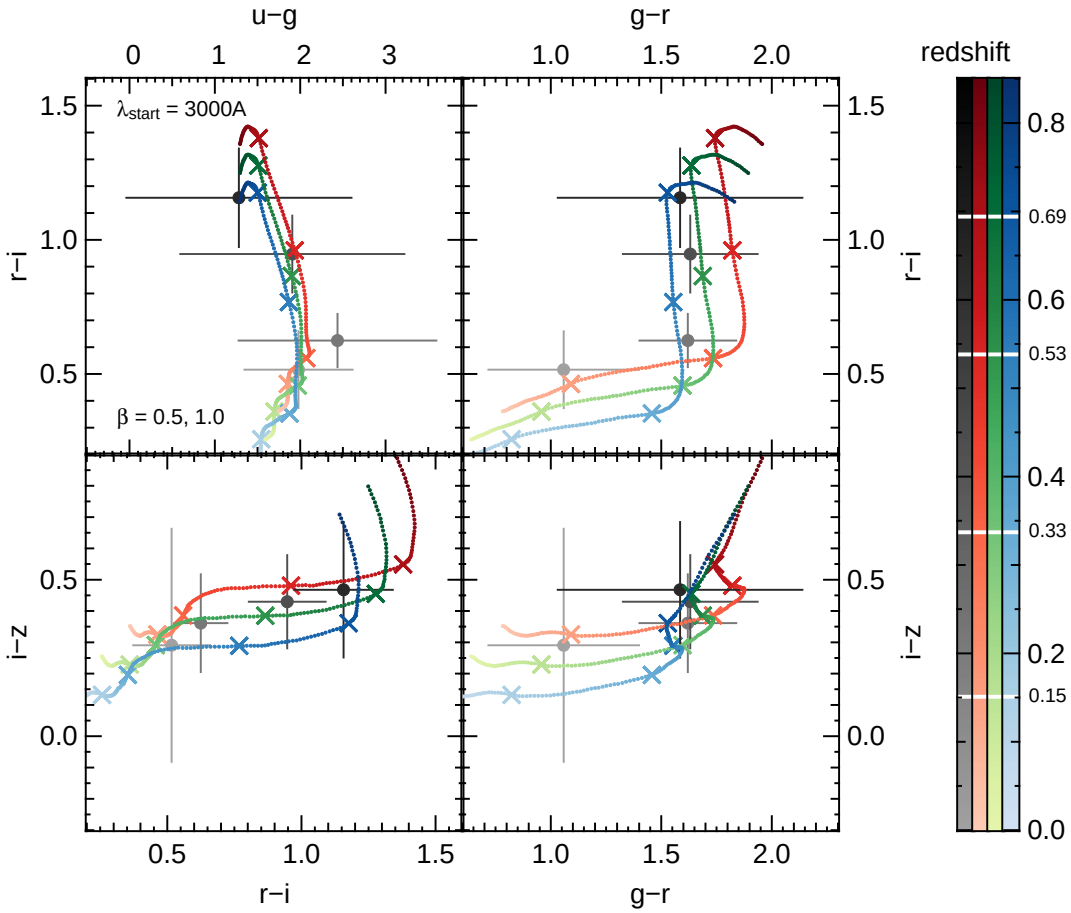

Figure 6.6: Color-color plots of the CMASS data and model SEDs. The BOSS data is split into four equally sized redshift bins. From these bins we calculate the mean color and plot it as points in the panel. The error bars denote the accompanying root mean square values of the color in each bin. The points and the error bars are color-coded with redshift (see the color bar on the right, where white lines indicate the mean redshift in each bin). The red shaded track shows the colors of the model SED of Figure 6.5 with varying redshift, where the crosses are at the same z as the data. Again, the redshift of the track is color coded and shown in the color bar. The green and blue shaded tracks originate from the same model, when multiplied by $\lambda^{-\beta}$ with $\beta = 0.5, 1$ for $\lambda \geq 3000\,\text{Å}$. A more detailed description is given in the text.

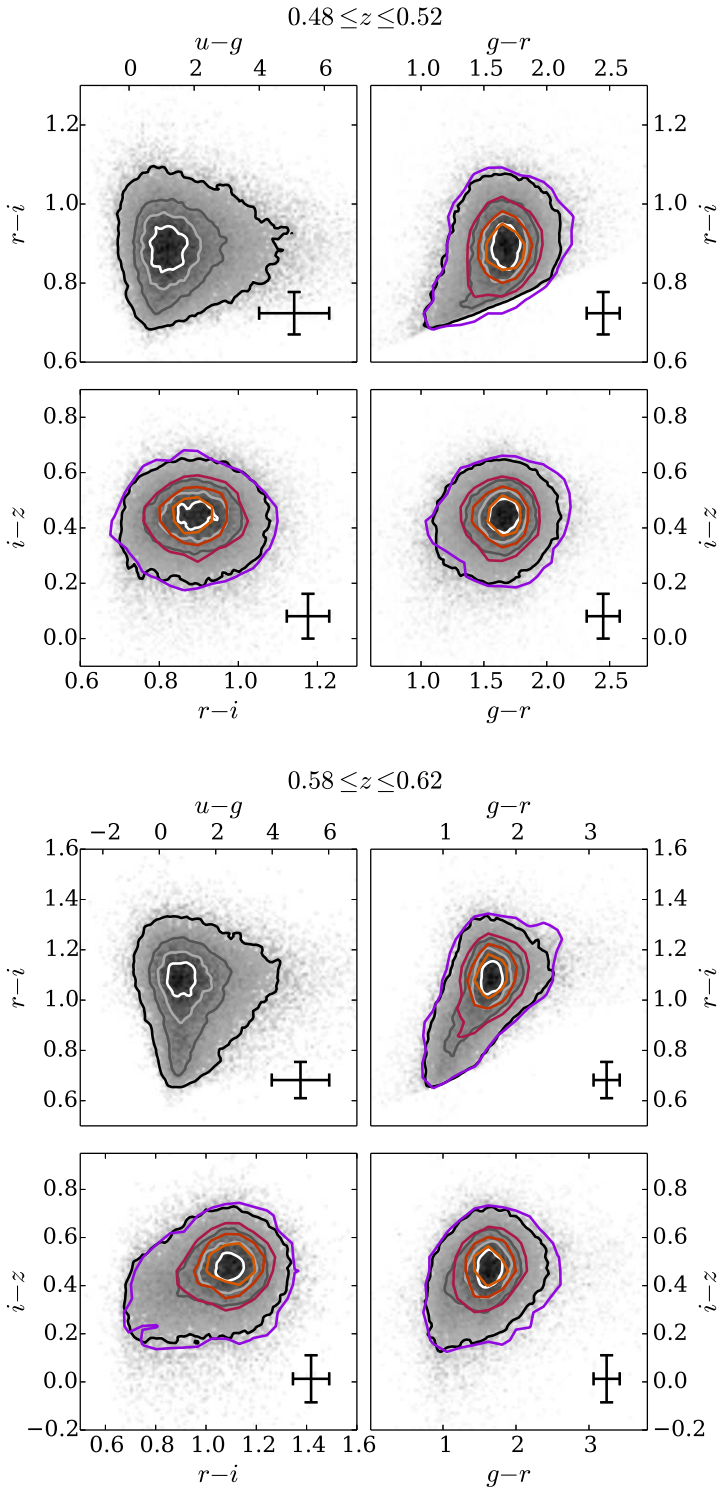

Figure 6.7: Color-color plots for CMASS galaxies at $0.48 \leq z \leq 0.52$ (upper panel) and $0.58 \leq z \leq 0.62$ (lower panel) analogous to Figure 6.3. The BOSS data is shown by gray shades and accompanying gray contours. Error bars denote the median error in color in this redshift bin. The resulting colors of the best fitting BC03 and BC03 $\lambda^{-\beta}$ models are shown by colored contours. Both sets of contours are drawn at the same frequency levels.

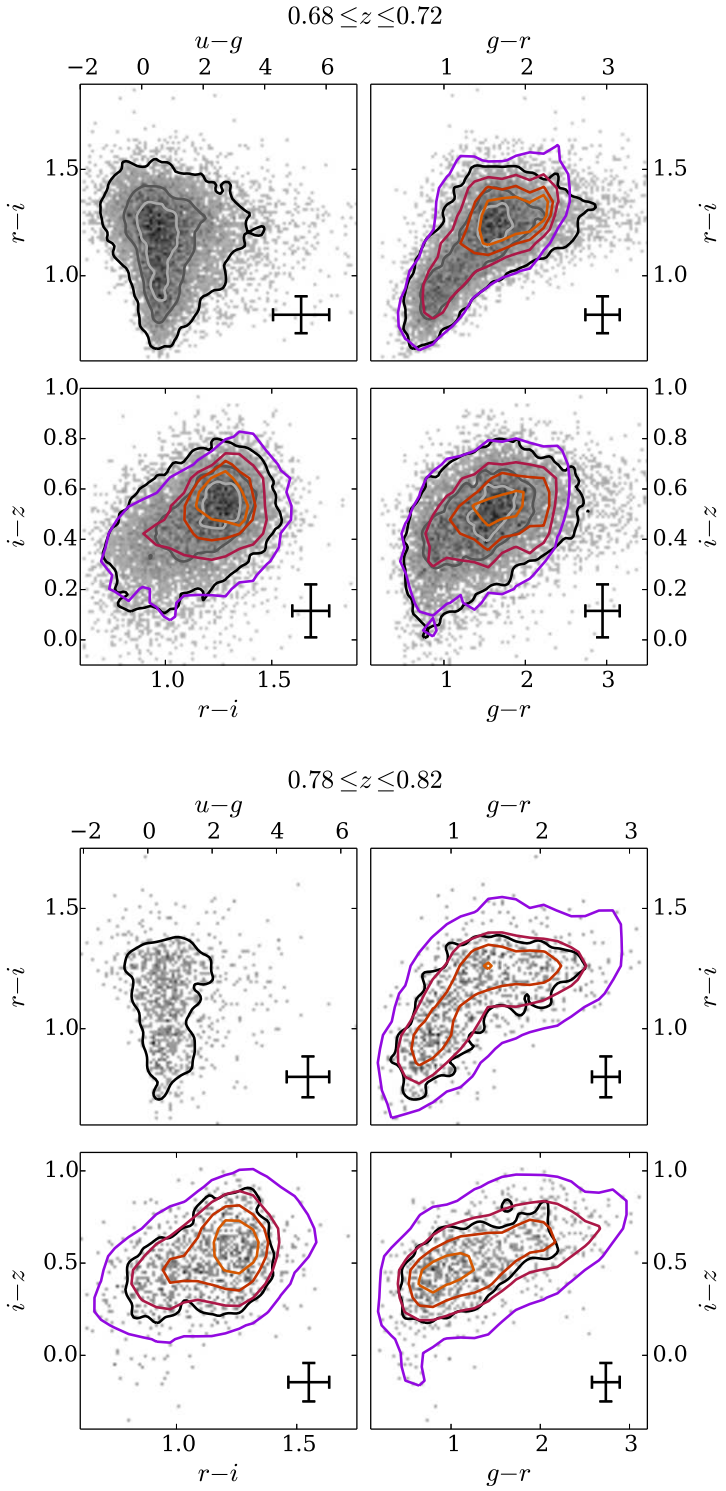

Figure 6.8: Color-color plots for CMASS galaxies and best fitting BC03 $\lambda^{-\beta}$ models at $0.68 \leq z \leq 0.72$ (upper panel) and $0.78 \leq z \leq 0.82$ (lower panel) analogous to Figure 6.4. See Figure 6.7 for a detailed explanation of the plot.

Comparing Figures 6.3 and 6.4 to Figures 6.7 and 6.8, we can observe an improvement on the match between the locations in color space populated by the best fitting models and the data. This is additionally confirmed by Figure C.1 in Appendix C, where we see that the resulting χ^2 values of the fits are more frequently found at lower χ^2 values for the BC03 $\lambda^{-\beta}$ models than for the original ones, hence improving the goodness of fit. Here, we want to specifically analyze the offsets in the $g-r$, $r-i$, $i-z$, and $g-z$ colors of the best fitting BC03 models and their modified variations from the data. We show in Figures 6.9 to 6.12 the deviations of the model colors, predicted by the best fitting SEDs, to the data, $\Delta\mathrm{col} = \mathrm{col}_{\mathrm{model}} - \mathrm{col}_{\mathrm{data}}$, for both setups and in all four considered z bins. The distribution in Δcol of the original BC03 models (SSPs, as well as decreasing and increasing SFR CSPs) is presented by red histograms, where the Gaussian curve best fitting the histogram is shown by a dashed orange line. The color offsets of the BC03 $\lambda^{-\beta}$ models are plotted by a gray histogram, with an according Gaussian drawn in black. $\Delta\mathrm{col} = 0$ is shown by a dashed black line, and we print the parameter values of the Gaussians in the plots.
While the mean deviations are more or less the same in all considered redshift bins, the standard deviation of Δcol is about 2 to 3 times higher for the original BC03 SEDs. Furthermore, the BC03 $\lambda^{-\beta}$ distributions of Δcol resemble a Gaussian much more closely than that of the original BC03, especially in the $r-i$ and $i-z$ colors, but also in $g-z$. We already pointed out that the $i-z$ color is overestimated by the BC03 models analyzed in G13 for $z \sim 0.4$. This is still true for $z \sim 0.5$ to 0.7, observable in Figures 6.9 to 6.11 (cf. Figs. 6.7 and 6.8). Concurrently, the predicted $r-i$ colors of the unmodified BC03 models are too blue up to $z \sim 0.6$, and too red in $z \sim 0.7$. In contrast to that, the colors of the BC03 $\lambda^{-\beta}$ SEDs yield very symmetric distributions, although with increased flanks in comparison to a Gaussian for higher redshifts.

In summary, we investigated two additional degrees of freedom, apart from the model SED parameters metallicity, SFH, age, burst, and dust, by modification of the red continuum slope through multiplying with $\lambda^{-\beta}$ at $\lambda \geq \lambda_i$ to match the colors of the CMASS galaxies to a better extent. When fitted to the data, the $\lambda^{-\beta}$ SEDs predict colors that deviate less from the data than the unmodified models they originate from, and yield lower χ^2 values. We will use the variety of best fitting SEDs (including the additional burst and dust components) as a basis to select from when we define a model set for photometric redshift template fitting codes.

6.2.2 Selection of Best Fitting SEDs for the New Template Set

We want to select a set of model SEDs from the best fitting models of the previous section that represent the data in terms of colors for each redshift bin. Since the color space we can construct from the SDSS colors is many-dimensional, we want to reduce the dimensions to two for simplicity of the selection process. We could, in principle, use a principal component analysis (PCA) for this task, but it has the caveat that the PCA eigenvectors are orthogonal. That means that if we chose to select models in the space which is spanned only by the two eigenvectors with the lowest variance, we could miss information carried by the third (or forth) component. In order to lose as few information as possible in the dimension reduction, we chose to map the data colors on a two-dimensional surface with a self-organizing map (SOM or Kohonen-map, Kohonen, 1982, 2001). A SOM is an artificial neural network (ANN) that provides a discrete representation of a set of higher dimensional data values in a plane (usually, but the output space can have higher dimensions too). The network is trained using

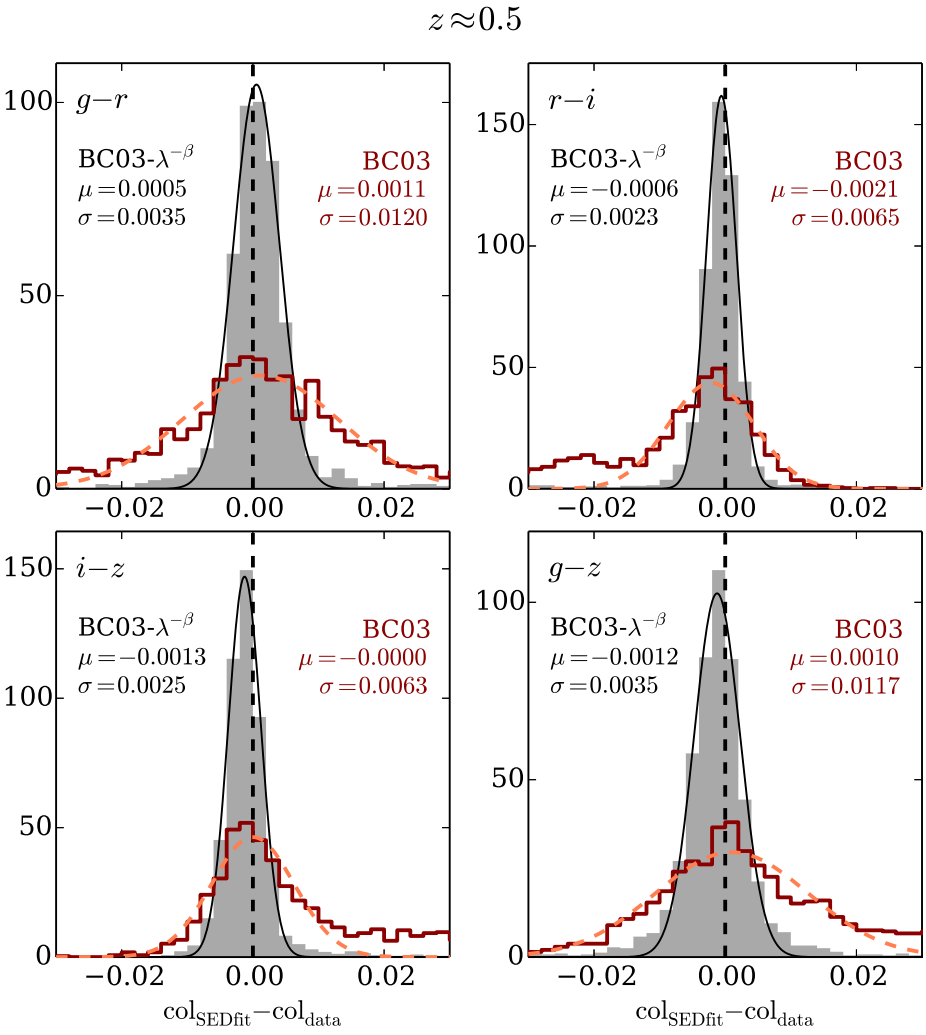

Figure 6.9: Deviations in $g-r$, $r-i$, $i-z$ and $g-z$ colors for BC03 models and BC03 $\lambda^{-\beta}$ models from the data within $0.48 \leq z \leq 0.52$. The Δcol distribution for the BC03 $\lambda^{-\beta}$ models are plotted in gray, and the best fitting Gaussian is indicated by a solid black line. The same distribution, but for unmodified BC03 models is given by a dark red line, where we plot the best fitting Gaussian by a dashed orange line. The black dashed vertical line highlights $\Delta\mathrm{col} = 0$. The best fitting parameters of the Gauss curves are given in the plots.

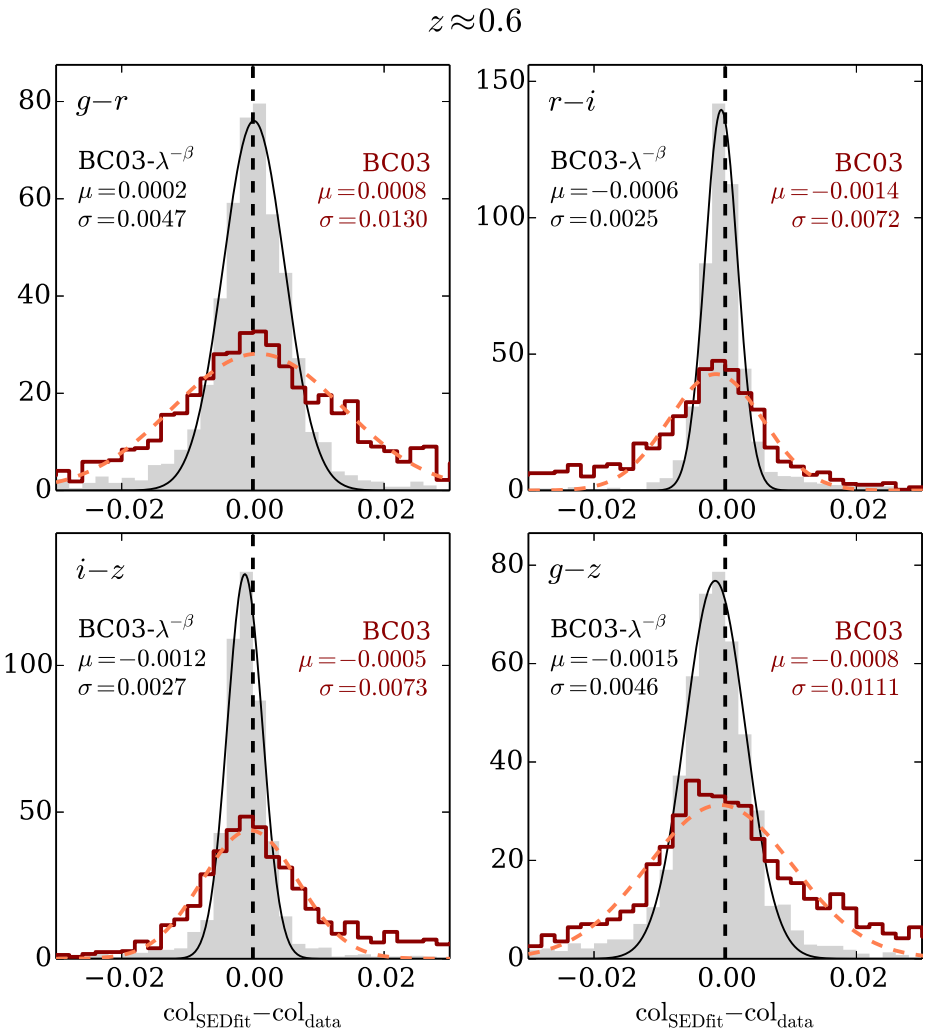

Figure 6.10: Deviations in $g-r$, $r-i$, $i-z$ and $g-z$ colors for BC03 models and BC03 $\lambda^{-\beta}$ models from the data within $0.58 \leq z \leq 0.62$. A detailed description of the plot is given in Figure 6.9.

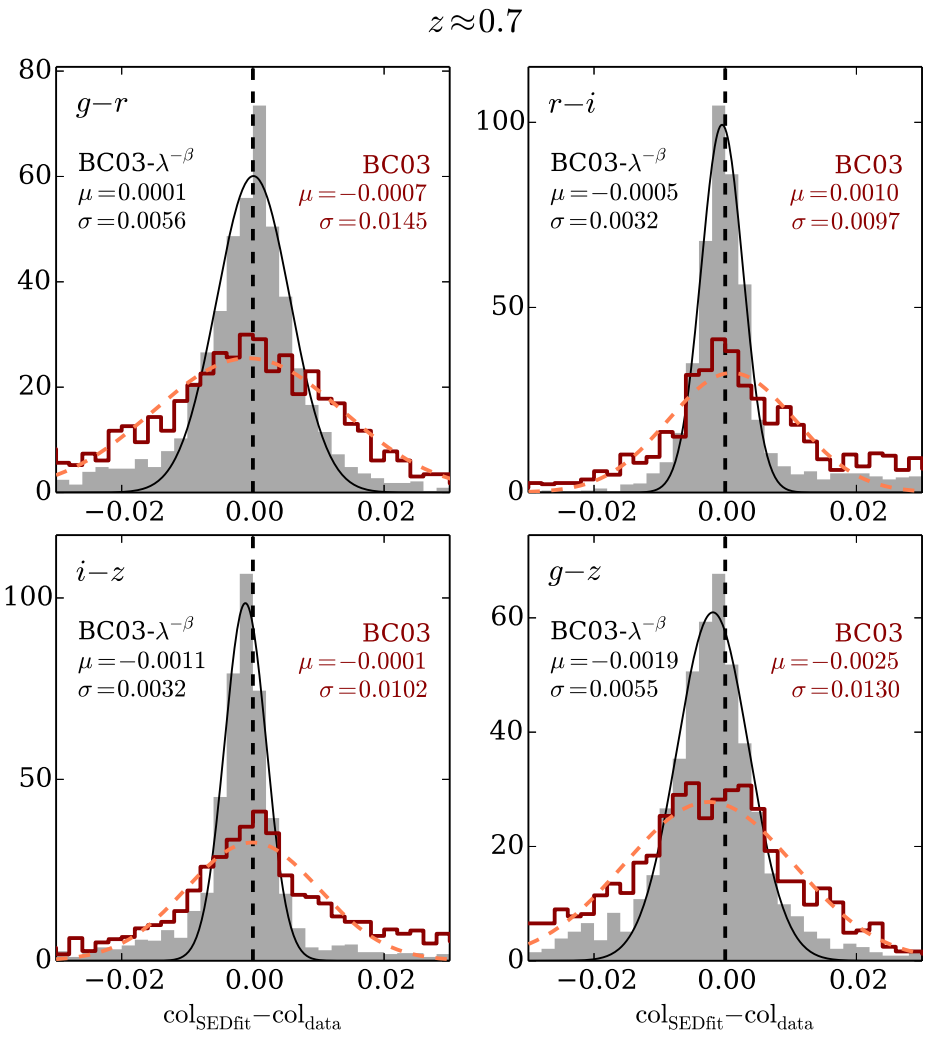

Figure 6.11: Deviations in $g-r$, $r-i$, $i-z$ and $g-z$ colors for BC03 models and BC03 $\lambda^{-\beta}$ models from the data within $0.68 \leq z \leq 0.72$. A detailed description of the plot is given in Figure 6.9.

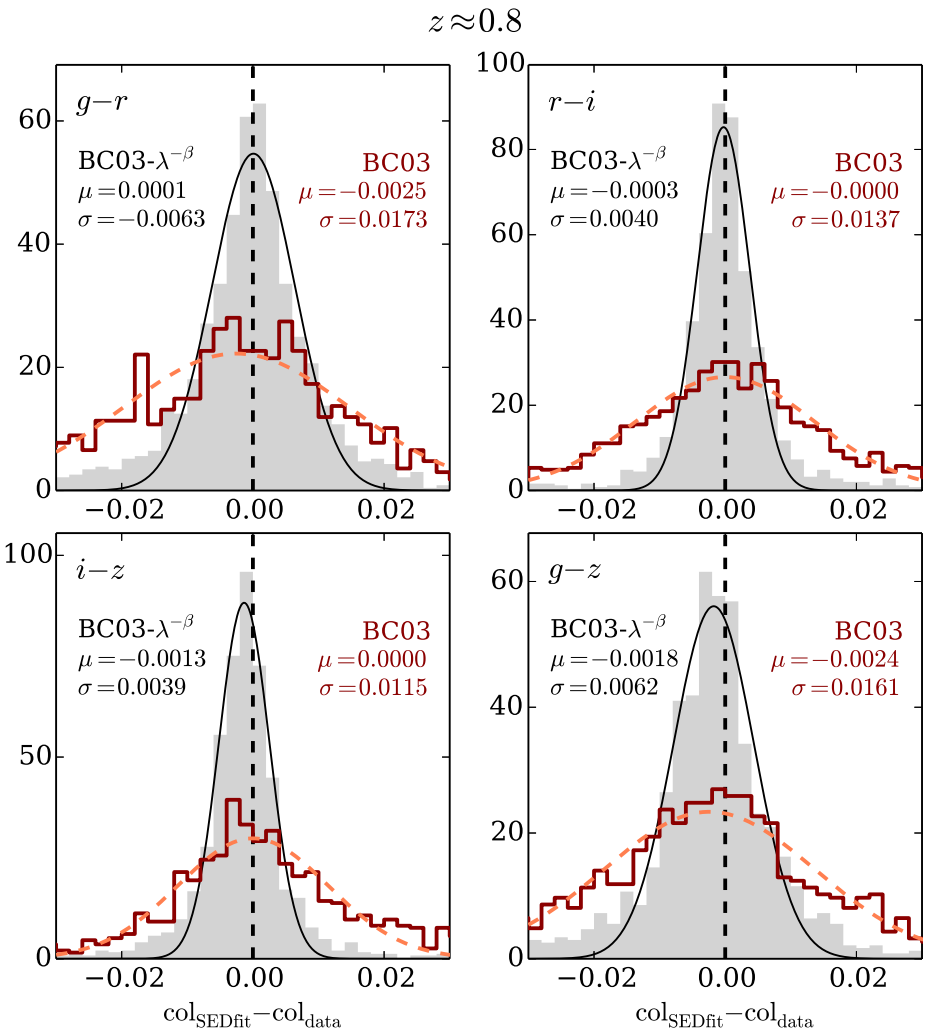

Figure 6.12: Deviations in $g-r$, $r-i$, $i-z$ and $g-z$ colors for BC03 models and BC03 $\lambda^{-\beta}$ models from the data within $0.78 \leq z \leq 0.82$. A detailed description of the plot is given in Figure 6.9.

unsupervised learning to map the data onto the two-dimensional surface in such a way that data points with similar values are located in nearby areas. Unlike the plane of the first two components of a PCA, the SOM still entails the information of the underlying many-dimensional data in its points. We create maps of the galaxy catalogs from the $g-r$, $r-i$, and $i-z$ colors of the four considered redshift bins. Furthermore, we can easily introduce also the absolute magnitudes M of the galaxies as a forth quantity to be mapped, since M can also hold information about the color (e.g., Baldry et al., 2004). We have to take care of the errors in the data by normalizing the colors to their mean value and dividing through the color errors determined by the uncertainties in the photometry. The absolute magnitudes of the galaxies are calculated by fitting the LRG model SEDs of G13 to the data at their spectroscopic redshifts. From the best fitting models we extract the absolute magnitude in the SDSS r band, which has to be normalized to a standard normal distribution to be comparable to the color values. Once this is done we train the SOM and create a surface with 30×30 discrete x- and y-values with assigned input values (i.e., colors and M_r). The positions of data points in the SOM are then determined by performing a nearest neighbor search.
We use a SOM algorithm implemented in python and provided in the `PyMVPA` package (Hanke et al., 2009). Figure 6.13 presents the SOM trained on the catalog with $0.48 \leq z \leq 0.52$ on the $g-r$, $r-i$, and $i-z$ colors and on M_r. The four panels are representations

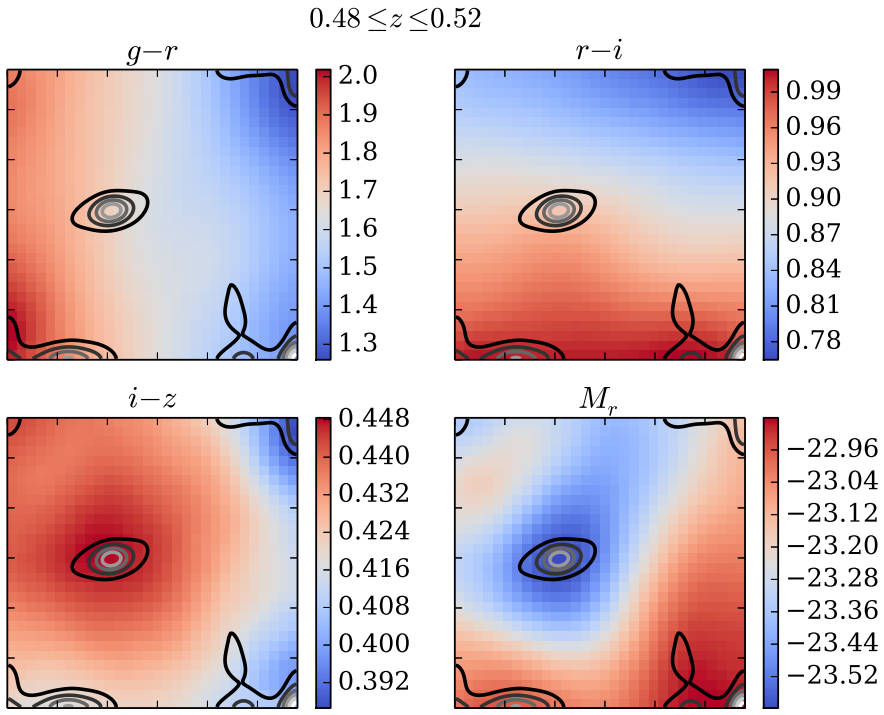

Figure 6.13: SOM with 30×30 bins for $g-r$, $r-i$, $i-z$, and M_r for galaxies within $0.48 \leq z \leq 0.52$, where the parameter values are color-encoded (see colorbars on the right of each panel). M_r is computed by fitting templates from G13 to the spectroscopic redshift. Contours show the number density N of the data in the map (note that the contours are smoothed to improve the clarity of the figure).

of the same SOM, but with the values of the four mapped parameters $g-r$, $r-i$, and $i-z$ colors and M_r encoded in color. The mapped underlying data are shown by contours. We can see from Figure 6.13 that the extreme values of $g-r$ and $r-i$ have the greatest separation in the map which is due to these colors having the largest range in values, also when taking measurement errors into account (which we did). We can see for example that the dense region in the middle of the panels corresponds to a higher concentration of the data at the respective color values. These are $g-r \sim 1.7$, $r-i \sim 0.9$, and $i-z \sim 0.45$, which represent also the areas of highest density in the color-color plots in the upper panels of Figure 6.7.

As previously mentioned, we aim to select a sample of model SEDs from the best fitting models explained in the last section, we want to take these density variations in the population of the data into account. Therefore, we want to identify clusters in the mapped data to select a model from each cluster that should represent the galaxies within the same cluster cell in terms of color and absolute r band magnitude. To perform the cluster search we employ a k-means clustering algorithm (Steinhaus, 1957) that partitions the two-dimensional maps into cluster cells. Each cell is a Voronoi bin (Voronoi, 1908), where two bins or clusters are separated from one another by a border orthogonal to the line connecting the cluster centers. Every data point belongs to the cluster with the nearest center. Specifically, we use the python k-means clustering algorithm included in the `scikit-learn` package (Pedregosa et al., 2011). Figure 6.14 presents the bins of the k-means clustering algorithm with $k = 6$ clusters applied

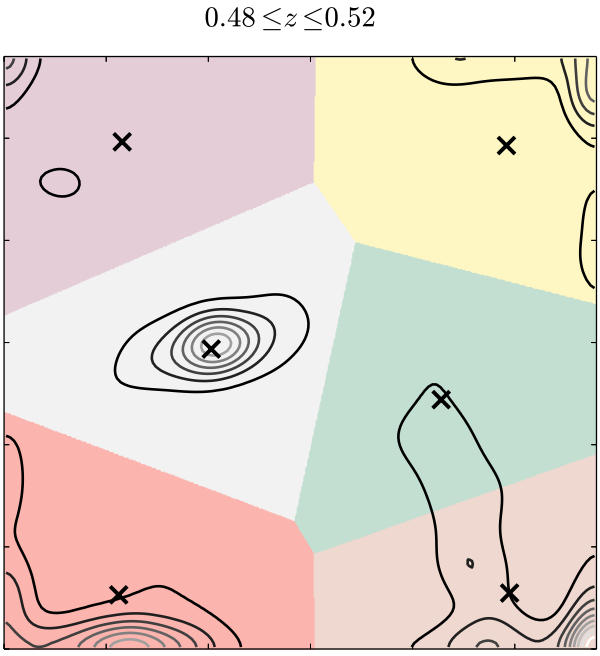

Figure 6.14: k-means clustering of the SOM for $0.48 \leq z \leq 0.52$ from Figure 6.13. The Voronoi bins are highlighted by different colors, and cluster centers are marked by black crosses. Contours show the number frequency of the data (note that the contours are smoothed to increase the clarity of the figure).

to the SOM of Figure 6.13, where the cluster centers are indicated by black crosses. The density of the data in the SOM is shown by contours in the plot.
To select a set of SEDs from the best fitting model SEDs which we will use in the following as template set for photo-zs. Therefore, for each cluster bin we take all models into account whose corresponding data points also lie in the same bin. Then, for each of these models separately, we estimate photometric redshifts on all objects within the cell and determine some resulting quality parameters, i.e., mean error, scatter and catastrophic failures of the photo-zs. For each cluster we chose a model that yielded the best photo-z results. Thereby, we have for each of the four redshift bins six models, hence 24 models in total. The such selected SEDs will be used in the following to estimate photo-zs.

6.3 Photometric Redshifts

In this section we analyze the photometric redshift results we get using our novel templates in combination with the `PhotoZ` code on the whole CMASS sample detailed in Section 6.1. Afterwards, we compare our photo-zs with the photometric redshifts provided by the SDSS database.

6.3.1 Photometric Redshifts with the Novel Template SEDs

To introduce the models created in the previous section into the `PhotoZ` code, we define the z prior such that μ_z of a model SED is the central value of the z bin of the catalog from which the model originates, wherefore $\mu_z = 0.5, 0.6, 0.7, 0.8$. As σ_z, we set a default value of 0.2 which leads to a prior $P_z(z|T)$ function wide enough to avoid focusing effects at certain redshift bin centers, while ensuring a smooth transition between them. The resolution in redshift of a photo-z run is 0.002 in the range of $z \in [0.0, 3.0]$. The allowed redshift range is much wider than that which is populated by the CMASS galaxies, such that we can analyze if the photo-z accuracy is diminished by $z_{\rm phot}$ values that are highly over- or underestimated. This is done because we would like to be able to run photo-z codes with the new models on catalogs with galaxies from larger redshift ranges and with more variations in SED type. We therefore want to make sure that small errors in $z_{\rm phot}$ are not due to a restriction in the redshift range.
To improve the priors by adapting them iteratively, we analyze the outcome for subsamples of objects which are fitted best by a specific model. Thereby we can adjust the redshift and luminosity priors for each model SED in order to reduce outliers and bias. Essentially, we decrease the value of σ_z whenever a model yields lower $z_{\rm phot}$ accuracies for redshifts further away from its bin center. Additionally, we allow μ_z to vary if the photo-z performance of a specific model can be enhanced. Furthermore, if we observe that a specific template provides very bad redshifts which cannot be resolved by adjusting the respective prior, we omit these models completely in following runs. This is mostly the case for models created from the highest two redshift bins. Since we chose the models only on account of their photo-z performance on a redshift bin, they might still yield a bad estimate in other redshift ranges. The resulting model set then consists of nine SEDs with adapted redshift priors. The luminosity priors where set initially to $\mu_M = 20.0$, $\sigma_M = 6.0$, and $p_M = 6.0$, to allow for a wide range of higher luminosities. The high exponent leads to a very flat functional behavior within σ_M, and to steep decreases in $P_M(M)$ at $\mu_M \pm \sigma_M$. Adjusting $P_M(M)$ in the iteration is not necessary, since we cannot detect outliers that are caused by a specific model which can be

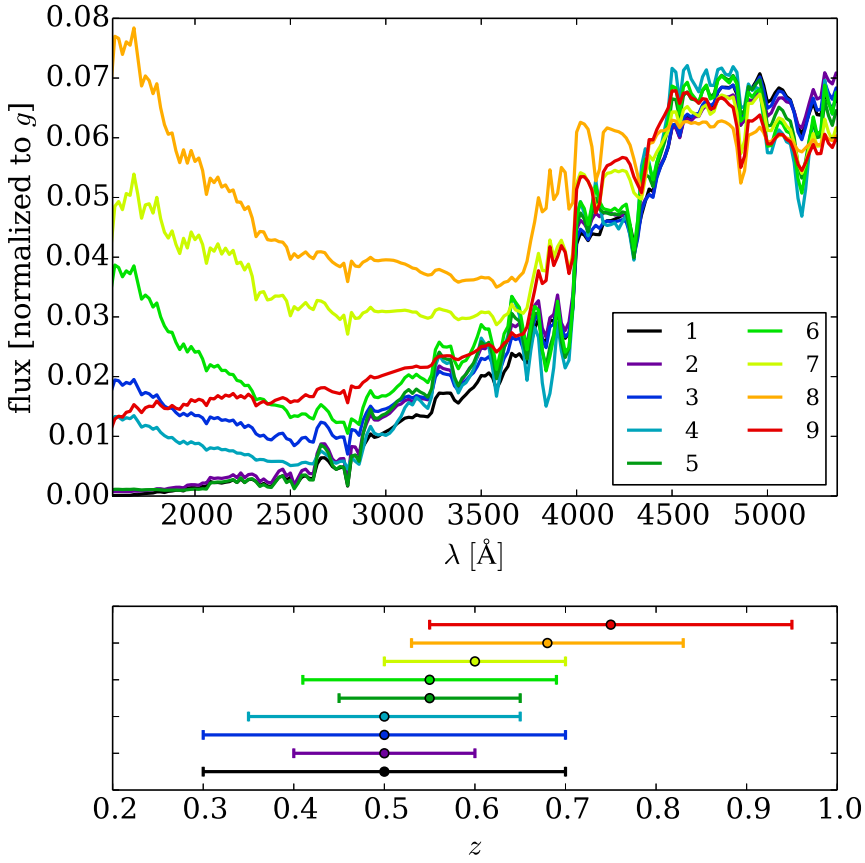

Figure 6.15: *Upper panel:* The nine surviving model SEDs within the range of the SDSS *ugriz* filter system at redshifts $0.45 \leq z \leq 0.9$ (the z range of our galaxy sample). *Lower panel:* Corresponding z prior parameters, μ_z and σ_z. The color code is the same as in the upper panel.

avoided through a different luminosity prior.

The model SEDs are shown in Figure 6.15, where we plot them in the wavelength range covered by SDSS at the redshifts of our catalog $0.45 \leq z \leq 0.9$. In the lower panel we present the redshift prior parameters μ_z, and σ_z which correspond to the model SEDs. Additionally, Figure 6.16 shows the colors of the nine models as a function of redshift. With the nine models we can account for the large spread of the data in most cases. Furthermore, we cover also the bluer parts in $g-r$, $r-i$, and $i-z$ (cf. Fig. 6.2). For lower redshifts, our models produce colors that cover only the bluer $r-i$ and $i-z$ ranges of the data. The physical parameters and the λ_i and β values of the nine model SEDs are summarized in Table 6.1. The model parameters explain the behavior of the SEDs in Figure 6.15. We see from Figure 6.15 that the model SEDs roughly follow a trend and become bluer with increasing redshift (except for the red colored SED with $\mu_z = 0.8$), which is mirrored in the values of Table 6.1. The e-folding timescale τ also roughly increases, as well as the burst fraction, making the resulting SEDs bluer on average. The red highlighted SED has lower fluxes than the orange and the light green SEDs in the UV part of the spectrum, not continuing the sequence. It is redder

model no.	**Z**	$\boldsymbol{\tau}$ **[Gyr]**	**age [Gyr]**	$\mathbf{A_V}$	**burst**	$\mathbf{A_{V,\,burst}}$	$\boldsymbol{\lambda_i}$ **[Å]**	$\boldsymbol{\beta}$
1 (black)	0.008	SSP	2.0	1.2	–	–	5000	2.0
2 (violet)	0.004	SSP	6.0	0.6	1.0 %	2.0	5500	1.5
3 (blue)	0.008	1.0	6.0	1.0	1.0 %	0.0	5000	1.5
4 (turquoise)	0.05	SSP	8.0	0.7	1.0 %	0.0	3500	2.0
5 (dark green)	0.02	SSP	4.0	0.0	1.0 %	1.0	5000	1.5
6 (green)	0.05	SSP	3.0	1.3	2.0 %	0.0	3500	2.0
7 (light green)	0.05	3.0	5.0	1.4	1.0 %	0.0	5000	2.0
8 (orange)	0.004	50.0	4.0	2.7	1.0 %	0.0	3000	1.5
9 (red)	0.02	3.0	4.0	2.2	2.0 %	2.0	4500	2.0

Table 6.1: Physical parameters of the nine surviving templates. The first column gives the numbering of the models according to Figure 6.15 and the plot color from Figures 6.15 and 6.16. The column "burst" is the mass fraction of the burst, whereas $A_{V,\,burst}$ is the burst extinction in V.

because of the high extinction values and because of its (relatively) high metallicity, $Z = Z_\odot$. In summary, although we can see a qualitative trend in the SEDs as a function of redshift prior (which originates from the spectroscopic redshifts of the underlying z bin with small adaptions), the trends in the physical parameters are not that evident. This is because they are degenerate and changes in one parameter can yield similar results to a variation in another parameter (e.g., the well-known age-metallicity degeneracy).

Before analyzing the photo-z performance of the new models, we want to introduce several metrics which we use to assess the photometric redshift quality, additionally to those introduced in Section 5.3. We decide to provide a large number of metrics to enable the reader to compare with other publications. The mean errors are characterized by the bias $\langle \Delta z \rangle$ and the mean absolute error $\langle |\Delta z| \rangle$, as well as by their corresponding values in the rest frame, $\langle \Delta z_{rf} \rangle$ and $\langle |\Delta z_{rf}| \rangle$. The root of the sample variance is denoted $\sigma(\Delta z_{rf})$, and $\sigma_{68,rf}$ is the half of the width of the Δz_{rf} distribution where 68.27 % of the sample is located, corresponding to a 1σ confidence interval. We furthermore calculate the fractions $\eta_{2\sigma}$, $\eta_{3\sigma}$ of galaxies within $\langle \Delta z_{rf} \rangle \pm i \cdot \sigma(\Delta z_{rf})$, where $i = 2, 3$ (cf. Carrasco Kind & Brunner, 2014b). If Δz_{rf} behaved as a perfect Gaussian, then $2 \cdot \sigma(\Delta z_{rf})$ and $3 \cdot \sigma(\Delta z_{rf})$ would be the 2σ and 3σ confidence intervals and therefore hold 95.45 % and 99.73 % of the objects. But the distribution of Δz_{rf} can be non-Gaussian and still yield reasonable values for $\eta_{2\sigma}$ and $\eta_{3\sigma}$ (compare Figs. 6.18 and 6.23 later on).
To evaluate the precision of the photo-z errors δz estimated by the code from the PDF (cf. Eq. 2.10 in Section 2.2.2), we introduce $\Delta z' \equiv \Delta z / \delta z$. If the errors are estimated correctly, the distribution of $\Delta z'$ should resemble a standard normal distribution (cf. Sánchez et al., 2014). We will therefore analyze the values of the mean $\langle \Delta z' \rangle$ and the standard deviation $\sigma_{\Delta z'}$ of the distribution of $\Delta z'$.
Lastly, we perform Kolmogorov-Smirnov (KS) tests on the distributions $N(z_{\mathrm{phot}})$ and $P(z_{\mathrm{phot}})$ (cf. Carrasco Kind & Brunner, 2014b), where the latter is the z_{phot} distribution derived by stacking the PDFs for all objects. The KS test value is the maximum absolute difference between the cumulative distribution functions $\int_{-\infty}^{z} P(\hat{z}) d\hat{z}$ of the probability density of $N(z_{\mathrm{phot}})$, or the stacked $P(z_{\mathrm{phot}})$, to the cumulative distribution of $N(z_{\mathrm{spec}})$. All the metrics are summarized in Appendix A.

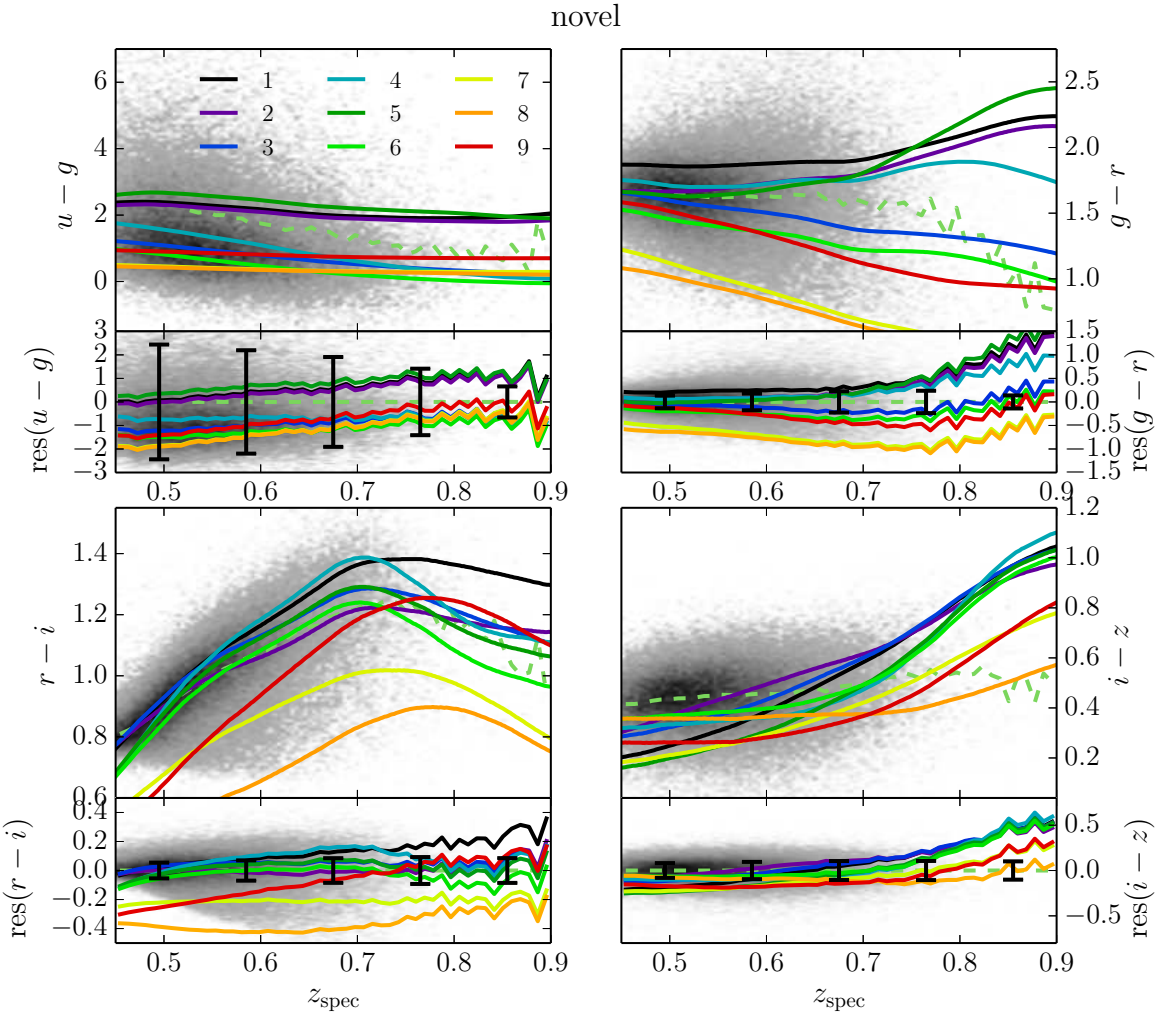

Figure 6.16: In the upper panels of the four major panels the CMASS colors are plotted as a function of redshift in gray, and the median is shown by a green dashed line. The lower panels are normalized to the median color and show the residuals. Error bars present the median data errors in five redshift bins. On top of that the predicted colors of the nine selected models are drawn. The color code matches that of Figure 6.15.

After this introduction of photo-z quality metrics we present the photometric redshift results of the CMASS sample with the novel template set and priors of Section 6.2.2 and Figure 6.15 in Figure 6.17. The upper left panel shows $z_{\rm spec}$ versus $z_{\rm phot}$, the middle left panel shows the photometric redshift rest frame error as a function of spec-z, and the lower left panel presents $z_{\rm spec}$ versus $|\Delta z_{rf}|$. We indicate the median Δz_{rf} and $|\Delta z_{rf}|$ values in the middle left panel by solid and dashed black lines, whereas in the lower panel we highlight the $\sigma_{68,rf}$ and $\sigma_{\rm NMAD}$ values as a function of the spectroscopic redshift also by solid and dashed black lines. Finally, the right panel presents the normalized redshift distributions for $z_{\rm spec}$ (gray filled histogram), the distribution derived from the single-value photometric redshifts $z_{\rm mode}$ (gray line) derived by the mode of $P(z)$, and the stacked PDFs (black line). The photometric redshift quality metrics discussed above are printed in the plot as well.
The bias has a small positive value for the lowest considered redshifts, then is close to zero, and is decreasing to negative values for higher redshifts $z \sim 0.7$. The overall mean value is still positive due to the small sample size at higher redshifts, visible by the number density in Figure 6.1 or the right panel of Figure 6.17. The scatter (lower left panel in Fig. 6.17) increases slightly with increasing spec-z, while the value of $\sigma_{\rm NMAD}$ stays more or less the same. This means, that the outliers (which are not considered in the calculation of $\sigma_{\rm NMAD}$) are predominantly responsible for an increase in $\sigma(\Delta z_{rf})$. The fraction of catastrophic outliers however is very small $\eta = 0.22\,\%$. In the right panel of Figure 6.17 we observe that the photo-z predictions from the mode of the PDF yield deviations from $N(z_{\rm spec})$ larger than for the case where the whole PDFs are considered. The excesses observed in $N(z_{\rm phot})$ are due to the overestimation mentioned previously. They are not anymore visible when we use the posterior $P(z)$ distributions in the $N(z)$ reconstruction, which means that the PDF should be favored for science analyses (cf., e.g., Mandelbaum et al., 2008; Hildebrandt et al., 2012). As for the error estimation of the `PhotoZ` code, we see from the values of $\langle \Delta z' \rangle$ and $\sigma_{\Delta z'}$, that $\Delta z'$ behaves very close to a standard normal distribution, which is the goal in the photometric redshift error predictions. This means that not only the approach for the calculation of δz is legitimate, but also that the models and priors create a reasonable $P(z)$ (from which δz is extracted).
Table 6.2 in Section 6.3.2 holds a summary of the derived quality metrics, together with results of public photo-zs from the SDSS, which we will analyze in Section 6.3.2.

To investigate if the estimated errors are reliable in identifying outliers, we assume the null hypothesis that an object is a photo-z outlier. The probability of an outlier being falsely classified as a non-outlier on account of δz is then the type I error α. Additionally, the type II error β gives the probability of a non-outlier being misclassified as an outlier by the estimated error. For the photo-zs of this work we get $\alpha = 0.93$ and $\beta = 3.5 \cdot 10^{-4}$. This means that an outlier is falsely classified as a non-outlier with probability α, but also that a non-outlier is almost never classified as an outlier (β). Although $\sigma(\Delta z')$ is close to, but slightly smaller than, 1, the deviation is probably caused by to shallow peaks around the mode of $P(z)$ for outliers, such that their errors are underestimated. The α and β value are summarized together with the results of the SDSS (Sec. 6.3.2) in Table 6.3.

The distribution of the photometric redshift errors in the rest frame is shown in Figure 6.18 by a gray filled histogram. In red we indicate the bias and the ranges of $\langle \Delta z_{rf} \rangle \pm i \cdot \sigma(\Delta z_{rf})$, where $i = 1, 2, 3$. Furthermore, we calculate the real interval widths where $95.45\,\%$, and $99.73\,\%$ of the galaxies are located and introduce them in the plot (black dash dotted and

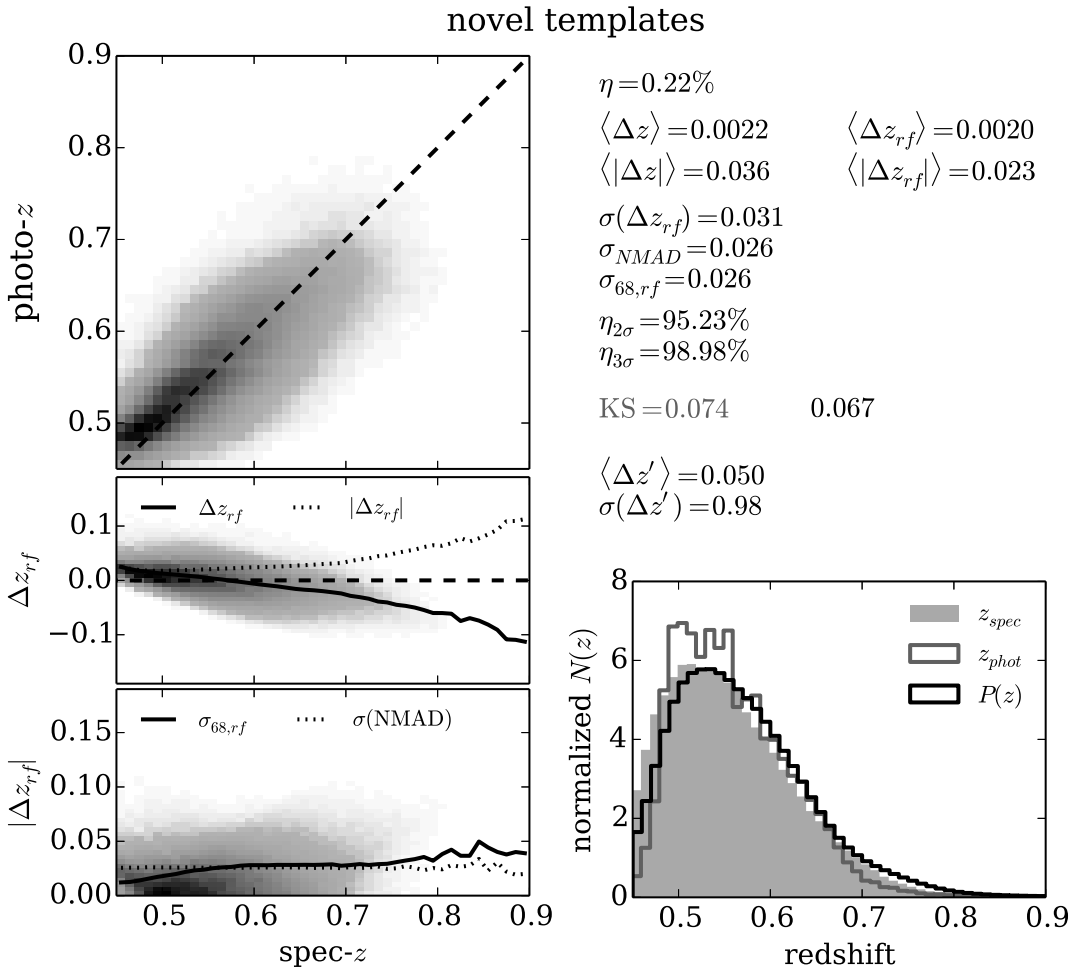

Figure 6.17: Photo-z results with novel templates, based on the BC03 $\lambda^{-\beta}$ models and selected through a SOM. *Left Panels:* The upper panel shows the spectroscopic versus the photometric redshift, where the dashed line is at $z_{\rm phot} = z_{\rm spec}$. The middle panel is the distribution of the rest frame photo-z error as a function of the spectroscopic redshift, where we plot the bias (solid line) as well as the mean absolute error (dashed line). Finally, the lower panel shows $z_{\rm spec}$ versus the absolute rest frame error, where we plot σ_{68} (solid) and $\sigma_{\rm NMAD}$ (dashed) on top. *Right Panel:* The relative frequency distributions of $z_{\rm spec}$ (gray filled histogram), $z_{\rm phot}$ (dark gray line), and the stacked PDF of all objects (black line). Additionally, the photo-z quality metrics are printed in the figure. We provide two values for the *KS* test, for the predictions of single $z_{\rm phot}$ values (from the mode of the PDF, gray), and for the PDF (black).

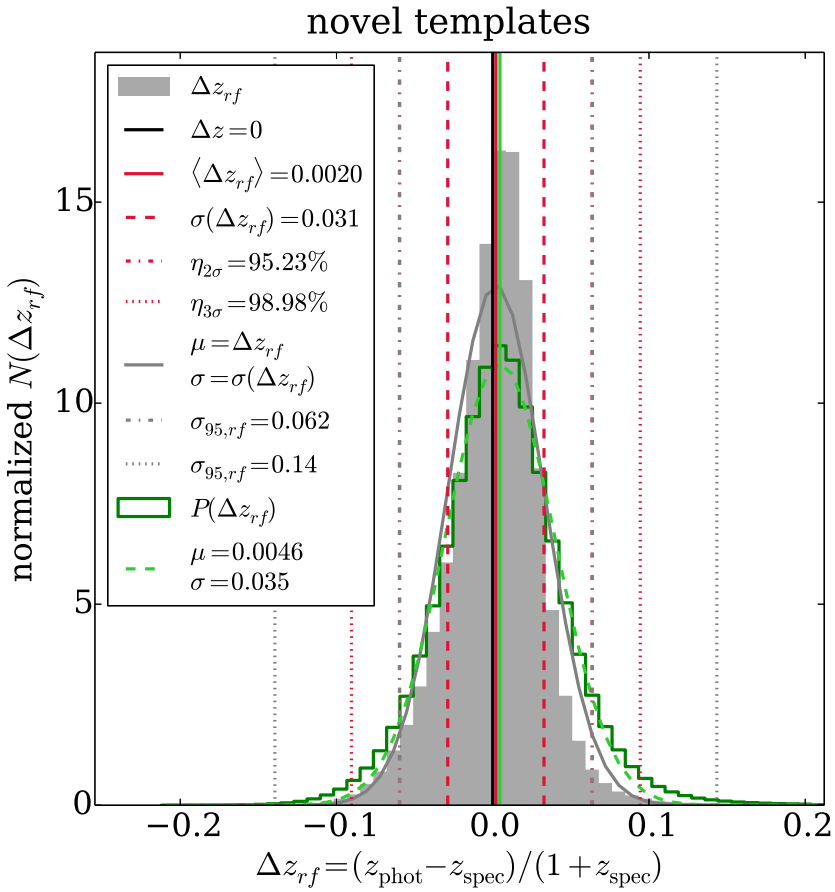

Figure 6.18: Photometric redshift rest frame error distribution (gray filled histogram) estimated with the novel templates and priors. $\Delta z = 0$ is highlighted by a solid black line, whereas the bias is shown by a solid red line. The steps at $\Delta z_{rf} \pm \sigma(\Delta z_{rf})$, $2\sigma(\Delta z_{rf})$, and $3\sigma(\Delta z_{rf})$ are represented by dashed, dash dotted, and dotted red lines, and the corresponding number fractions are given in the legend. We also plot a Gaussian with $\mu = \langle \Delta z_{rf} \rangle$, and $\sigma = \sigma(\Delta z_{rf})$ with a dotted black curve. Furthermore, dash dotted and dotted gray vertical lines show the real $\sigma_{95,rf}$ and $\sigma_{99,rf}$ intervals. Since $\sigma_{95,rf} \approx 2\sigma(\Delta z_{rf})$ we cannot discern the red and the gray dash dotted lines. The Δz_{rf} derived from the stacked PDFs is shown by a green solid histogram, to which we fit a Gaussian highlighted by a dashed green curve. The best fitting parameters of the Gaussian are given in the legend.

dotted vertical lines). Finally, we calculate the Δz_{rf} distribution from the stacked PDFs and fit a Gaussian to the histogram (green lines). The peak of the Δz_{rf} distribution is slightly shifted to the right, which is again due to the overestimated photo-zs at the lowest redshifts. Concerning the resemblance to a Gaussian, we observe from the values given in the plot that $\sigma_{95,rf} \approx 2\sigma(\Delta z_{rf})$ (which is why we cannot discern the lines in Figure 6.18) and $\sigma_{99,rf} > 3\sigma(\Delta z_{rf})$. The latter is due to the outliers which can be scattered far from the spec-z value due to the allowed fitting range. The fractions $\eta_{2\sigma}$ and $\eta_{3\sigma}$ of objects within $2\sigma(\Delta z_{rf})$ and $3\sigma(\Delta z_{rf})$ have reasonable values. While $\eta_{2\sigma} = 95.23\,\%$ is very close to the desired value of $95.45\,\%$, $\eta_{3\sigma} = 98.98\,\%$ is slightly smaller by only $\sim 0.8\,\%$. The Δz_{rf} distribution derived from the PDFs is slightly broader than the distribution of the individual photo-z results.

This is a consequence of the asymmetry of the PDFs that often have higher probabilities for redshifts greater than the most probable z.

Figure 6.19 presents the deviations of magnitudes, predicted by the models in a photo-z run, and the data. A Gaussian is fitted to the histograms whose best fitting parameters are printed in the panels. For the r and i band magnitudes the Δmag distributions are very narrow, while the expectation values of the Gaussians are near zero. This is due to the depth of the photometry in these filters, which is greatest in i and r, and the photometry has the smallest measurement uncertainties. For the g and the z band magnitudes however, the broadness of the Δmag distributions increases, which is mostly an effect of the more shallow photometry in these bands, but also of the respective magnitude ranges of the galaxies. Furthermore, the mean magnitude deviations in these two bands are higher than in r and i. While μ of the best fitting Gaussian is still small in comparison to σ for the g band, it is relatively high in the z band. We selected the models for the photo-z estimation only on account of their photo-z performance, and not how well they match the data in terms of magnitudes. However, this does not necessarily imply that the BC03 $\lambda^{-\beta}$ models are not able to fit the data in terms of magnitudes. Indeed, calculating the average magnitude offsets of the SED fitting of Section 6.2.1, yield maximum values of $[0.6, 0.9, 0.3, 1.1] \cdot 10^{-3}$ in g, r, i, and z for the four redshift bins.

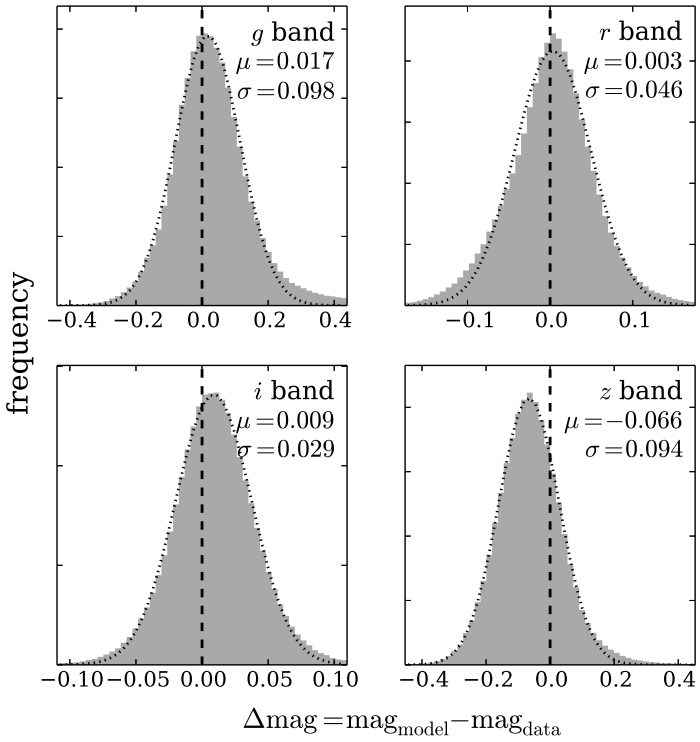

Figure 6.19: Deviations of the predicted magnitudes of the novel model SEDs from the data.

6.3.2 Comparison to SDSS Photometric Redshifts

We present the photometric redshifts from the SDSS-III database in comparison to the results of our code-template combination. The SDSS database provides photo-zs from two different empirical methods. One code uses a k-d tree nearest neighbor fit to derive the redshifts (hereafter KF after Csabai et al., 2007).

The SDSS-KF results for our catalog are shown in Figure 6.20, which is equivalent to Figure 6.17 and also shows the quality metrics. From the middle panel of Figure 6.20 we see that the location of the densest part of the population lies close to the $\Delta z = 0$ line for low redshifts $z \lesssim 0.6$. In all higher z_{spec} regions the photometric redshift is systematically underestimated which leads to the low negative value in the bias. An underestimation is also present in the results from Section 6.3.1, but it is slightly higher in the SDSS-KF case. From the lower panel of Figure 6.20 we see that also here the rise in $\sigma_{68,rf}$ is mostly caused by outliers, which are not considered in the calculation of σ_{NMAD} (which stays more or less constant). Taking the estimated photometric redshift errors into account, the relatively high bias does not decrease, which is mirrored in the higher absolute value of $\langle \Delta z' \rangle$, but even still increases. Hence, the bad estimates of the photo-zs are not recognized by the error estimation. This can also be seen in the value of $\sigma_{\Delta z'}$ which is well beyond 1 and means that the errors are on average underestimated. This also affects the results of a significance test on outlier classification (cf. Sec. 6.3.1). The errors type I and II read (at two significant figures) $\alpha = 1.0$ and $\beta = 7.9 \cdot 10^{-4}$ (see also Tab. 6.3 for comparison with other results), confirming that the errors are underestimated.

In terms of $N(z_{\mathrm{phot}})$ the low photo-z values at higher z_{spec} do not change the shape significantly, which is due to the low sample size in these regions, and results in the low value of *KS*. We cannot compare to the $N(z)$ results for PDFs, since they are not provided.

As before, we show in Figure 6.21 the Δz_{rf} distribution. Comparing with our case from Figure 6.18, we see that here the results are slightly shifted to the left, an effect due to the underestimated photo-zs at higher redshifts. Also in the case of the SDSS-KF redshifts, the $\sigma_{95,rf}$ and $\sigma_{99,rf}$ values do not perfectly coincide with $2\sigma(\Delta z_{rf})$ and $3\sigma(\Delta z_{rf})$, but are elevated. $\eta_{2\sigma} = 95.25\,\%$ is very close to the desired value, while $\eta_{3\sigma} = 98.43\,\%$ deviates by $\sim 1.3\,\%$.

In comparison to the redshifts we get with our novel templates and code, the quality metrics for bias, error, and scatter are all higher in the case of the SDSS-KF photo-zs. Furthermore, the estimated redshift errors are more reliable in our code-template combination. The KS test yields better results in the case of this work, since the SDSS-KF photo-zs are systematically underestimated.

The second photometric redshift results published by the SDSS are from another empirical code by Carliles et al. (2010), which uses random forests to predict photo-zs (hereafter SDSS-RF). Figure 6.22 presents the results of the SDSS-RF code. In this case, the outlier rate is significantly higher than in both previous cases, as well as the mean and the absolute rest frame errors. Above $z \sim 0.55$ the photo-zs are systematically underestimated to a greater extent than in the former two cases, yielding the low bias. Furthermore, the scatter values $\sigma(\Delta z_{rf})$, $\sigma_{68,rf}$, and σ_{NMAD} have also increased in comparison to the results of this work, as well as compared to the SDSS-KF redshifts.

In the distribution of $N(z_{\mathrm{phot}})$ we can observe the lack of higher photo-z values which accumulate around $z \sim 0.55$. Because of that, the KS test value is higher. In terms of their

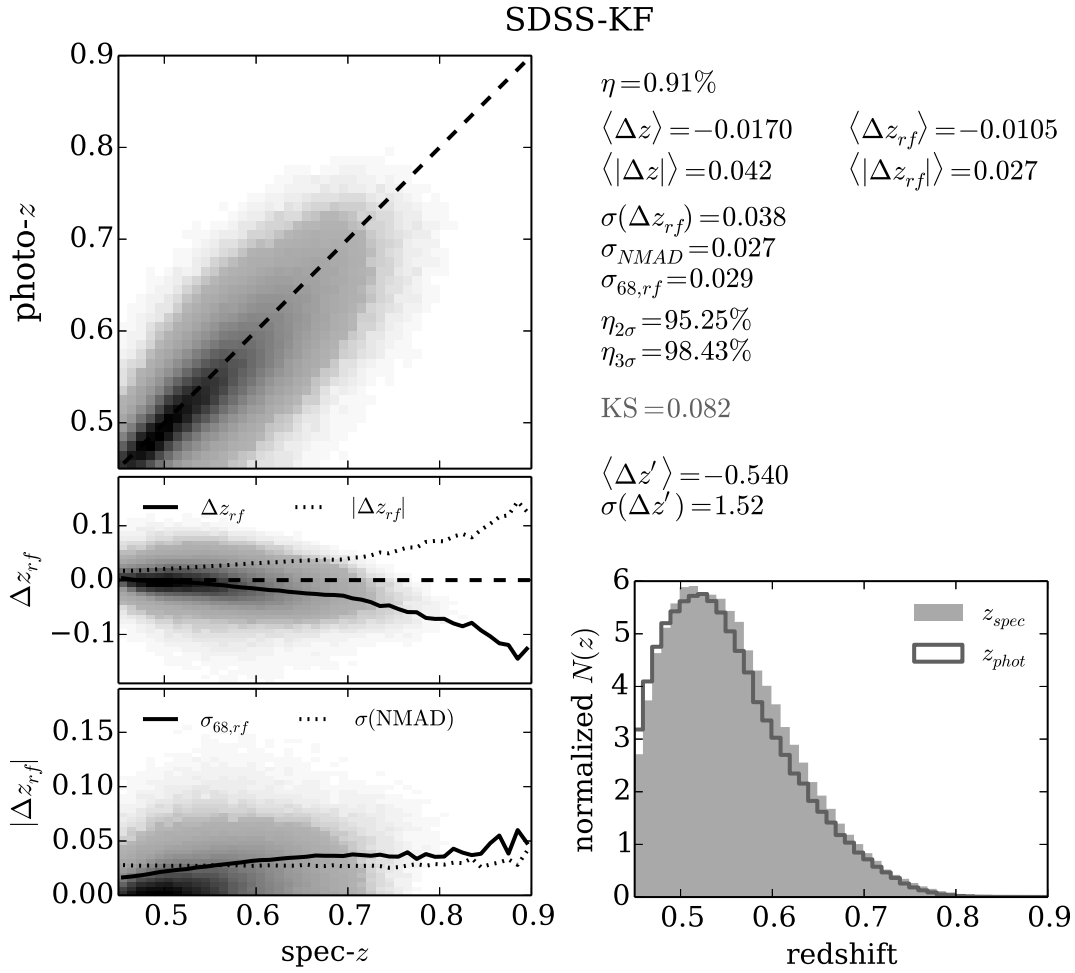

Figure 6.20: Photo-z results published by the SDSS from the KF approach. The KS value is determined from the single-value photo-z estimates. See Figure 6.17 for a detailed caption.

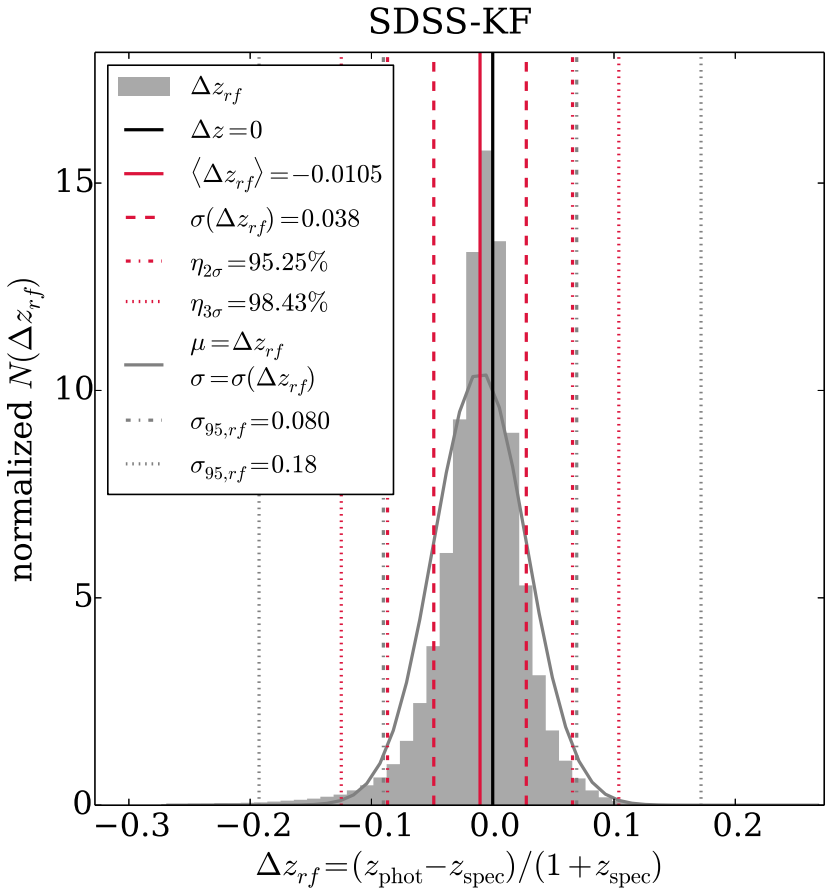

Figure 6.21: Photometric redshift rest frame error distribution (gray filled histogram) estimated by the SDSS-KF method. A detailed explanation is given in Figure 6.18.

estimated photometric redshift errors, the SDSS-RF method outperforms the SDSS-KF approach, but still produces a larger bias in $\langle \Delta z' \rangle$. The errors are furthermore overestimated, visible in the $\sigma_{\Delta z'}$ value which is lower than the desired value of 1. Again, this affects also the outlier classification derived from δz. In the case of SDSS-RF photo-zs, the null hypothesis of an object being an outlier yields $\alpha = 0.86$ and $\beta = 0.015$ (shown together with the previous results in Table 6.3). The higher values of δz yield a smaller probability of an object being misclassified as a non-outlier, but also raises β (although only to a small value). So although the errors are overestimated on average, outliers are still misclassified at a significance level of 0.86.

In Figure 6.23 we plot the Δz_{rf} distribution. The shift to the left is present to an even greater extent than in the previous case, and strongly deviating from a Gaussian with $\mu = \langle \Delta z_{rf} \rangle$ and $\sigma = \sigma(\Delta z_{rf})$ also plotted in Figure 6.23. This is a result of the underestimation of the photo-zs. Furthermore, as in the previous cases, the value of $\sigma_{95,rf}$ is marginally higher than $2\sigma(\Delta z_{rf})$, while the deviation is greater in $\sigma_{99,rf}$. On the other hand $\eta_{2\sigma} = 95.35\,\%$ is even closer to $95.45\,\%$ than in both previous cases, while $\eta_{3\sigma} = 97.93\,\%$

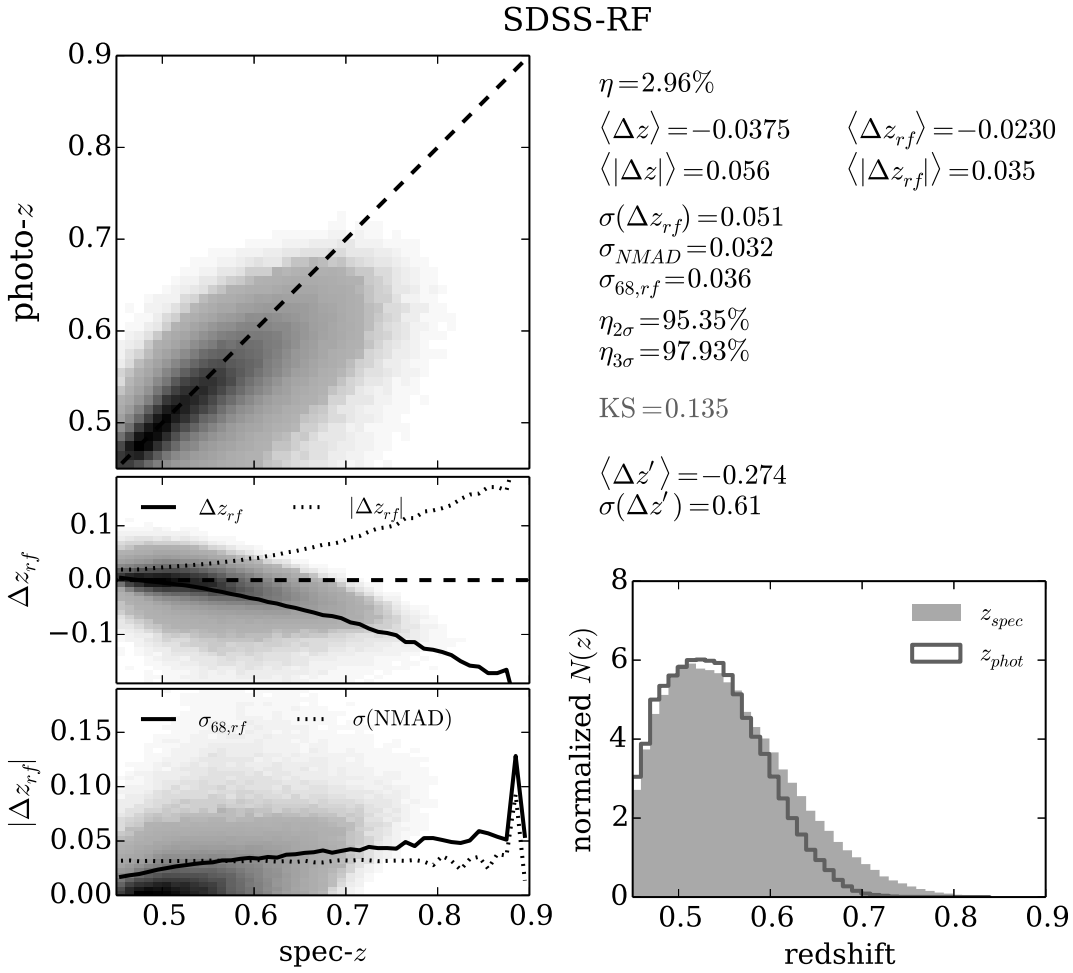

Figure 6.22: Photo-z results published by the SDSS from the empirical SDSS-RF approach. See Figure 6.17 for a detailed description.

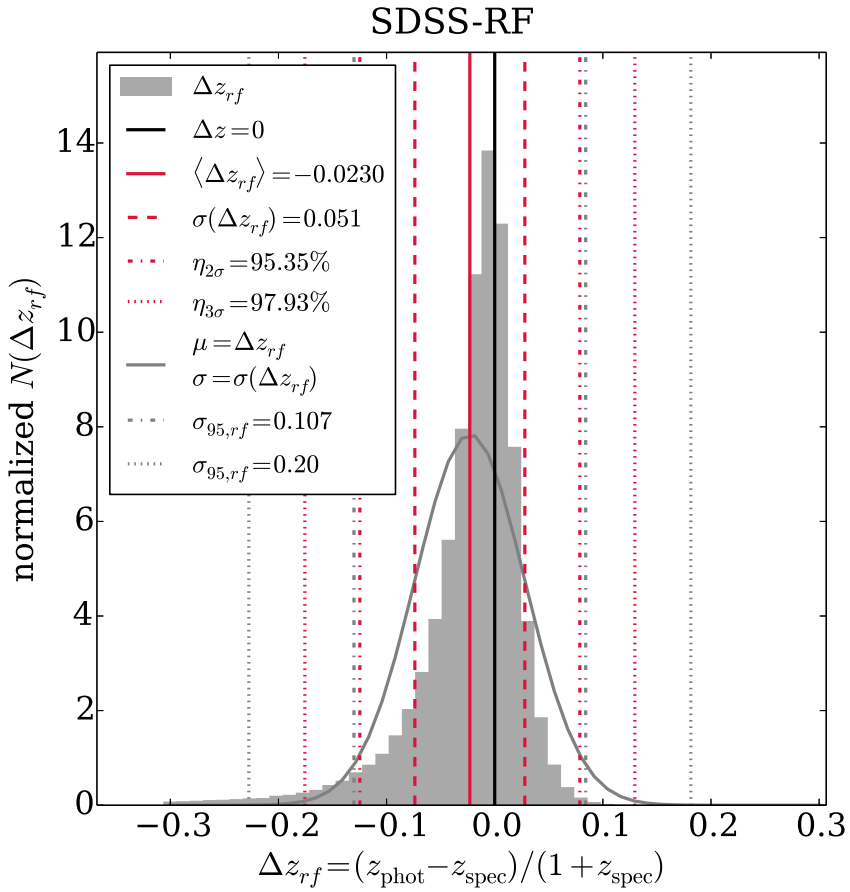

Figure 6.23: Photometric redshift rest frame error distribution of the empirical random forest method (SDSS-RF) and published by the SDSS. A detailed explanation is given in Figure 6.18.

deviates more strongly, $\sim 1.8\,\%$.

We present a summary of the photometric redshift quality metrics of the results from the template fitting of this work and the two SDSS photo-zs in Table 6.2. Our model-prior combination produces the lowest outlier fraction, bias, and mean absolute error compared to the SDSS photo-z results. When considered as a function of spec-z the photometric redshifts derived in this work are slightly biased to lower $z_{\rm phot}$ values for higher $z_{\rm spec}$, but not to the same extent of the SDSS codes. Furthermore, the different scatter values calculated ($\sigma(\Delta z_{rf})$, $\sigma_{68,rf}$, and $\sigma_{\rm NMAD}$) are also lower for photo-zs of this work. The distribution of redshift errors are more similar to a Gaussian for the models, priors, and code of this work, which can be observed in Figures 6.18, 6.21, and 6.23. Considering the number fractions within $2\sigma(\Delta z_{rf})$ and $3\sigma(\Delta z_{rf})$ all three codes perform similarly well. However, this does not mean that the Δz_{rf} distributions are necessarily good approximations of Gaussian distributions with $\mu = \Delta z_{rf}$ and $\sigma = \sigma(\Delta z_{rf})$ which can be observed in Figures 6.18, 6.21 and 6.23.
When we evaluate the similarity between the redshift distributions, from which spectroscopic and photometric redshifts are sampled, through a KS test, the photo-zs of this work yield

setup	η [%]	$\langle \Delta z_{rf} \rangle$	$\langle \lvert \Delta z_{rf} \rvert \rangle$	$\sigma_{68,\mathrm{rf}}$	$\langle \Delta z' \rangle$	$\sigma(\Delta z')$	**KS**
this work	0.22	0.0020	0.023	0.026	0.050	0.98	0.067 (0.074)
SDSS-KF	0.91	−0.011	0.027	0.029	−0.54	1.5	0.082
SDSS-RF	3.0	−0.023	0.035	0.036	−0.27	0.61	0.14

Table 6.2: Summary of photometric redshift quality metrics of the template fitting results with the novel templates and priors used with the `PhotoZ` code, the SDSS photo-zs of the KF code, and the random forest code (SDSS-RF). Values for KS in the case of this work are derived from $P(z)$, whereas the values in brackets are calculated using only the z_{mode} predictions.

	this work	**SDSS-KF**	**SDSS-RF**
$\boldsymbol{\alpha}$	0.93	1.0	0.86
$\boldsymbol{\beta}$	$3.5 \cdot 10^{-4}$	$7.9 \cdot 10^{-4}$	$1.4 \cdot 10^{-2}$

Table 6.3: Significance test results for outlier classification. The null hypothesis is that an object is an outlier, and α and β are the type I and II errors respectively (rounded to two significance figures).

the best results for predictions of the mode which are improved when considering the whole $P(z)$ distribution.

The results of a significance test with the null hypothesis being that an object is a photometric redshift outlier is shown in Table 6.3. Not one of the codes produces errors that can reliably predict outliers. Albeit the SDSS-RF produces the best α, this is mostly due to the overestimated errors (cf. Tab. 6.2).

6.3.3 Deviations in Color Predictions

In the last section we analyzed the photometric redshift results with our code in comparison to the SDSS photo-zs. The redshifts from the SDSS-RF method were outperformed by those derived with the SDSS-KF code. We want to analyze if the photo-zs of this work can produce the CMASS galaxy colors to a better extent than if we fit the same models to the photometric redshifts of the SDSS-KF method.

Figure 6.24 presents the differences between the model predictions and the data in the $g-r$, $r-i$, $i-z$, and $g-z$ colors. The prediction of the photo-zs of this work are plotted in gray, the best fitting Gaussian is presented by a black curve, and its parameters are given in the plot. When fitted to the photometric redshifts of the SDSS-KF approach, the BC03 $\lambda^{-\beta}$ models predict colors that are represented by a red histograms in Figure 6.24, to which we also fit a Gauss curve. In all four colors, the standard deviation of the Δcol distribution is larger for SDSS-KF. Furthermore, the mean deviation from the data colors is greater for SDSS-KF results in the $i-z$ and $g-z$ colors. μ is of the same order in $r-i$ for both cases. It is smaller for SDSS-KF photo-zs in $g-r$, but there both values are very low. The fact that the SDSS-KF $r-i$, $i-z$, and $g-z$ colors are shifted bluewards from the distribution in colors of this work is a result of the SDSS-KF photo-zs being more heavily underestimated. All considered Δcol distributions resemble their best fitting Gaussians well in the inner parts, with increased flanks. However, this increase is more significant in for the SDSS-KF photo-zs in $r-i$. Therefore, the photometric redshifts derived in this work not only generally have better quality metrics than the SDSS photo-zs, but can furthermore recreate the colors of the CMASS galaxies to a better extent.

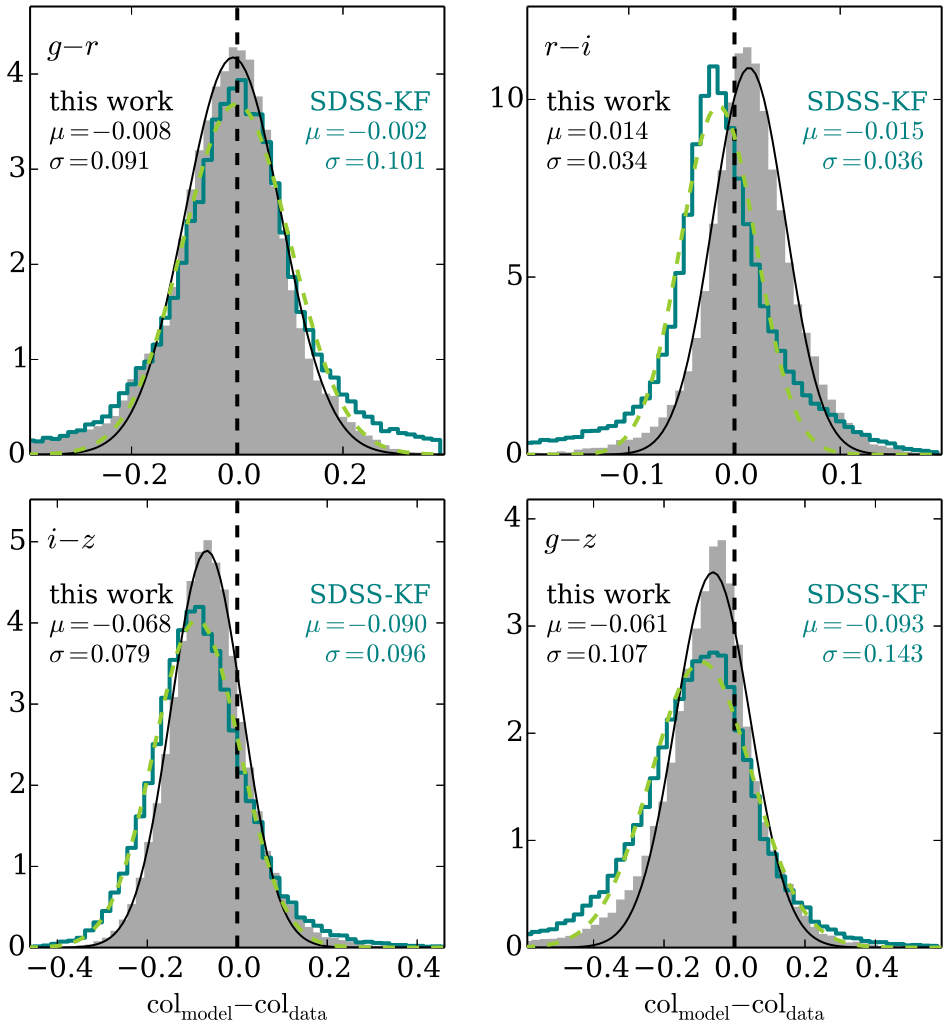

Figure 6.24: Deviation in $g-r$, $r-i$, $i-z$, and $g-z$ colors for photometric redshift results with the novel templates (gray filled histogram) and model predictions when fitting the same templates to the SDSS-KF photo-zs (blue histogram). Gaussians are fitted to both distributions, and are shown by a solid black line and a dashed green line respectively. The best fitting parameters of the Gaussians are printed in the panels.

6.4 Summary

In this work we created a set of model SEDs that are designed to match the colors of the BOSS CMASS sample and provide accurate photometric redshifts. We first analyzed the colors of the LRG model of Maraston et al. (2009) which was created to match the $g-r$ and $r-i$ colors of the LRG sample of SDSS-II (Eisenstein et al., 2001), and found that we cannot use a singular age and evolution configuration that matches the data in all colors. Therefore, we created models for four redshift bins of widths $\Delta z = 0.04$ centered on $z_{\rm spec} = 0.5, 0.6, 0.7, 0.8$ with the stellar population synthesis code from Bruzual & Charlot (BC03, 2003). We generated SSPs and CSPs with decreasing SFHs at various metallicities, and sampled the models from a wide age range. These models were then fitted in superposition to a burst component and dust extinction to the data at known spectroscopic redshifts. In order to better recreate the colors of the CMASS galaxies we introduce additional degrees of freedom in modifying the model SEDs by multiplication with $\lambda^{-\beta}$ for $\lambda > \lambda_i$ with several values for λ_i and β. We showed that the BC03 models modified in this way indeed are a better match to the colors of the CMASS galaxies (Figs. 6.3 and 6.4, 6.7 and 6.8, and 6.9 to 6.12) and also yield better χ^2 values from the fitting (Fig. C.1).
From these best fitting SEDs we selected a small subset that should cover the region in color space and absolute magnitude in r of the CMASS sample. We therefore projected the CMASS galaxy colors and M_r of the four z bins onto two-dimensional planes using a self-organizing map. Afterwards we partition the plane in six clusters for each redshift bin with a k-means clustering algorithm and select one model SED per cluster cell that produces the best photo-z for galaxies within the same cell. We estimate photometric redshifts with a template fitting code and with the selected models and analyze their individual performance. Thereby we modify their redshift priors to improve on the photo-zs, but also decide to omit some of the templates which do not yield accurate photo-zs on the whole sample, regardless how the priors are modified. We then compared the photometric redshift results with the photo-zs two empirical method published by SDSS and calculate several metrics that assess the quality of the photo-zs, their estimated errors, and their distribution. We found that the photo-zs with the generated models of this work produce better values in all quality metrics. Furthermore, we observed that including the stacked PDFs yield better results in the reconstruction of $N(z_{\rm spec})$ and yield better results in a KS test. Concerning the estimated errors δz, a significance test shows that none of the three considered δz results provides a significant classification of outliers. However, the probabilities of a non-outlier being misclassified is very small for all three considered cases.
Finally, we compared the predicted colors of the novel model SEDs when fitted to the photo-zs of this work and to the better of the two SDSS redshifts, to the data. We found that the deviations from the data are smaller for photo-zs of this work.

Chapter 7

Photometric Redshifts from Dark Energy Survey Science Verification Data

I participated in a project comparing the photometric redshift performance of numerous photo-z codes with data from the *Dark Energy Survey* (DES, Flaugher, 2005). It was published by Carlos Sánchez et al. in the *Monthly Notices of the Royal Astronomical Society* (Sánchez et al., 2014). The work presented in this chapter is part of this publication, but presented in less detail, and with a focus on my contribution.

Introduction

DES aims to improve our understanding of the nature of dark energy, and therefore the acceleration of the Universe. Combining Clusters, BAOs, Weak Lensing, and Supernova analyses, the DES forecast predicts errors of, e.g., $\sigma(\Omega_{\mathrm{DE}}) = 0.004$ on the dark energy density, and $\sigma(w_0) = 0.061$ on the dark energy equation of state. The survey is carried out with the 4 m Victor M. Blanco Telescope, situated at the Cerro Tololo Inter-American Observatory (CTIO) high in the Chilean Andes. It will probe the southern sky during a five year period. DES started officially in August 2013, after a science verification (SV) phase from November 2012 to February 2013 following the commissioning of the camera. The *Dark Energy camera* (DECam, Flaugher et al., 2012; Diehl & Dark Energy Survey Collaboration, 2012) takes images in the *grizY* filters during the survey, where *griz* are Sloan filters, and the central wavelength of the Y filter is at 9899 Å, i.e., redwards of the z band. For science verification, imaging data were also taken in the u band, although this is not part of the main survey. DES will observe $\sim 5000\,\mathrm{deg}^2$, which is about one eighth of the sky, out to an depth of $i < 24$. It will survey approximately 300 million galaxies up to a redshift of $z \sim 1.4$. The SV footprint was chosen such that it overlaps with several spectroscopic surveys, the VVDS (Le Fèvre et al., 2005), zCOSMOS (Lilly et al., 2007), and ACES (Cooper et al., 2012), with an approximate survey area of $\sim 3\,\mathrm{deg}^2$ for each field. The exposure times, and therefore photometric depths, vary between filters and fields and we would like to refer the reader to Sánchez et al. (2014, Sec. 2, Tab. 1) for details.

The SV data consist of two subsamples, a *main* sample with the same depth as in the main survey, and a *deep* sample with three or more times the exposure time of the *main* sample. Furthermore, the data are divided into a calibration or training sample and an validation or

query sample. The former (where the nomenclature depends on whether one uses a template fitting or empirical photo-z method) should be used to re-calibrate the zero point offsets and prior probabilities, or, in the case of empirical methods, for the training of the estimator. To re-calibrate the zero points of the data is often useful if the data is not well calibrated but shows systematic offsets. After the calibration/training the photometric redshifts are estimated for the query sample, using the recalibration (or trained model respectively) from the training sample.
Since spectroscopic samples suffer from selection effects and are shallower than the photometric data (cf. Sec. 2) a weighting scheme is applied to the calibration data to ensure its magnitude and color distributions are the same as those of the full sample (Sánchez et al., 2014, Sec. 3.1). The weighting procedure compares local densities in magnitude and color spaces of the training sample with the full sample, and assigns a weight to each galaxy according to the ratio between the density of the calibration to the full data set using a nearest neighbor algorithm (Lima et al., 2008). These weights are also used when the photometric redshift quality metrics are calculated.
The DES photo-z working group proposed several tests from which we present the most important ones in this chapter. The first (second) test consists of calibrating/training on the *main* (or *deep*) sample and validating on the *main* (or *deep*) sample. The *main-main* test (first test) is the most important, because it corresponds to the real configuration of DES. The reason for the *deep-deep* test (second test) is to analyze the photo-z performance for higher S/N ratios. In the third test we calibrate on the *deep* sample and validate on the *main* sample. The reason for this *deep-main* test is to evaluate whether the photo-z performance is enhanced when the calibration/training is done on a sample with higher S/N. For the first three tests, second photo-z runs including u band data were performed, in order to investigate their importance for the accuracy of photometric redshift estimates. It is reasonable to assume that the photo-zs will improve for lower redshift galaxies when using the u band, since the 4000 Å break lies in the g band for small redshifts, wherefore data in u gives important information about the fluxes bluewards of 4000 Å, i.e., helps to distinguish between intrinsically red and blue galaxies.
We emphasize that the tests performed are blind tests: Although spectroscopic redshifts are available for the validation sets, we do not use their information in the photo-z estimation, but solely to evaluate the photo-z performance.

Setup of Photo-z Runs

Before discussing the results, we summarize the setup of the `PhotoZ` code used here. The model set we use contains SEDs ranging from star forming (blue) to passively evolving (red) galaxies. It includes model SEDs from Bender et al. (2001), which were created from spectroscopically observed objects in the Hubble Deep Field North. Another three templates (an S0, Sac, and an Sbc galaxy) are from Mannucci et al. (2001, where we already analyzed the colors of the S0 galaxy in Section 4.2.2), and two empirical SEDs (of an Scd and an Sbc galaxy) are from Coleman et al. (1980). Our model set additionally includes 13 SEDs from Ilbert et al. (2006) which are based on CWW spectra and are optimized to match local star-forming galaxies. This model setup is essentially a combination of template sets already used in the past for photometric redshift estimation (e.g., Bender et al., 2001; Brimioulle et al., 2013, and Sec. 2.2.2 and references therein). The models are presented in Figure 7.1, where we (loosely) classify them according to their SED shape. Furthermore, we incorporate the

template SEDs created for LRGs in Chapter 5. We do not show them in Figure 7.1 since they were discussed extensively in the sections of Chapter 5. Furthermore, we do not use the models created in Chapter 6, since they were not yet established at the time the work on the DES-SV data was done.

The redshift and luminosity priors have the form (cf. Sec. 2.2.2, Eq. 2.9)

$$P(x|T) \propto \exp\left[-\ln(2)\left(\frac{x-\mu_x}{\sigma_x}\right)^{p_x}\right],$$

where $x = z, M$, and μ_x, σ_x, and p_x are defined individually for each model SED. Setting μ_x, σ_x, and p_x accordingly, we can, for instance, decrease the probability of observing red models at higher redshifts ($z \gtrsim 0.9$), or avoid fits that are too bright in absolute magnitude. In addition to that, we adapt the z and M (absolute magnitude) priors for every model SED in such a way that photometric redshift outliers in the *main* and *deep* DES-SV training sets are less likely. Therefore, we identify their location in the z vs. M space and modify the priors in such a way that they assign smaller probabilities to those regions. This is done solely if the outliers of a template are isolated from good photometric redshift estimates of the same template in the z vs. M space. In Figure 7.2 we show an example of a model SED in our template set and its photo-z performance on the DES-SV data with and without including u band data. To assign lower probabilities to the location in z vs. M_B space inhabited by photo-z outliers, we chose to include the model twice with high exponents to create steep flanks of the z prior. The prior parameter values for this specific model are $\mu_{z,1} = 0$, $\sigma_{z,1} = 0.5$, and $p_{z,1} = 60$, as well as $\mu_{z,2} = 2.0$, $\sigma_{z,2} = 1.35$, and $p_{z,2} = 60$. Not all the priors had to be adapted in similar harsh manners, we presented here an extreme case.

The calibration is performed by an iterative adaption of the zero point offsets for the training catalogs. We use the median magnitude offsets between the data and the model predictions when fitting to spectroscopic redshifts, while optimizing the photo-z performance. The calibration is done separately for photo-z runs on catalogs with and without u band data. The zero point derived from this method are summarized in Table 7.1 for the *main* and the *deep* sample, with and without the u band.

	u	**g**	**r**	**i**	**z**	**Y**
main	—	0.00187	0.00492	-0.0257	0.0362	0.101
	-0.198	0.0157	0.00651	-0.0262	0.0332	0.0987
deep	—	0.0111	0.00404	-0.0236	0.0409	0.115
	-0.191	0.0296	0.00925	-0.0234	0.0377	0.111

Table 7.1: Zero point offsets of the DES *ugrizY* filters derived from the median differences between the data and the magnitudes predicted by the models when fitted to the spectroscopic redshift. The upper main row is calibrated on the *main* and the lower main row is calibrated on the *deep* sample. The two upper sub rows are calibrated without used u band data, whereas the lower sub rows are the results including u.

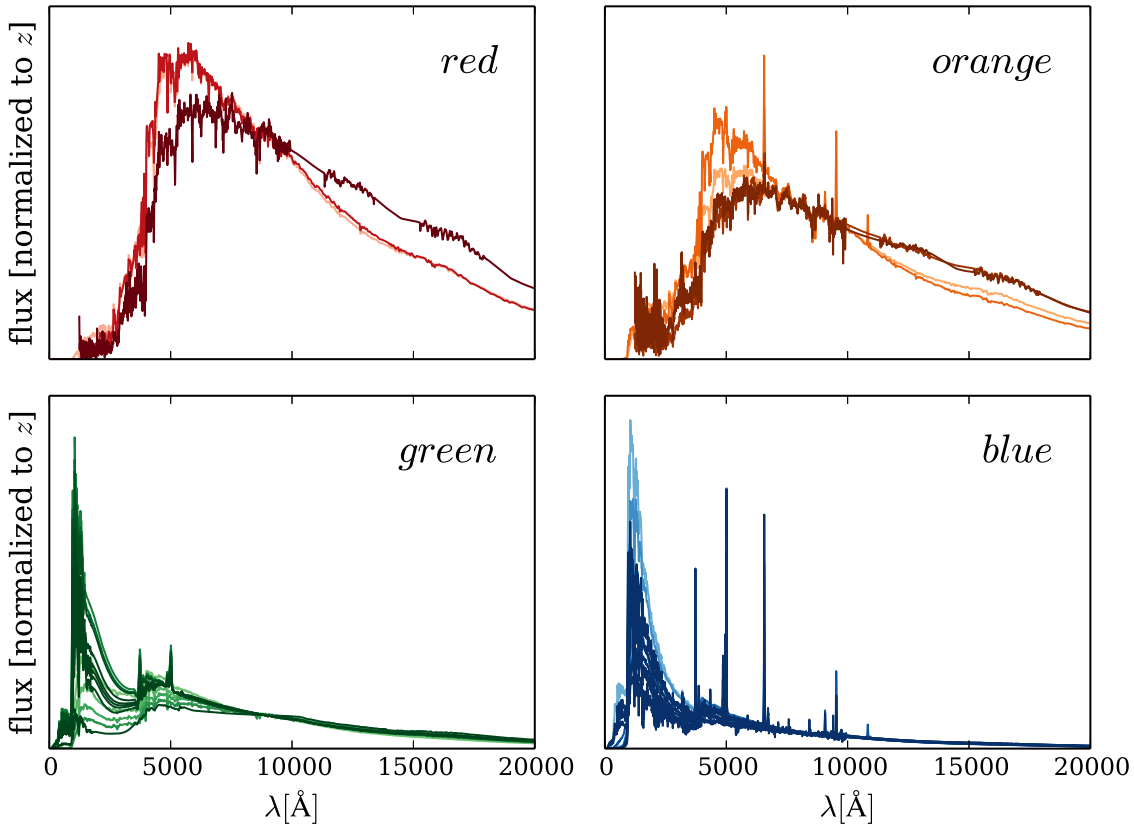

Figure 7.1: Model SEDs used to determine photo-zs for the DES Science Verification data. We classified the models loosely according to the SED properties into *red*, *orange*, *green*, and *blue*. The template SEDs tagged *red* are the same we showed in terms of redshift versus color in Fig. 4.7. They show strong 4000 Å breaks and only absorption lines. *Yellow* SEDs exhibit an increase in flux in the bluer parts of the spectrum, and mostly only small emission lines. The SEDs termed *green* have even more flux in the blue wavelengths and more prominent emission lines. Lastly, *blue* templates show very strong emission lines and a very high flux in the blue part of their spectra, such that the continuum of their SEDs is (apart from the spectral lines) almost proportional to $\lambda^{-\beta}$. The SEDs were normalized to their flux in the z band.

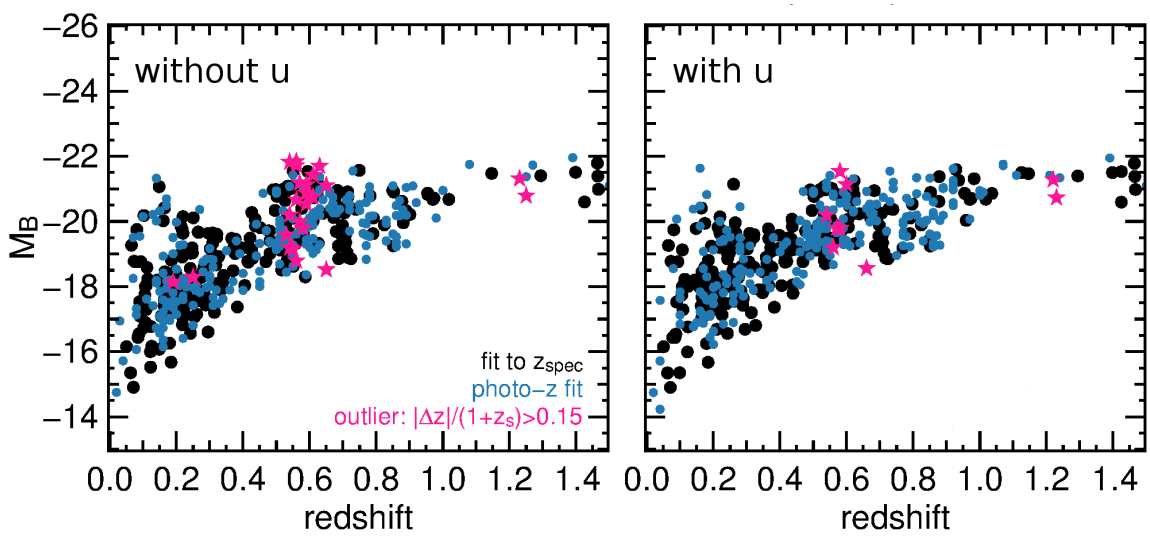

Figure 7.2: Example for the prior probability adaption in z vs. M_B space. The left panel shows results without the u band, whereas the right panel has u included. Black dots are derived from the model when fitted to the spectroscopic redshift and blue dots are from photometric redshifts. Magenta stars highlight outlier in the photo-z estimation.

Photometric Redshift Results

In the following we present the photometric redshift results obtained in the three tests, and focus on the *main-main* test, for which we also show the corresponding plots. We omit to include plots for all tests to prevent a cluttering of this chapter. The resulting quality metrics (already introduced in Secs. 5.3 and 6.3, see also Appendix A for a summary) of the *main-main* test are summarized in Table 7.2 and are calculated using the weighting scheme of Lima et al. (2008) discussed above and in Sánchez et al. (2014). Furthermore, we cut off 10 % of the galaxies with highest estimated photo-z errors which is allowed by the DES photometric redshift requirements. The bias for both runs (with and without u) is relatively high, but does not differ remarkably from one another. We can see, however, a significant improvement on the scatter when including u band data. The quality of the errors estimated by the `PhotoZ` code are also very similar in both setups. The average of the normalized error, which is the ratio between actual and estimated photo-z error $\Delta z' = \Delta z/\delta z$ (where δz was defined in Sec. 2.2.2, Eq. 2.10), has a relatively high negative value, but the scatter is nearly that of a normal distribution. The fact that the average value of $\Delta z'$ is negative means that the estimated photo-z error does not cancel out the bias. The RMS of the normalized error is nearly one, which means that δz only slightly overestimates the real photo-z error. The *KS* value is the result of a Kolmogorov-Smirnov test, which is a measure for how similar the distributions $N(z_{\mathrm{phot}})$ and $N(z_{\mathrm{spec}})$ are. *KS* differs only marginally for runs with and without u band data. This is shown also shown in Figure 7.4, which we will discuss later.

We plot spectroscopic versus photometric redshift in Figure 7.3, where the upper panel shows photo-z results without using the DECam u band data, and the lower panel shows results including u. We observe that the photo-z results improve particularly for $z_{\mathrm{spec}} \lesssim 0.6$, where the u band gives important information about the spectral range below 4000 Å. The

	$\boldsymbol{\eta}$	$\langle \boldsymbol{\Delta z} \rangle$	$\boldsymbol{\sigma(\Delta z)}$	$\langle \boldsymbol{\Delta z'} \rangle$	$\boldsymbol{\sigma(\Delta z')}$	**KS**
without u band	9.8%	-0.029	0.14	-0.28	0.95	0.051
with u band	5.2%	-0.029	0.12	-0.26	0.95	0.052

Table 7.2: Photometric redshift results for *main-main* test without u band and including the u band. Definitions of the metrics can be found in Appendix A. They are calculated using the weights of Lima et al. (2008) discussed in the text.

4000 Å break moves between the g and r band at about $z \sim 0.4$, where the importance of u band information decreases. Still, we can see from Figure 7.1 (e.g., from the *green* SEDs) that u band data entails clues about the steepness in the blue parts of the spectra, i.e., the UV slope $\beta := \log(f_\lambda)/\log(\lambda)$ where $\lambda \in [1276, 2490]$ Å (Calzetti et al., 1994). Figure 7.3 also highlights the difference in the photo-z quality of red and blue galaxies. Galaxies best fitted by a red model, i.e., LRG models from G13 or *red PZstandard* models from Figure 7.1, are plotted in red, whereas the rest of the *main* sample is blue in the plots. The red galaxies generally have better photometric redshift estimates, which is due to the 4000 Å break. The blue galaxies on the other hand do not exhibit such strong features in their SEDs (cf. Figure 7.1) in the wavelength range observed by DES. Their spectral features that could serve as good photo-z indicators mostly consist of emission lines that are not resolved by broadband observations.

Figure 7.4 shows the redshift distributions $N(z)$ of z_{spec} and z_{phot} with and without u band data. While we often concentrate on predicting the redshifts of individual objects correctly, for some science, e.g., cosmic shear studies, the overall redshift distribution is of greater interest. In contrast to the estimation of individual redshifts, we cannot observe an improvement on $N(z)$ when we include u band data in the photo-z estimation. This is mostly because the errors in the photometric redshifts without u are not systematic offsets and small in comparison to the width of the overall distribution, and therefore only slightly change the shape of $N(z)$ in comparison to the results including u.

In Figure 7.5 we show the deviations in magnitude Δmag of the predicted model magnitudes from the validation data when fitted to the spectroscopic redshift and after applying our zero point offsets. The upper panels present the results without the u band data, whereas the lower panels include u. We fit a Gaussian to every Δmag histogram and include the resulting parameters in the respective panels. We see that through our zero point offsets we achieve very small mean Δmag values of order $\mathcal{O}(10^{-3})$. This is reassuring and means that our model set is able to represent the data. The histograms of the r, i, and z band magnitude offsets are very well fitted by a Gaussian, whereas we observe higher flanks in the histograms of the other filter bands, which is likely due to an increase in the photometric measurement errors.

Table 7.3 presents the photo-z quality metrics for the *deep-deep* and the *deep-main* tests. We observe that deviations in the metrics do not differ more than at the 10% level from the *main-main* test in most cases. The results in outlier rate and bias improve when we consider higher S/N data in the *deep-deep* test, whereas the scatter is more or less the same. The absolute values of $\langle \Delta z' \rangle$ slightly decreases compared to Table 7.2, and $\sigma_{\Delta z'}$ is also nearer to the desired value of 1. The KS test value worsens in comparison with the *main-main* test in the case without u, and stays the same for the run with u band data.
In the *deep-main* test also the first two metrics improve, whereas the scatter stays the same

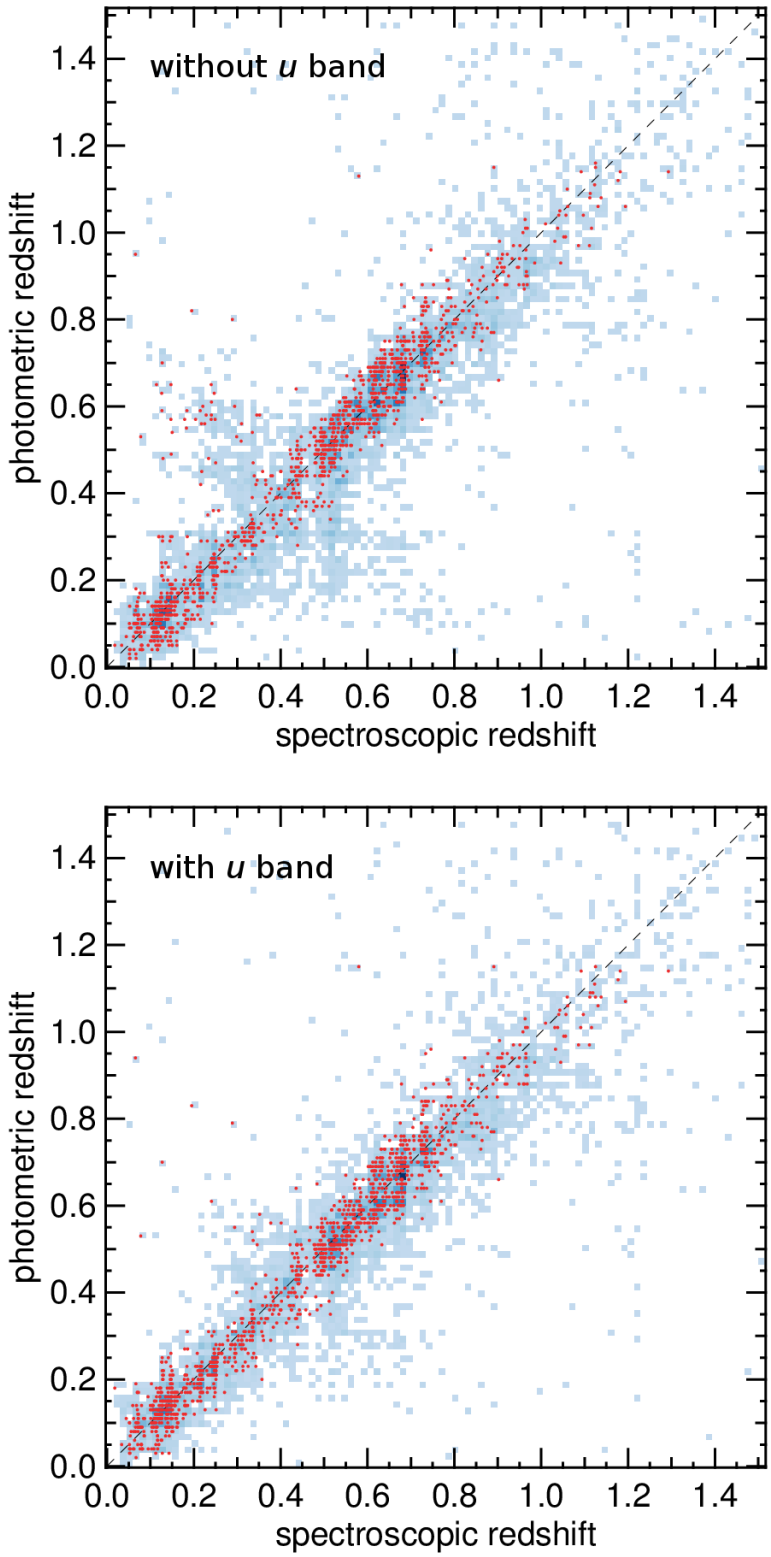

Figure 7.3: Spectroscopic versus photometric redshift results of the *main-main* test on the DES-SV data without using u band data (upper panel) and including u (lower panel). Galaxies with best fitting red models (LRG models and *red PZstandard* models in Fig. 7.1) are highlighted in red, whereas the rest of the sample is plotted in a blue mesh.

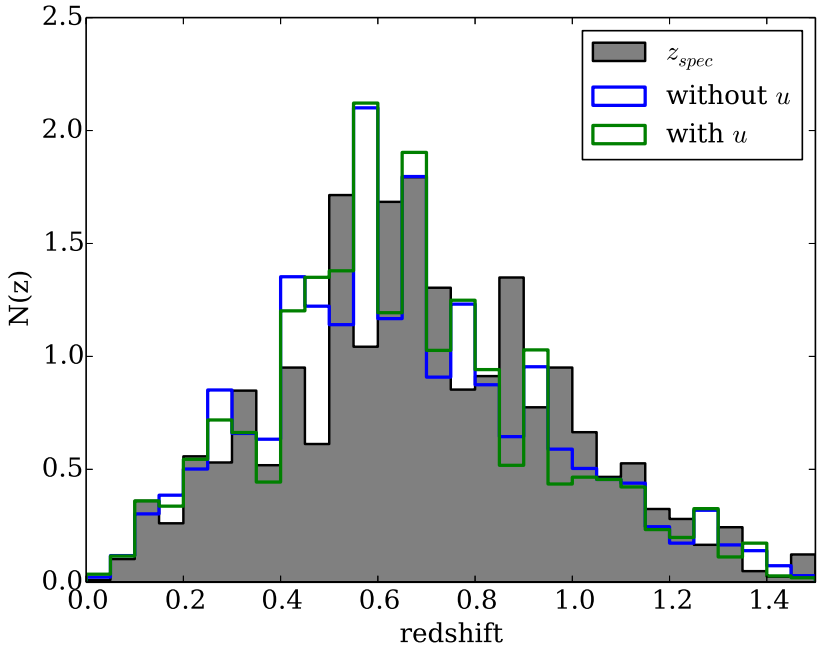

Figure 7.4: Photometric redshift distribution for *main-main* test without u band (blue) and including the u band (green). The spectroscopic redshift distribution is plotted in gray.

compared to the calibration on the *main* data. The bias in $\langle \Delta z' \rangle$ worsens slightly, whereas $\sigma_{\Delta z'}$ improves to a very small degree. Finally, the *KS* values show an increase in their values compared to Table 7.2.

	η	$\langle \Delta z \rangle$	$\sigma(\Delta z)$	$\langle \Delta z' \rangle$	$\sigma(\Delta z')$	KS
			deep-deep test			
without u band	8.7%	−0.020	0.15	−0.25	1.0	0.065
with u band	3.9%	−0.017	0.12	−0.24	1.1	0.052
			deep-main test			
without u band	9.7%	−0.031	0.14	−0.30	0.96	0.056
with u band	5.1%	−0.026	0.12	−0.28	0.96	0.055

Table 7.3: Photometric redshift results for *deep-deep* and *deep-main* tests without u band and including the u band. Definitions of the metrics can be found in Appendix A. They are calculated using the weights of Lima et al. (2008) discussed in the text.

Summary

In comparison to the other template fitting codes tested in Sánchez et al. (2014), our template-prior combination is among the best $\sim 30\,\%$ concerning bias and scatter. Our photometric redshift results can reproduce the $N(z_{\rm phot})$ distribution best compared to the other investi-

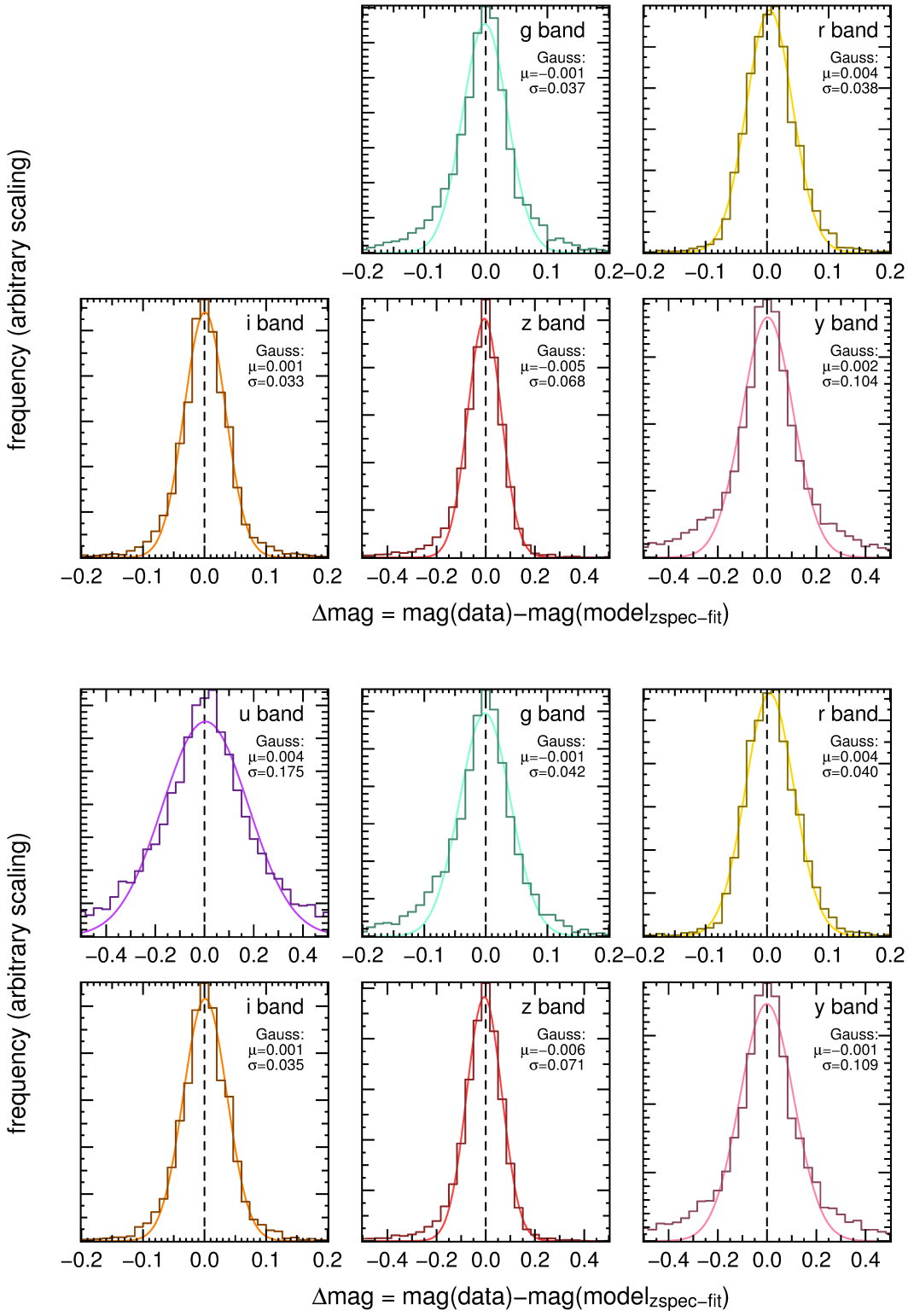

Figure 7.5: Deviations of predicted magnitudes from the model SEDs to the validation *main* data after zero point calibration on the *main* calibration set. The upper five panels show Δmag for a calibration without u band, whereas the lower panels use the u band data. The curves represent Gaussians when fitted to the histograms of Δmag. μ and σ of the Gauss curves are given in each panel.

gated template fitting methods.
In conclusion we can say that photo-z results of higher S/N data are more accurate than of lower S/N data, which is not surprising given the precision of the data. The errors of the photometric redshifts show no systematics, i.e., are not systematically over- or underestimating the real errors. However, the absolute values of the bias cannot be decreased by taking the estimated errors into account, hence the negative values of $\Delta z'$. The metrics describing how close the $N(z_{\mathrm{phot}})$ and $N(z_{\mathrm{spec}})$ distributions are, worsen when we consider the deeper calibration data, except for the KS value in the *deep-deep* run if we include u band data in the photo-z estimation. In the *deep-main* test this is probably due to the fact that we calibrate on deeper data than the validation data and therefore these two samples differ from one another.
Finally, in all runs the inclusion of u band data significantly improves the quality of the individual photo-z predictions in terms of catastrophic outlier rate, bias, and scatter. u band information is specifically important for any template based photo-z code to be able to distinguish between red and blue galaxies on account of their fluxes redwards of 4000 Å for lower redshifts, but it also includes clues about the steepness in the blue part of the spectrum for higher z.

Chapter 8

Physical Properties of a Young $z \sim 6$ Galaxy Quintuply Lensed by the Cluster RXC J2248.7-4431

This chapter presents work I have done to estimate the physical properties of a galaxy lensed by the cluster RXC J2248.7-4431. This section was published as part of a strong lensing analysis led by Anna Monna, Stella Seitz, Natascha Greisel, and Thomas Eichner, et al. in the *Monthly Notices of the Royal Astronomical Society* (Monna et al., 2014, hereafter Mo14). We refer to Mo14 for a detailed description of the data reduction, photometric and strong lensing analyses of the high-z quintuply lensed galaxy. We also refer to Balestra et al. (2013), Boone et al. (2013), and Karman et al. (2014) for details about the spectroscopic confirmation of the $z \sim 6$ nature of this source and the properties inferred by its observed spectra. Here we want to give a very brief introduction into gravitational lensing and a short summary of the strong lensing part of Mo14.

Light is bent while passing through the gravitational fields of masses (*gravitational lensing*). Therefore, if a massive object lies in the line of sight between a light source and an observer, the image of the source will be affected. The mass structures, the *gravitational lenses*, can be other galaxies, galaxy clusters, or even stars or planets. Gravitational lensing will cause a distortion and magnification of the image or even create multiple images of the background source (*strong lensing*). For a very conclusive introduction to gravitational lensing we would like to refer to Schneider et al. (2006).
Lensing analyses are used to derive the projected total mass distribution of the structures acting as gravitational lenses. Strong gravitational lensing allows to reconstruct the density profile of galaxies and galaxy clusters. In parametric lensing techniques, analytic mass profiles are used to describe the cluster smooth dark matter halo and the structures within.
In Mo14, a strong lensing analysis is performed to reconstruct the cluster mass profile of the cluster RXC J2248.7-4431. This is a very massive cluster with a virial mass[1] of $M_{200} > 2.5 \cdot 10^{15}\, M_{\odot}$ (Guzzo et al., 2009), which is located at a redshift of $z \sim 0.384$. It was observed as part of the *Cluster Lensing And Supernovae survey with Hubble* (CLASH,

[1]The virial radius r_{200} is the radius of a sphere with mean mass density is $\rho = \Delta_c \rho_{crit} = 200 \rho_{crit}$. The virial mass M_{200} is then the mass enclosed in a sphere with radius r_{200}, $M_{200} = 4\pi/3 \cdot \Delta_c \rho_{crit} \cdot r_{200}^3$.

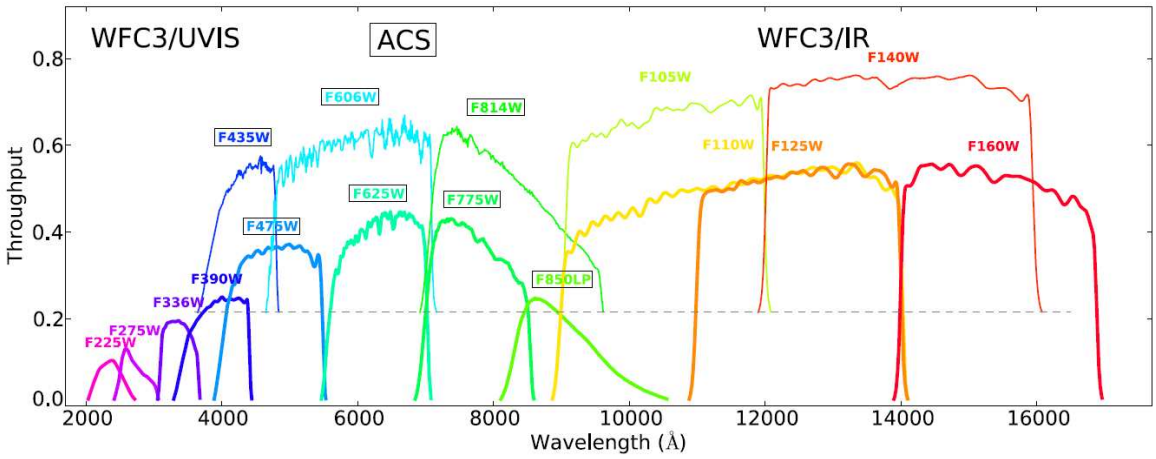

Figure 8.1: CLASH filter transmission curves.

Postman et al., 2012) targeting 25 clusters with the *Hubble Space Telescope* (HST) and combining ground- and space-based photometry for the reconstruction of the cluster mass profiles. Photometric data was taken through 16 overlapping broadband filters ranging from ~ 2000 to $\sim 17,000$ Å, which are shown in Figure 8.1. Mo14 identified five lensed images of a star forming galaxy in the core of RXC J2248, employing the dropout technique (cf. Sec. 2). An image of the inner part of the cluster and the five images of the background galaxy are given in Figure 8.2. The central image lies in the cluster center, and is detectable only after removing the brightest cluster galaxy (BCG) in the NIR images. Unfortunately, the photometry of ID1 is strongly contaminated by the nearby galaxy, whereas ID4 is detected only at a 3σ level in the first detection band $f814$ at ~ 814 nm. From the five images of the lensed galaxy, ID2 and ID3 have the best photometry, which is detected at a 5σ level in the first detection band $f814$. The photometric redshift estimate for the lensed galaxy lies at $z \approx 5.88$ and $z \approx 6.04$ for the two images (ID2&3 henceforth) with non-contaminated photometry. The photometry is shown together with the photometric redshift results in Figure 8.3. The object was spectroscopically confirmed after the publication of Mo14, and the spectroscopic redshift was measured to $z_{\rm spec} = 6.11$ for the images ID2-4 (Balestra et al., 2013; Boone et al., 2013; Karman et al., 2014).

In the following we estimate the physical properties of the lensed system by SED fitting to the observed photometry in the CLASH filters with `SEDfit`. The `SEDfit` algorithm was explained in Section 5.1. We use the approximate mean photo-z value of the four images (ID1-4) of $z_{\rm phot} = 5.9$ for the SED fitting. As noted already above, the images ID2 and ID3 are the ones with the best photometry, which do not suffer from any contamination by nearby sources, so we focus on these images for the SED fit. As basis for the template set for the fitting we chose the SSP model SEDs of BC03 with a Chabrier IMF (Chabrier, 2003) and Padova 1994 evolutionary tracks. From these we create CSPs with the software `GALAXEV`. As mentioned in Section 4.2 the SFH of galaxies is commonly described by a so-called τ *model* that follows:

$$\psi(t) \propto \exp(-t/\tau), \tag{8.1}$$

where $\psi(t)$ is the star formation rate, and τ the (positive) e-folding timescale. t is the time

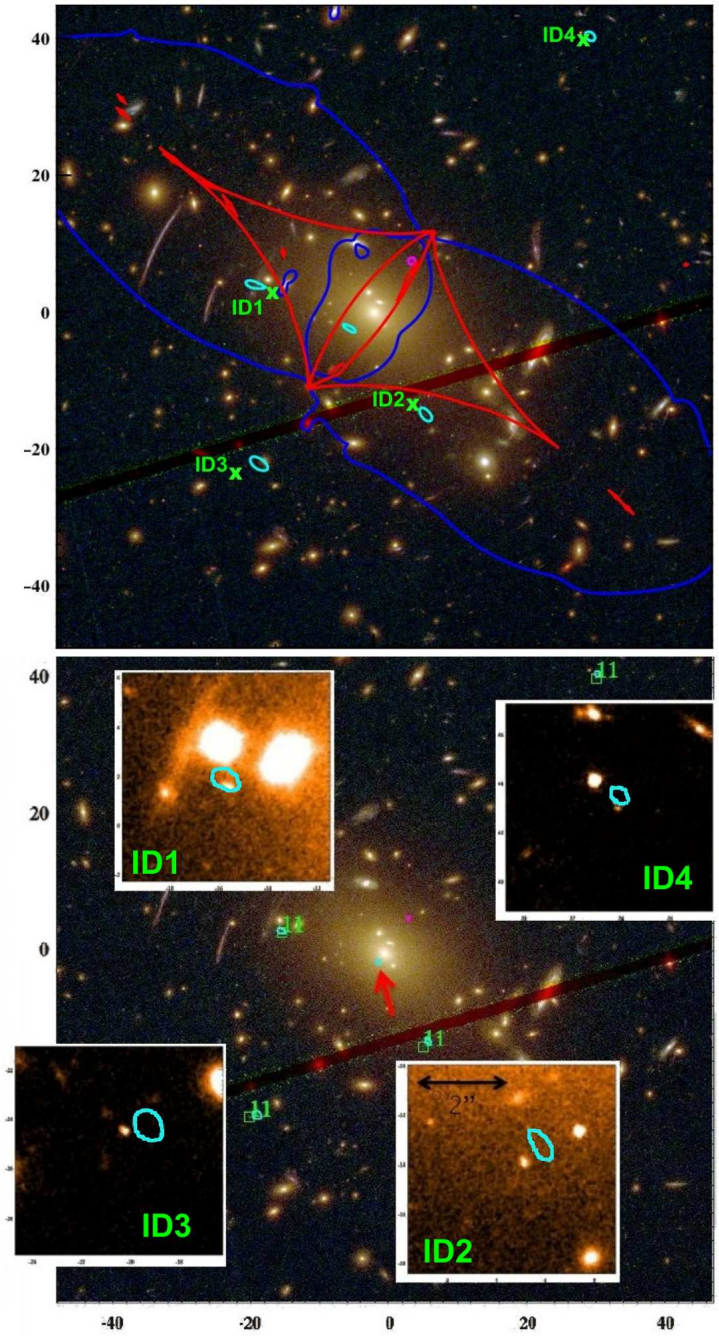

Figure 8.2: HST color composite images of the inner region of RXC J2248. *Upper Panel*: critical lines (blue) and caustics (red), for the source at $z \sim 6$, for the best strong lensing model. The green squares show the ID1-4 positions, while in cyan we show the images that the strong lensing model predicts for the source. *Lower Panel:* Predictions of the images for the system ID1-4 from the final best model. The green squares are the observed positions and the cyan contours are the predicted images for the source. We show the zoom on the predicted images ($8.0'' \times 8.0''$) in the HST/$f110w$ image. The red arrow indicates the position of the 5th central image predicted for this system.

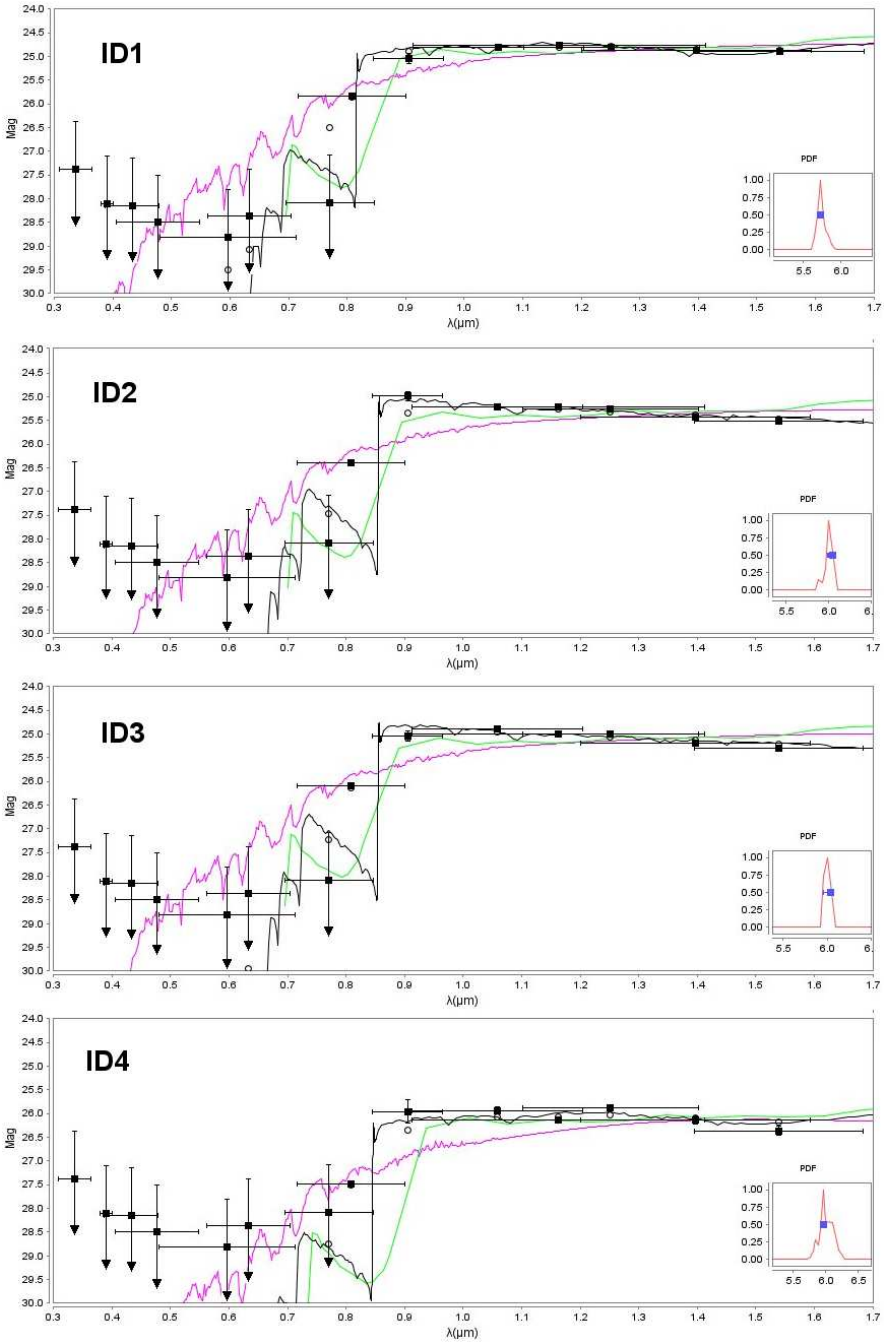

Figure 8.3: Photometric redshift results with the software `LePhare` (Arnouts et al., 1999; Ilbert et al., 2006). Black squares are the observed magnitudes, open circles are the predicted magnitudes, and arrows represent upper limits in the detection. The black, green and magenta lines are the galaxy, quasar and stellar templates respectively. The PDF of the redshift for the galaxy template is given in the lower right corner of each fit.

that has elapsed since the start of star formation, i.e., the age of the galaxy. This function describes well the SFHs of local galaxies, but likely does not hold for $z \gtrsim 1$ (e.g., Maraston et al., 2010; Lilly et al., 2013). In fact, Maraston et al. (2010) showed that fitting model SEDs with exponentially decreasing SFRs to star-forming galaxies at $z \sim 2$ yields unrealistic young ages because the galaxy spectrum is then dominated by the young stellar population. Moreover, they showed that τ models with negative τ lead to more physical results.

Here we create CSPs for τ models with both positive and negative τ values, and therefore decreasing and increasing SFR, as these should set lower and upper limits to the SFR of the investigated galaxy. We generate CSPs with metallicities of $Z = 0.005$, 0.02, 0.2, 0.4, 1, $2.5 Z_\odot$ and e-folding timescales of $\tau = \pm 0.01$, ± 0.1, ± 0.5, ± 1.0, ± 2.0, ± 3.0, ± 4.0, and ± 5.0 Gyr for the τ model. The created CSPs are extracted at 24 different ages equally distributed in logarithmic space between 0.1 Myr and 3 Gyr. Additionally to the CSPs, we extract SSPs with the same metallicities and ages. For model ages that are higher than the age of the Universe at redshift 5.9 the SED fitting code will assign probabilities of zero. The extinction is allowed to take values between $A_\mathrm{V} = 0.0$ and 3.0 with steps of 0.1. In this case we do not include a burst model, since the galaxy is expected to be very young (the age of the Universe at $z = 5.9$ is ~ 1.0 Gyr).
We perform the SED fitting for model sets containing the SSPs and CSPs with increasing and decreasing SFH separately, and afterwards with all models combined. The results are summarized in Table 8.1.
Using the SSPs models, the best fits yield subsolar metallicities (0.2 and $0.005\, Z_\odot$ for ID2 and ID3 respectively) and very young ages (0.1 and 1.5 Myr for ID2 and ID3 respectively). The results when using the CSPs with increasing and decreasing SFR are similar to one another. The best fits yield the same values for the metallicities (0.2 and $0.005\, Z_\odot$) and comparable small ages (0.5 and 1.5 Myr). In summary, for the three different model sets (SSPs, and CSPs with $\tau \leqslant 0$) the results for metallicities, extinctions, and ages are essentially the same (except for the age of ID2, which is 0.1 Myr for the CSPs, and 0.5 Myr for the SSPs). Actually this is what we would expect given the small ages resulting from the best fits. Within these short timescales ($t \ll \tau$) the galaxies could not evolve significantly which is why the stellar populations (and therefore SEDs) of the CSPs are very similar to one another and also to the SSPs. When including all models (CSPs and SSPs) in the SED fitting, we get that ID2 is better fitted by a CSP model with increasing SFR, while an SSP model is preferred for ID3. In any case, the differences in the stellar populations are, as mentioned above, not very high at these young ages which is why the results for ID2&3 are in good agreement.
All the results are shown in greater detail in Appendix D, Figures D.1 to D.8, where we plot the best fitting SEDs and the marginalized likelihood distributions in parameter spaces, as well as in Figure D.9 which displays the PDFs of the model ages. Through interpolation of the age PDF, we calculate that ID2&3 have ages within the interval of $[0.1, 330]$ Myr (centered on the PDF) at a 2σ confidence. The same was done for the masses, which we estimate to be within $[0.3, 7.5] \cdot 10^8 \mathrm{M}_\odot$ at the same level of confidence. The best fitting masses and the age and mass intervals are also summarized in Table 8.1. The reduced χ^2 values of all fitting results lie between 10 and 12.7 and differ only marginally when the underlying model set is changed.

We furthermore estimate the UV slope β (Table 8.1) from the best fitting SED of each run by a linear fit of $\log(\lambda)$ versus $\log(f_\lambda)$ within $\lambda \in [1276, 2490]$ Å (see Calzetti et al., 1994),

ID	τ [Gyr]	Z [$Z_\odot$]	age [Myr]	$M_\star$ [$10^9\,M_\odot$]	β_{UV}	A_V
			SSPs			
ID2	SSP	0.2	0.1 [0.1, 45]	0.093 [0.17, 0.28]	-2.90 ± 0.02	0.2
ID3	SSP	0.005	1.5 [0.1, 111]	0.21 [0.034, 0.79]	-2.64 ± 0.02	0.4
			$\tau < 0$ *model CSPs*			
ID2	-1.5	0.2	0.5 [0.1, 274]	0.093 [0.034, 0.25]	-2.90 ± 0.02	0.2
ID3	-0.1	0.005	1.5 [0.1, 301]	0.23 [0.059, 0.71]	-2.62 ± 0.02	0.4
			$\tau > 0$ *model CSPs*			
ID2	0.1	0.2	0.5 [0.1, 291]	0.093 [0.034, 0.27]	-2.90 ± 0.02	0.2
ID3	0.1	0.005	1.5 [0.1, 383]	0.23 [0.059, 0.78]	-2.62 ± 0.02	0.4
			all combined			
ID2	-1.5	0.2	0.5 [0.1, 270]	0.093 [0.033, 0.26]	-2.90 ± 0.02	0.2
ID3	SSP	0.005	1.5 [0.1, 330]	0.21 [0.058, 0.75]	-2.64 ± 0.02	0.4

Table 8.1: Best fitting parameters from SED fitting CSPs, SSPs, and all models combined. Positive values for τ stand for an exponentially decreasing SFR with $\psi(t) \propto \exp(-t/\tau)$, while negative values denote an increasing SFR. We interpolate the PDF of model ages (see also Fig. D.9 in Appendix D) and calculate the age interval within which the probability of a fit is 95.45 % (corresponding to a 2σ confidence interval). These 2σ intervals are given in brackets in the column *age*. The same was done for the stellar masses, where the corresponding intervals are in column $M_\star$. Note that the age and mass values of the best fitting model do not necessarily lie in the calculated intervals.

getting results which are in very good agreement with $\beta = -2.89 \pm 0.25$ estimated using the observed NIR colors in Section 4 of Mo14.
We also combine the likelihoods from the SED fits of the two candidates, and the results are the same as the previous ones, with equal constraints on ages and masses. Moreover, we repeat the SED fits with SSP and CPS, using this time the combined photometry of the two lensed images ID2&ID3, and the SED fits lead to same constraints on masses and ages.
We supplement our photometric dataset with mosaic mid-IR imaging data from Spitzer obtained with the *Infrared Array Camera* (IRAC) in channel1 (3.6 µm) and channel2 (4.5 µm), from the survey Use of massive clusters as cosmological lenses (PI: G. Rieke, program ID 83). We analyze whether the inclusion of additional information through the shallow data from the IRAC of the Spitzer Telescope affects our results. We repeat the previous SED fitting but include the upper limits estimated in the 3.6µm and 4.5 µm filters in the photometry. In this step we only focus on ID3, since the ID2 IRAC photometry is significantly contaminated by the nearby cluster members. The best fitting SSP model has an age of 1.5 Myr within the 2σ interval of [0.1, 30] Myr and mass of $2.2 \cdot 10^8\,M_\odot$ within $[0.4, 5.2] \cdot 10^8\,M_\odot$. For the increasing and decreasing SFH we get respectively $\tau = -2.5$ Gyr and $\tau = 2.5$ Gyr, ages of 2.5 Myr within [0.1, 222] Myr and 2.5 Myr within [0.1, 212] Myr, and masses of $2.2 \cdot 10^8\,M_\odot$ within $[0.6, 4.7] \cdot 10^8\,M_\odot$. In these cases, the best fitting models require a higher τ than the previous results, in which we get $\tau = \pm 0.1$ Gyr. However, in Appendix D (Figures D.1 to D.4 for increasing and decreasing SFH models) we show that the probability distributions of τ are very flat, so that slight changes in the input photometry can change the best fitting

	τ [Gyr]	Z [$Z_\odot$]	age [Myr]	$M_\star$ [$10^9\ M_\odot$]	β_{UV}	A_V	Ref.
R11	SSP	0.005	3.5 $[2.9, 1.3 \cdot 10^3]$	0.2 $[0.3, 11.4]$	-2.03 ± 0.02	0.8	Mo14
	0.5	—	$[640, 940]$	$6.3^{+2.8}_{-1.2}$	-2	—	R11
Z12	0.1	0.005	300 $[4.7, 700]$	1.3 $[0.2, 1.5]$	-2.32 ± 0.02	0.0	Mo14
	—	0.5	180	1	-2.50 ± 0.06	—	Z12

Table 8.2: Best fitting parameters for the objects of R11 and Z12 in the clusters A383 and MACS0329 respectively. Columns are the same as in Tab. 8.1, and the last column gives the reference.

values for τ very easily. Finally, when combining all the models, the best fit is provided by the SSP model with age of 1.5 Myr within [0.1, 211] Myr, and best mass of $2.2 \cdot 10^8\ M_\odot$ within $[0.6,\ 4.8] \cdot 10^8\ M_\odot$. The metallicities and the dust contents of all the cases remain the same as in the previous results. Thus, the inclusion of IRAC upper limits for ID3 leads to best fitting parameters which are in agreement with the previous ones for all the cases, and to slightly smaller 2σ intervals for the age and mass.
It it planned to observe the cluster RXC J2248 with the Spitzer telescope[2] in the future. These observations will give photometry in the visible wavelength region in the rest frame of the spectrum, and will enable us to constrain the galaxy properties (like age and stellar mass) more accurately.

To compare our results with properties of young galaxies from the literature, we perform the same fitting procedure (using CSPs and SSPs models) on other known $z \sim 6$ lensed sources selected in the field of the CLASH clusters (i.e., Richard et al., 2011; Zitrin et al., 2012; Bradley et al., 2014), for which we have the same photometry as for our candidate. The SED fitting performed in our work uses similar model parameters as Zitrin et al. (2012, hereafter Z12). The main difference is in the lowest model ages, which is 0.1 Myr in our work and 5 Myr in Z12. For the $z \sim 6.2$ quadruply lensed galaxy in the field of MACS0329, we get results that are consistent with Z12, for a low mass ($M \sim 10^9\ M_\odot$) young galaxy. Our best age (300 Myr) is slightly higher, but still consistent within the 2σ confidence level.
Richard et al. (2011, hereafter R11), unlike us, adopt in their SED fitting models a Salpeter IMF and smaller ranges for the metallicity (Z within $[0.2,\ 1]\ Z_\odot$) and ages (within 10 Myr and 1 Gyr). Our best age is much younger than the age range predicted in R11, although our 2σ confidence level age interval of $\sim [3, 10^3]$ Myr covers the age range of $[640, 940]$ Myr given in R11. As a consequence, also the stellar mass is found to be quite different, given that our best value is much smaller than the mass estimated by R11 (although they are anyhow consistent within our 2σ confidence level mass interval). In Table 8.2 we summarize our results for these two lensed systems, and provide the physical properties estimated in the reference works.

We also perform the SED fitting for the 208 galaxy candidates at $z \sim 6$ from Bradley et al. (2014, hereafter B14). We plot in Fig. 8.4 the best fitting ages versus masses for these candidates, together with the results we get for our system, R11 and Z12. Our candidates have age and mass similar to many young candidates from B14. In Figure 8.5 we show the histograms for best age, with the lower and upper age limits, for B14 and our candidates.

[2] http://ssc.spitzer.caltech.edu/warmmission/scheduling/approvedprograms/ddt/frontier/

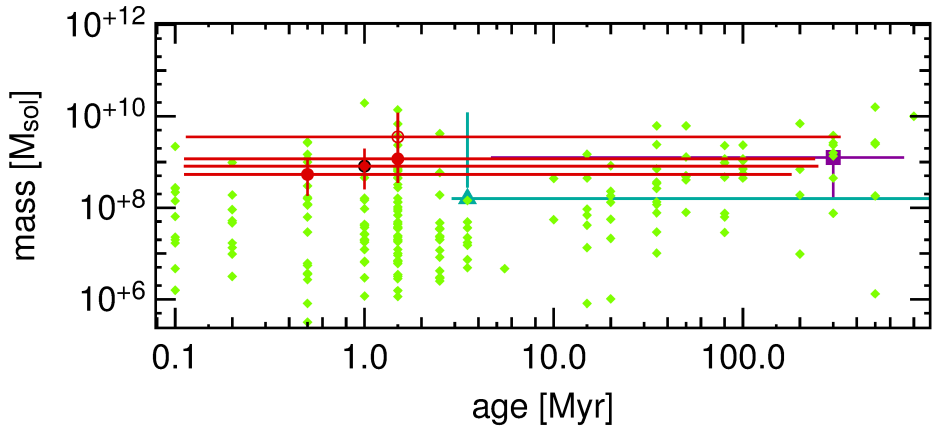

Figure 8.4: Age and mass results for SED fitting runs with all SFHs combined on the objects from the literature and the system discussed in this paper. The plot shows the objects of R11 (green triangle), Z12 (purple square), B14 (green diamonds), and ID2&3 (red points). Error bars show the intervals in mass and age within which the total probability reads 95.45 % (corresponding to a 2σ confidence interval). Note that the best fitting values therefore do not necessarily lie within these intervals. To improve on the clarity of the figure we did not plot the intervals for the B14 objects.

Albeit the 2σ confidence level intervals are large, we see that our multiple lensed system is definitely among the young sample of $z \sim 6$ galaxies. We also estimate the UV slope β from the best SED fitting templates for all the sources and we get that our candidate belongs to the sample of galaxies with very steep UV slope ($\beta \sim 2.6 - 2.9$). Moreover, despite the large 2σ confidence level intervals on the ages, our SED fitting suggests a very young best age for our system, compared to the other $z \sim 6$ multiply lensed galaxies (Z12 and R11), while the Z12 and R11 have confirmed older ages, in agreement with the results from the literature.

Summary

We estimate the physical properties of the high-z system, lensed by the galaxy cluster RXC J2248.7-4431, which may very well be a progenitor of present massive ellipticals. We perform a SED fitting of the observed photometry of the system, using SSP and CSP SED templates, and derive that our candidate has subsolar metallicity ($Z < 0.2\,Z_{\odot}$), low dust content ($A_V \sim 0.2 - 0.4$), stellar mass $M_{\star} \sim 108\,M_{\odot}$ and best age of ~ 1 Myr. Although the age is not well constrained, we can set an upper limit of ~ 300 Myr, given the 2σ uncertainties we get. The young age is an indication that the high-z galaxy could be a progenitor of present massive elliptical galaxies. We verify that including the shallow IRAC photometric upper limits in the SED fitting leads to similar results, with slightly smaller 2σ intervals on the mass and age. Finally, we compare our predicted ages and masses, with the $z \sim 6$ candidates selected up to the date of the publication in the field of all CLASH clusters, and we find that our multiply lensed galaxy has a young age and low mass, similar to some of the objects from B13. However, compared to the other known multiply lensed $z \sim 6$ galaxies, our SED fitting suggests a younger age for our multiple lensed system. This source adds to the several multiple lensed objects known at high-z, characterized for being five times lensed, with a central image identified in the very inner region of the lensing cluster.

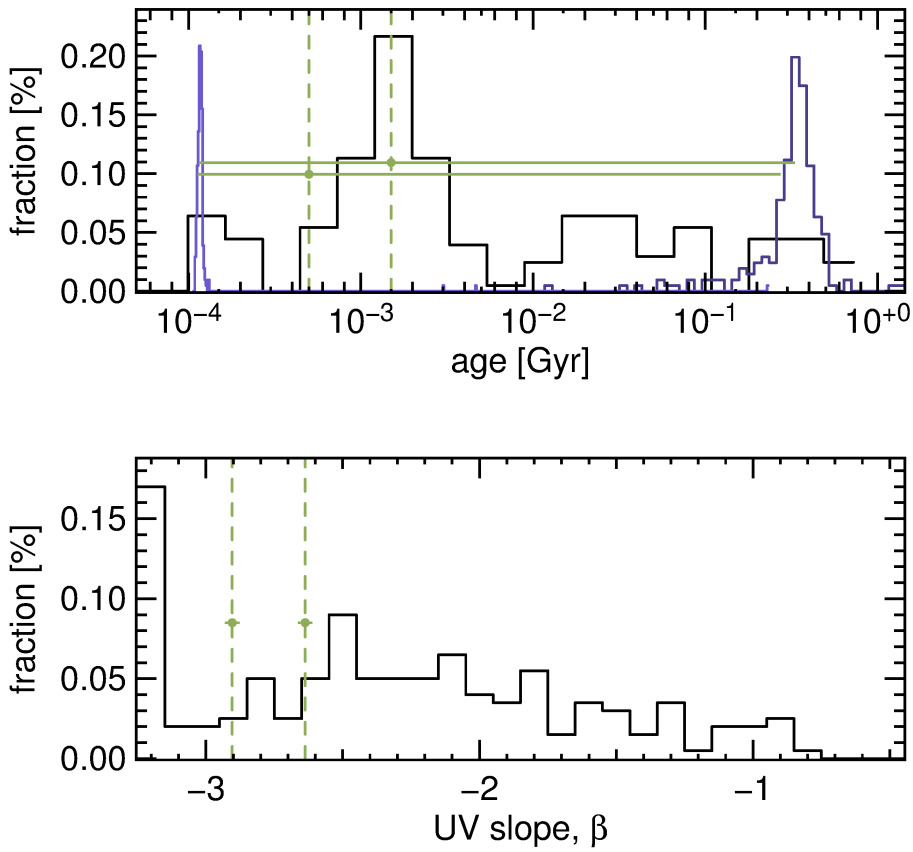

Figure 8.5: *Upper panel*: Histogram of the best ages for the $z \sim 6$ candidates from R11, Z12, and B14 (in black). We plot in lighter and darker blue the age lower and upper limits respectively for the 2σ confidence level intervals of the $z \sim 6$ candidates. In green we plot ID2&3. The 2σ confidence level intervals are too large to well constrain the ages of the sources, however our multiple lensed system surely belongs to the young sample of $z \sim 6$ galaxies. *Lower panel:* Histogram of the UV slope β as measured from the best fitting SED templates for the B14 candidates (in black) and for our high-z galaxy (in green).

Chapter 9

Summary and Conclusion

The main goal of this thesis was the estimation of accurate photometric redshifts with template fitting methods through an improvement of the template set. Concurrently, we aimed to improve our knowledge on the physical and the rest frame properties of the studied galaxy samples. We therefore searched for SEDs with known properties that matched the photometry of the galaxies. Furthermore, we tried to find small model sets that represent the analyzed galaxy samples while providing accurate photo-zs when used in template fitting codes.

In the beginning we analyzed the colors as functions of redshift of public model SEDs that are used with photometric redshift template fitting codes in comparison to SDSS-II LRG data. We found that no model was able to produce the colors of the LRGs throughout the redshift range of the sample. Even though the passively evolving M09 model showed promising results, it neither could match all the colors nor could it account for their spread. This gave raise to the need of new model SEDs that can match the SDSS-II LRG colors and for SEDs that are specifically designed for different redshifts and hence galaxy populations. We therefore performed fits of SSPs and CSPs of BC03 with decreasing SFHs to the photometry of the SDSS-II LRGs with the software `SEDfit`. Apart from the SFH, the fitting process included extinction by dust, a burst model in superposition and extinction of the burst as degrees of freedom. We performed the fit on five different redshift intervals to account for evolution with redshift. From this large set of best fitting models we chose a much smaller subsample that represents the data within a grid that includes the region in the $u-g$ vs. $r-i$ color spaces which is densely populated by the LRGs. We additionally selected models outside the grid to account for galaxies with peculiar colors.
The selected models were then used in a photo-z estimation with the Bayesian template fitting code `PhotoZ`. The overall bias in the photo-z error is very small $\langle \Delta z \rangle = -0.0004$, but it lies between -0.01 and 0.01 when tracked as a function of redshift. Furthermore, the fraction of "catastrophically" wrong z_{phot} estimates is small $\eta = 0.12\,\%$, as is the RMS scatter $\sigma(\Delta z) = 0.027$. The SDSS published photo-z results from a template fitting approach and two empirical methods. Compared to the redshifts of SDSS, the bias of the photo-zs of this work changes much less strongly as a function of z which is important for any science application in which the redshifts are used.
Since we created the model SEDs without any prior knowledge about their properties at wavelengths bluer than the SDSS u band, we compared the predicted UV colors to GALEX

observations of early-type galaxies at $0.05 \leq z \leq 0.12$ of R12. We found that the UV colors of our LRG models match the colors of the quiescent, star-forming, and composite samples of R12. After that, we analyzed the behavior of the new models when fitted to the spectroscopic redshift of the LRG sample as functions of luminosity and redshift. For local LRGs with $z \leq 0.1$ we observed that the best fitting models show increasing fluxes in the blue part of the spectrum with decreasing luminosity. The fluxes in the blue spectral region are mostly caused by young stars in the galaxy. We confirmed this observation by inspecting the observed SDSS spectra from the same galaxies which exhibit the same trend. Another confirmation was obtained by investigating the fraction of LRGs that show signs of star formation on account of their line indices (from the SDSS spectra), i.e., the LRGs with $\log_{10}(\mathrm{OIII/H}\beta) < 0.7 - 1.2(\log_{10}(\mathrm{NII/H}\alpha) - 0.4)$.

Afterwards, we generated model SEDs for the BOSS CMASS galaxies which are similar to the SDSS-II LRGs, but extend to higher redshifts $z \lesssim 0.9$. We analyzed the colors of the M09 model for different formation redshifts and also constant ages and observe (similar to the SDSS-II LRGs) that it cannot produce the same color range that the data shows and that it differs from the data in some redshift ranges and colors. In order to create a set of template SEDs for photo-z estimation we again fitted model SEDs to the data similar to SED fit previously performed on the SDSS-II LRGs. The mismatch in colors of the best fitting models and the data, which was already present for the LRG sample, is even more prominent for the CMASS galaxies. It is predominant in the $i - z$ color, wherefore we modified the red continuum slope in order to match the data. The observed $i - z$ colors require smaller values (bluer SEDs) and we modified the models by multiplication with $\lambda^{-\beta}$ for $\lambda \geq \lambda_i$. Introduced in the SED fitting, the best fitting $\lambda^{-\beta}$ models match the color space inhabited by the CMASS galaxies to a much better extent than the original BC03 models.
We selected a subset from the best fitting models which should represent the data in terms of their colors and absolute magnitude. Therefore, we projected the four-dimensional space spanned by $g - r$, $r - i$, $i - z$, and M_r onto a plane with a self organizing map. We then used a k-means clustering algorithm to identify six clusters of the data in the SOM and create Voronoi cells around them. From each cell a model which yields accurate photometric redshift results was selected for the later photo-z template set. Afterwards, we analyzed the photo-z performance on the whole CMASS sample for each model separately and concurrently adopted redshift priors where it improved the redshift quality. Additionally, when a change in priors could not improve the bad photo-z quality of a model, we erased it from the final template set which consists of nine models in the end.
We then analyzed the photo-z accuracy of the new models in comparison to the two photo-z results from SDSS. The redshifts estimated with our novel model set outperform those published by SDSS in terms of bias and scatter. Furthermore, the redshift errors derived from the PDFs of the results from this work are much more reliable than the SDSS errors, albeit they are not reliable in classifying outliers.
Finally, we compared the colors that are predicted from our models with our photo-zs and those with the better of the two SDSS photo-zs respectively to the data. Also in this analysis the photometric redshifts from this work provide better results.

We furthermore estimated photo-zs for mixed samples of red and blue galaxies, where we present our photometric redshift results for the DES science verification data. The models of the SDSS-II LRGs were combined with three further red templates as well as a number of

SEDs showing few to very high star formation activity. The DES provided a *main* sample which was observed at the main survey depth, and a *deep* sample with deeper photometry. Recalibration of the photometric zero point offsets was performed on separate samples with *main* and *deep* photometry, after which we estimated the photo-zs for the query *main* and *deep* samples, yielding the combinations: *main-main*, *deep-deep*, and *deep-main*. The latter was a test to evaluate if the calibration/training on deep data could improve the photometric redshifts of the *main* catalog. We could not observe an improvement in the photo-zs of the *main* catalog when we recalibrated on the *deep* galaxies, in fact the KS test value increased slightly. Calibrating on and evaluating redshifts of the *deep* sample produced photo-zs that ware more accurate than those of the *main* sample in terms of their outlier fraction.
We also studied the impact of u band data on the photo-z accuracy which is important for the distinction of blue and red galaxies at lower redshifts. Indeed, including u band data in the $z_{\rm phot}$ estimation approximately reduced the outlier fraction by a factor of 2 and decreased the scatter noticeably.

In the last chapter we studied the physical properties of a quintuply lensed system located at high redshift $z \sim 6$. We found that the galaxy has subsolar metallicity and a low dust content. The SED fitting results suggest an age of $\sim$1 Gyr which is not well constrained, but is lower than 300 Myr at a 2σ confidence. Therefore, the lensed system is possibly a progenitor of massive red elliptical galaxies at of low redshift Furthermore, we derived a low stellar mass of $M_\star \sim 10^2 M_\odot$. Compared to results for high-z galaxies in the literature, our galaxy is similar to some of the galaxies investigated by B13, but has lower ages in comparison to samples from other publications.

We discussed the advantages and disadvantages of template fitting codes and empirical methods in Chapter 2. From Chapters 5 and 6 we learned that template fitting methods can outperform empirical approaches, although large representative spectroscopic training sets are available, if only the template SEDs match the data in question. However, some studies indicate that empirical approaches outperform template fitting methods, especially when mixed samples of blue and red galaxies are analyzed and the spectroscopic training sample is representative (Abdalla et al., 2011; Sánchez et al., 2014). This is mostly the case if the model set in template fitting methods is not a good representation of the data and/or the accuracy suffers from degeneracies in colors that are not lifted by the prior probabilities (Hildebrandt et al., 2010). In the future, a combination of empirical and template fitting approaches seem to be most promising in providing precise photo-zs, to incorporate the advantages of both methods. This can be done, e.g., by combining the $P(z)$ distributions of one or more empirical and a template fitting method, or by deciding the goodness of the photo-z estimate on the basis of the shape of the PDF (e.g., Carrasco Kind & Brunner, 2014a).
Furthermore, adaptive improvements of the redshift and luminosity priors of templates on the basis of a spectroscopic training set (or a set with accurate photometric redshifts) could improve the photo-z accuracy with the caveat that the priors would have to be retrained for different observations. As can be seen from Figures 5.9 in Section 5.5.2 and Figure6.1 in Section 6.1, the redshift and luminosity are not independent from one another, as treated by the function of the prior $P(\boldsymbol{\mu}|T) = P(M|T) \cdot P(z|T)$ (Eq. 2.7 in Sec. 2.2.2). Instead, the prior probability should be set up as a 2-dimensional function $P(\boldsymbol{\mu}|T) = P(M, z|T)$.
Estimating the physical properties of galaxies from photometry requires a good model for the

star formation history and therefore the stellar content. While we do not claim to be able to recover the physical properties from SDSS broad-band filters, we saw that there are systematic offsets of the colors from the models to the data with the considered model SEDs. These have to be addressed in a more physical way by changing the function for the SFR and/or by superposition of SEDs with different properties. We already took the latter into account when we fitted a star burst component in superposition to a main SED component, but the burst had a fixed metallicity, SFH, and age. In future analyses a more complex SFH could be considered, or multi-component SEDs (Díaz García et al., submitted 2014), since age or metallicity can vary within galaxies (e.g., Tamura et al., 2000; Sánchez-Blázquez et al., 2014). However, we expect that progress on the subject of galaxy population studies has the most prospect for spectroscopic or narrow-band photometric surveys, because of their ability to resolve spectral lines which are important clues for the derivation of the underlying stellar population.

Bibliography

Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Anderson, K., Anderson, S. F., Annis, J., Bahcall, N. A., Baldry, I. K., Bastian, S., Berlind, A., et al. 2004: *The Second Data Release of the Sloan Digital Sky Survey*, AJ, 128, 502

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Allende Prieto, C., An, D., Anderson, K. S. J., Anderson, S. F., Annis, J., Bahcall, N. A., & et al. 2009: *The Seventh Data Release of the Sloan Digital Sky Survey*, ApJS, 182, 543

Abdalla, F. B., Banerji, M., Lahav, O., & Rashkov, V. 2011: *A comparison of six photometric redshift methods applied to 1.5 million luminous red galaxies*, MNRAS, 417, 1891

Ahn, C. P., Alexandroff, R., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Aubourg, É., Bailey, S., Bastien, F. A., & et al. 2014: *The Tenth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Apache Point Observatory Galactic Evolution Experiment*, ApJS, 211, 17

Arnouts, S., Cristiani, S., Moscardini, L., Matarrese, S., Lucchin, F., Fontana, A., & Giallongo, E. 1999: *Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North*, MNRAS, 310, 540

Babbedge, T. S. R., Rowan-Robinson, M., Gonzalez-Solares, E., Polletta, M., Berta, S., Pérez-Fournon, I., Oliver, S., Salaman, D. M., Irwin, M., & Weatherley, S. J. 2004: *IMPZ: a new photometric redshift code for galaxies and quasars*, MNRAS, 353, 654

Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004: *Quantifying the Bimodal Color-Magnitude Distribution of Galaxies*, ApJ, 600, 681

Balestra, I., Vanzella, E., Rosati, P., Monna, A., Grillo, C., Nonino, M., Mercurio, A., Biviano, A., Bradley, L., Coe, D., Fritz, A., et al. 2013: *CLASH-VLT: spectroscopic confirmation of a z = 6.11 quintuply lensed galaxy in the Frontier Fields cluster RXC J2248.7-4431*, A&A, 559, L9

Barnes, J. E. & Hernquist, L. 1996: *Transformations of Galaxies. II. Gasdynamics in Merging Disk Galaxies*, ApJ, 471, 115

Baum, W. A. 1962: *Photoelectric Magnitudes and Red-Shifts*, in IAU Symposium, Vol. 15, Problems of Extra-Galactic Research, ed. G. C. McVittie, 390

Bell, E. F., Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., Kleinheinrich, M., Wisotzki, L., & McIntosh, D. H. 2004: *Nearly 5000 Distant Early-Type Galaxies in COMBO-17: A Red Sequence and Its Evolution since z˜1*, ApJ, 608, 752

Bender, R. 1988: *Velocity anisotropies and isophote shapes in elliptical galaxies*, A&A, 193, L7

Bender, R., Appenzeller, I., Böhm, A., Drory, N., Fricke, K. J., Gabasch, A., Heidt, J., Hopp, U., Jäger, K., Kümmel, M., Mehlert, D., et al. 2001: *The FORS Deep Field: Photometric Data and Photometric Redshifts*, in Deep Fields, ed. S. Cristiani, A. Renzini, & R. E. Williams, 96

Bender, R., Burstein, D., & Faber, S. M. 1993: *Dynamically hot galaxies. II - Global stellar populations*, ApJ, 411, 153

Bender, R., Doebereiner, S., & Moellenhoff, C. 1988: *Isophote shapes of elliptical galaxies. I - The data*, A&AS, 74, 385

Bender, R., Kormendy, J., Cornell, M. E., & Fisher, D. B. 2014: *The Cluster Velocity Dispersion of the Abell 2199 cD Halo of NGC 6166*, ArXiv e-prints

Bender, R. & Moellenhoff, C. 1987: *Morphological analysis of massive early-type galaxies in the Virgo cluster*, A&A, 177, 71

Benítez, N. 2000: *Bayesian Photometric Redshift Estimation*, ApJ, 536, 571

Benítez, N., Ford, H., Bouwens, R., Menanteau, F., Blakeslee, J., Gronwall, C., Illingworth, G., Meurer, G., Broadhurst, T. J., Clampin, M., Franx, M., et al. 2004: *Faint Galaxies in Deep Advanced Camera for Surveys Observations*, ApJS, 150, 1

Binney, J. & Merrifield, M. 1998, Galactic Astronomy, Princeton paperbacks (Princeton University Press)

Blanton, M. R., Dalcanton, J., Eisenstein, D., Loveday, J., Strauss, M. A., SubbaRao, M., Weinberg, D. H., Anderson, Jr., J. E., Annis, J., Bahcall, N. A., Bernardi, M., et al. 2001: *The Luminosity Function of Galaxies in SDSS Commissioning Data*, AJ, 121, 2358

Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000: *Photometric redshifts based on standard SED fitting procedures*, A&A, 363, 476

Bonnett, C. 2013: *Using neural networks to estimate redshift distributions. An application to CFHTLenS*, ArXiv e-prints

Boone, F., Clément, B., Richard, J., Schaerer, D., Lutz, D., Weiß, A., Zemcov, M., Egami, E., Rawle, T. D., Walth, G. L., Kneib, J.-P., et al. 2013: *An extended Herschel drop-out source in the center of AS1063: a normal dusty galaxy at z = 6.1 or SZ substructures?*, A&A, 559, L1

Bradley, L. D., Zitrin, A., Coe, D., Bouwens, R., Postman, M., Balestra, I., Grillo, C., Monna, A., Rosati, P., Seitz, S., Host, O., et al. 2014: *CLASH: A Census of Magnified Star-forming Galaxies at $z \sim 6-8$*, ApJ, 792, 76

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008: *EAZY: A Fast, Public Photometric Redshift Code*, ApJ, 686, 1503

Brimioulle, F., Lerchster, M., Seitz, S., Bender, R., & Snigula, J. 2008: *Photometric redshifts for the CFHTLS-Wide*, ArXiv e-prints

Brimioulle, F., Seitz, S., Lerchster, M., Bender, R., & Snigula, J. 2013: *Dark matter halo properties from galaxy-galaxy lensing*, MNRAS, 432, 1046

Brunner, R. J., Connolly, A. J., Szalay, A. S., & Bershady, M. A. 1997: *Toward More Precise Photometric Redshifts: Calibration Via CCD Photometry*, ApJ, 482, L21

Bruzual, G. & Charlot, S. 1993: *Spectral evolution of stellar populations using isochrone synthesis*, ApJ, 405, 538

Bruzual, G. & Charlot, S. 2003: *Stellar population synthesis at the resolution of 2003*, MNRAS, 344, 1000

Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., & Storchi-Bergmann, T. 2000: *The Dust Content and Opacity of Actively Star-forming Galaxies*, ApJ, 533, 682

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994: *Dust extinction of the stellar continua in starburst galaxies: The ultraviolet and optical extinction law*, ApJ, 429, 582

Caon, N., Capaccioli, M., & D'Onofrio, M. 1993: *On the Shape of the Light Profiles of Early Type Galaxies*, MNRAS, 265, 1013

Capaccioli, M., Piotto, G., & Rampazzo, R. 1988: *Photographic and CCD surface photometry of nine early-type galaxies*, AJ, 96, 487

Carliles, S., Budavári, T., Heinis, S., Priebe, C., & Szalay, A. S. 2010: *Random Forests for Photometric Redshifts*, ApJ, 712, 511

Carrasco Kind, M. & Brunner, R. J. 2013: *TPZ: photometric redshift PDFs and ancillary information by using prediction trees and random forests*, MNRAS, 432, 1483

Carrasco Kind, M. & Brunner, R. J. 2014a: *Exhausting the information: novel Bayesian combination of photometric redshift PDFs*, MNRAS, 442, 3380

Carrasco Kind, M. & Brunner, R. J. 2014b: *SOMz: photometric redshift PDFs with self-organizing maps and random atlas*, MNRAS, 438, 3409

Chabrier, G. 2003: *Galactic Stellar and Substellar Initial Mass Function*, PASP, 115, 763

Charlot, S. & Bruzual, A. G. 1991: *Stellar population synthesis revisited*, ApJ, 367, 126

Coe, D., Benítez, N., Sánchez, S. F., Jee, M., Bouwens, R., & Ford, H. 2006: *Galaxies in the Hubble Ultra Deep Field. I. Detection, Multiband Photometry, Photometric Redshifts, and Morphology*, AJ, 132, 926

Coleman, G. D., Wu, C., & Weedman, D. W. 1980: *Colors and magnitudes predicted for high redshift galaxies*, ApJS, 43, 393

Collister, A. A. & Lahav, O. 2004: *ANNz: Estimating Photometric Redshifts Using Artificial Neural Networks*, PASP, 116, 345

Connolly, A. J., Csabai, I., Szalay, A. S., Koo, D. C., Kron, R. G., & Munn, J. A. 1995: *Slicing Through Multicolor Space: Galaxy Redshifts from Broadband Photometry*, AJ, 110, 2655

Cooper, M. C., Yan, R., Dickinson, M., Juneau, S., Lotz, J. M., Newman, J. A., Papovich, C., Salim, S., Walth, G., Weiner, B. J., & Willmer, C. N. A. 2012: *The Arizona CDFS Environment Survey (ACES): A Magellan/IMACS Spectroscopic Survey of the Chandra Deep Field-South*, MNRAS, 425, 2116

Csabai, I., Budavári, T., Connolly, A. J., Szalay, A. S., Győry, Z., Benítez, N., Annis, J., Brinkmann, J., Eisenstein, D., Fukugita, M., Gunn, J., et al. 2003: *The Application of Photometric Redshifts to the SDSS Early Data Release*, AJ, 125, 580

Csabai, I., Connolly, A. J., Szalay, A. S., & Budavári, T. 2000: *Reconstructing Galaxy Spectral Energy Distributions from Broadband Photometry*, AJ, 119, 69

Csabai, I., Dobos, L., Trencséni, M., Herczegh, G., Józsa, P., Purger, N., Budavári, T., & Szalay, A. S. 2007: *Multidimensional indexing tools for the virtual observatory*, Astronomische Nachrichten, 328, 852

Davies, R. L., Sadler, E. M., & Peletier, R. F. 1993: *Line-strength gradients in elliptical galaxies*, MNRAS, 262, 650

Dawson, K. S., Schlegel, D. J., Ahn, C. P., Anderson, S. F., Aubourg, É., Bailey, S., Barkhouser, R. H., Bautista, J. E., Beifiori, A., Berlind, A. A., Bhardwaj, V., et al. 2013: *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, AJ, 145, 10

de Vaucouleurs, G. 1948: *Recherches sur les Nebuleuses Extragalactiques*, Annales d'Astrophysique, 11, 247

de Vaucouleurs, G. 1953: *On the distribution of mass and luminosity in elliptical galaxies*, MNRAS, 113, 134

Díaz García, L. A., Cenarro Lagunas, J., & et al. submitted 2014: *MUFFIT*, MNRAS

Diehl, H. T. & Dark Energy Survey Collaboration. 2012: *The Dark Energy Survey & Camera (DECam)*, in American Astronomical Society Meeting Abstracts, Vol. 219, American Astronomical Society Meeting Abstracts #219, #413.05

Djorgovski, S. & Davis, M. 1987: *Fundamental properties of elliptical galaxies*, ApJ, 313, 59

Dressler, A., Lynden-Bell, D., Burstein, D., Davies, R. L., Faber, S. M., Terlevich, R., & Wegner, G. 1987: *Spectroscopy and photometry of elliptical galaxies. I - A new distance estimator*, ApJ, 313, 42

Drory, N., Bender, R., & Hopp, U. 2004: *Comparing Spectroscopic and Photometric Stellar Mass Estimates*, ApJ, 616, L103

Drory, N., Feulner, G., Bender, R., Botzler, C. S., Hopp, U., Maraston, C., Mendes de Oliveira, C., & Snigula, J. 2001: *The Munich Near-Infrared Cluster Survey - I. Field selection, object extraction and photometry*, MNRAS, 325, 550

Drory, N., Salvato, M., Gabasch, A., Bender, R., Hopp, U., Feulner, G., & Pannella, M. 2005: *The Stellar Mass Function of Galaxies to z˜5 in the FORS Deep and GOODS-South Fields*, ApJ, 619, L131

Dunlop, J. S. 2013: *Observing the First Galaxies*, in Astrophysics and Space Science Library, Vol. 396, Astrophysics and Space Science Library, ed. T. Wiklind, B. Mobasher, & V. Bromm, 223

Einstein, A. 1916: *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik, 354, 769

Eisenstein, D. J., Annis, J., Gunn, J. E., Szalay, A. S., Connolly, A. J., Nichol, R. C., Bahcall, N. A., Bernardi, M., Burles, S., Castander, F. J., Fukugita, M., et al. 2001: *Spectroscopic Target Selection for the Sloan Digital Sky Survey: The Luminous Red Galaxy Sample*, AJ, 122, 2267

Eisenstein, D. J., Weinberg, D. H., Agol, E., Aihara, H., Allende Prieto, C., Anderson, S. F., Arns, J. A., Aubourg, É., Bailey, S., Balbinot, E., & et al. 2011: *SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems*, AJ, 142, 72

Faber, S. M. & Jackson, R. E. 1976: *Velocity dispersions and mass-to-light ratios for elliptical galaxies*, ApJ, 204, 668

Faber, S. M., Worthey, G., & Gonzales, J. J. 1992: *Absorption-Line Spectra of Elliptical Galaxies and Their Relation to Elliptical Formation*, in IAU Symposium, Vol. 149, The Stellar Populations of Galaxies, ed. B. Barbuy & A. Renzini, 255

Fan, X., Carilli, C. L., & Keating, B. 2006: *Observational Constraints on Cosmic Reionization*, ARA&A, 44, 415

Feldmann, R., Carollo, C. M., Porciani, C., Lilly, S. J., Capak, P., Taniguchi, Y., Le Fèvre, O., Renzini, A., Scoville, N., Ajiki, M., Aussel, H., et al. 2006: *The Zurich Extragalactic Bayesian Redshift Analyzer and its first application: COSMOS*, MNRAS, 372, 565

Feulner, G., Gabasch, A., Salvato, M., Drory, N., Hopp, U., & Bender, R. 2005: *Specific Star Formation Rates to Redshift 5 from the FORS Deep Field and the GOODS-S Field*, ApJ, 633, L9

Flaugher, B. 2005: *The Dark Energy Survey*, International Journal of Modern Physics A, 20, 3121

Flaugher, B. L., Abbott, T. M. C., Angstadt, R., Annis, J., Antonik, M. L., Bailey, J., Ballester, O., Bernstein, J. P., Bernstein, R. A., Bonati, M., Bremer, G., et al. 2012: *Status of the Dark Energy Survey Camera (DECam) project*, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8446, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Franx, M., Illingworth, G., & Heckman, T. 1989: *Multicolor surface photometry of 17 ellipticals*, AJ, 98, 538

Franx, M., van Gorkom, J. H., & de Zeeuw, T. 1994: *Evidence for axisymmetric halos: The case of IC 2006*, ApJ, 436, 642

Freeman, P. E., Newman, J. A., Lee, A. B., Richards, J. W., & Schafer, C. M. 2009: *Photometric redshift estimation using spectral connectivity analysis*, MNRAS, 398, 2012

Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996: *The Sloan Digital Sky Survey Photometric System*, AJ, 111, 1748

Gabasch, A., Bender, R., Seitz, S., Hopp, U., Saglia, R. P., Feulner, G., Snigula, J., Drory, N., Appenzeller, I., Heidt, J., Mehlert, D., et al. 2004: *The evolution of the luminosity functions in the FORS Deep Field from low to high redshift. I. The blue bands*, A&A, 421, 41

Gabasch, A., Goranova, Y., Hopp, U., Noll, S., & Pannella, M. 2008: *A deep i-selected multiwaveband galaxy catalogue in the COSMOS field*, MNRAS, 383, 1319

Gerdes, D. W. 2009: *Photometric Redshifts Using Boosted Decision Trees*, in Bulletin of the American Astronomical Society, Vol. 41, American Astronomical Society Meeting Abstracts #213, #483.03

Greisel, N., Seitz, S., Drory, N., Bender, R., Saglia, R. P., & Snigula, J. 2013: *Photometric Redshifts and Systematic Variations in the Spectral Energy Distributions of Luminous Red Galaxies from SDSS DR7*, ApJ, 768, 117

Gruen, D., Brimioulle, F., Seitz, S., Lee, C.-H., Young, J., Koppenhoefer, J., Eichner, T., Riffeser, A., Vikram, V., Weidinger, T., & Zenteno, A. 2013: *Weak lensing analysis of RXC J2248.7-4431*, MNRAS, 432, 1455

Gruen, D., Seitz, S., Brimioulle, F., Kosyra, R., Koppenhoefer, J., Lee, C.-H., Bender, R., Riffeser, A., Eichner, T., Weidinger, T., & Bierschenk, M. 2014: *Weak lensing analysis of SZ-selected clusters of galaxies from the SPT and Planck surveys*, MNRAS, 442, 1507

Gunn, J. E., Carr, M., Rockosi, C., Sekiguchi, M., Berry, K., Elms, B., de Haas, E., Ivezić, Ž., Knapp, G., Lupton, R., Pauls, G., et al. 1998: *The Sloan Digital Sky Survey Photometric Camera*, AJ, 116, 3040

Gunn, J. E. & Peterson, B. A. 1965: *On the Density of Neutral Hydrogen in Intergalactic Space.*, ApJ, 142, 1633

Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., Kent, S. M., et al. 2006: *The 2.5 m Telescope of the Sloan Digital Sky Survey*, AJ, 131, 2332

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

Guzzo, L., Schuecker, P., Böhringer, H., Collins, C. A., Ortiz-Gil, A., de Grandi, S., Edge, A. C., Neumann, D. M., Schindler, S., Altucci, C., & Shaver, P. A. 2009: *The REFLEX galaxy cluster survey. VIII. Spectroscopic observations and optical atlas,*, A&A, 499, 357

Hanke, M., Halchenko, Y. O., Sederberg, P. B., Hanson, S. J., Haxby, J. V., & Pollmann, S. 2009: *PyMVPA: a Python Toolbox for Multivariate Pattern Analysis of fMRI Data*, Neuroinformatics, 7, 37

Hildebrandt, H., Arnouts, S., Capak, P., Moustakas, L. A., Wolf, C., Abdalla, F. B., Assef, R. J., Banerji, M., Benítez, N., Brammer, G. B., Budavári, T., et al. 2010: *PHAT: PHoto-z Accuracy Testing*, A&A, 523, A31

Hildebrandt, H., Erben, T., Kuijken, K., van Waerbeke, L., Heymans, C., Coupon, J., Benjamin, J., Bonnett, C., Fu, L., Hoekstra, H., Kitching, T. D., et al. 2012: *CFHTLenS: improving the quality of photometric redshifts with precision photometry*, MNRAS, 421, 2355

Hubble, E. P. 1926: *Extragalactic nebulae.*, ApJ, 64, 321

Ilbert, O., Arnouts, S., McCracken, H. J., Bolzonella, M., Bertin, E., Le Fèvre, O., Mellier, Y., Zamorani, G., Pellò, R., Iovino, A., Tresse, L., et al. 2006: *Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey*, A&A, 457, 841

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H. J., Sanders, D. B., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floc'h, E., Mobasher, B., et al. 2009: *Cosmos Photometric Redshifts with 30-Bands for* $2-\deg^2$, ApJ, 690, 1236

Johnson, H. L. & Morgan, W. W. 1953: *Fundamental stellar photometry for standards of spectral type on the revised system of the Yerkes spectral atlas*, ApJ, 117, 313

Kaiser, N., Burgett, W., Chambers, K., Denneau, L., Heasley, J., Jedicke, R., Magnier, E., Morgan, J., Onaka, P., & Tonry, J. 2010: *The Pan-STARRS wide-field optical/NIR imaging survey*, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7733, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Kaiser, N., Tonry, J. L., & Luppino, G. A. 2000: *A New Strategy for Deep Wide-Field High-Resolution Optical Imaging*, PASP, 112, 768

Karman, W., Caputi, K. I., Grillo, C., Balestra, I., Rosati, P., Vanzella, E., Coe, D., Christensen, L., Koekemoer, A. M., Kruehler, T., Lombardi, M., et al. 2014: *MUSE integral-field spectroscopy towards the Frontier Fields Cluster Abell S1063: I. Data products and redshift identifications*, ArXiv e-prints

Khochfar, S. & Burkert, A. 2003: *The Importance of Spheroidal and Mixed Mergers for Early-Type Galaxy Formation*, ApJ, 597, L117

Kohonen, T. 1982: *Self-organized formation of topologically correct feature maps*, Biological Cybernetics, 43, 59

Kohonen, T. 2001, Self-Organizing Maps

Koo, D. C. 1999: *Overview - Photometric Redshifts: A Perspective from an Old-Timer[!] on their Past, Present, and Potential*, in Astronomical Society of the Pacific Conference Series, Vol. 191, Photometric Redshifts and the Detection of High Redshift Galaxies, ed. R. Weymann, L. Storrie-Lombardi, M. Sawicki, & R. Brunner, 3

Kormendy, J. & Djorgovski, S. 1989: *Surface photometry and the structure of elliptical galaxies*, ARA&A, 27, 235

Kormendy, J., Fisher, D. B., Cornell, M. E., & Bender, R. 2009: *Structure and Formation of Elliptical and Spheroidal Galaxies*, ApJS, 182, 216

Kriek, M., van Dokkum, P. G., Franx, M., Förster Schreiber, N. M., Gawiser, E., Illingworth, G. D., Labbé, I., Marchesini, D., Quadri, R., Rix, H.-W., Rudnick, G., et al. 2006: *Direct Measurements of the Stellar Continua and Balmer/4000 Å Breaks of Red $z > 2$ Galaxies: Redshifts and Improved Constraints on Stellar Populations1,*, ApJ, 645, 44

Kroupa, P. 2001: *On the variation of the initial mass function*, MNRAS, 322, 231

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. ., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., & et al. 2011: *Euclid Definition Study Report*, ArXiv e-prints

Le Borgne, D. & Rocca-Volmerange, B. 2002: *Photometric redshifts from evolutionary synthesis with PÉGASE: The code Z-PEG and the z=0 age constraint*, A&A, 386, 446

Le Fèvre, O., Vettolani, G., Garilli, B., Tresse, L., Bottini, D., Le Brun, V., Maccagni, D., Picat, J. P., Scaramella, R., Scodeggio, M., Zanichelli, A., et al. 2005: *The VIMOS VLT deep survey. First epoch VVDS-deep survey: 11,564 spectra with $17.5 \leq I_A B \leq 24$, and the redshift distribution over $0 \leq z \leq 5$*, A&A, 439, 845

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013: *Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos*, ApJ, 772, 119

Lilly, S. J., Le Fèvre, O., Renzini, A., Zamorani, G., Scodeggio, M., Contini, T., Carollo, C. M., Hasinger, G., Kneib, J.-P., Iovino, A., Le Brun, V., et al. 2007: *zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering $0 < z < 3$ in the COSMOS Field*, ApJS, 172, 70

Lima, M., Cunha, C. E., Oyaizu, H., Frieman, J., Lin, H., & Sheldon, E. S. 2008: *Estimating the redshift distribution of photometric galaxy samples*, MNRAS, 390, 118

Loh, E. D. & Spillar, E. J. 1986: *Photometric redshifts of galaxies*, ApJ, 303, 154

Longhetti, M. & Saracco, P. 2009: *Stellar mass estimates in early-type galaxies: procedures, uncertainties and models dependence*, MNRAS, 394, 774

Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999: *A Modified Magnitude System that Produces Well-Behaved Magnitudes, Colors, and Errors Even for Low Signal-to-Noise Ratio Measurements*, AJ, 118, 1406

Madau, P. 1995: *Radiative transfer in a clumpy universe: The colors of high-redshift galaxies*, ApJ, 441, 18

Malumuth, E. M. & Kirshner, R. P. 1981: *Dynamics of luminous galaxies*, ApJ, 251, 508

Mancone, C. L. & Gonzalez, A. H. 2012: *EzGal: A Flexible Interface for Stellar Population Synthesis Models*, PASP, 124, 606

Mandelbaum, R., Seljak, U., Hirata, C. M., Bardelli, S., Bolzonella, M., Bongiorno, A., Carollo, M., Contini, T., Cunha, C. E., Garilli, B., Iovino, A., et al. 2008: *Precision photometric redshift calibration for galaxy-galaxy weak lensing*, MNRAS, 386, 781

Mannucci, F., Basile, F., Poggianti, B. M., Cimatti, A., Daddi, E., Pozzetti, L., & Vanzi, L. 2001: *Near-infrared template spectra of normal galaxies: k-corrections, galaxy models and stellar populations*, MNRAS, 326, 745

Maraston, C. 1998: *Evolutionary synthesis of stellar populations: a modular tool*, MNRAS, 300, 872

Maraston, C. 2005: *Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies*, MNRAS, 362, 799

Maraston, C., Pforr, J., Renzini, A., Daddi, E., Dickinson, M., Cimatti, A., & Tonini, C. 2010: *Star formation rates and masses of $z \sim 2$ galaxies from multicolour photometry*, MNRAS, 407, 830

Maraston, C. & Strömbäck, G. 2011: *Stellar population models at high spectral resolution*, MNRAS, 418, 2785

Maraston, C., Strömbäck, G., Thomas, D., Wake, D. A., & Nichol, R. C. 2009: *Modelling the colour evolution of luminous red galaxies - improvements with empirical stellar spectra*, MNRAS, 394, L107

Martin, D. C., Fanson, J., Schiminovich, D., Morrissey, P., Friedman, P. G., Barlow, T. A., Conrow, T., Grange, R., Jelinsky, P. N., Milliard, B., Siegmund, O. H. W., et al. 2005: *The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission*, ApJ, 619, L1

Moles, M., Aguerri, J. A. L., Alfaro, E. J., Benítez, N., Broadhurst, T., Cabrera-Caño, J., Castander, F. J., Cepa, J., Cerviño, M., Fernández Soto, A., González Delgado, R. M., et al. 2008: *The ALHAMBRA Survey*, in Astronomical Society of the Pacific Conference Series, Vol. 390, Pathways Through an Eclectic Universe, ed. J. H. Knapen, T. J. Mahoney, & A. Vazdekis, 495

Molino, A., Benítez, N., Moles, M., Fernández-Soto, A., Cristóbal-Hornillos, D., Ascaso, B., Jiménez-Teja, Y., Schoenell, W., Arnalte-Mur, P., Pović, M., Coe, D., et al. 2014: *The ALHAMBRA Survey: Bayesian photometric redshifts with 23 bands for 3 deg*, MNRAS, 441, 2891

Monna, A., Seitz, S., Greisel, N., Eichner, T., Drory, N., Postman, M., Zitrin, A., Coe, D., Halkola, A., Suyu, S. H., Grillo, C., et al. 2014: *CLASH: $z \sim 6$ young galaxy candidate quintuply lensed by the frontier field cluster RXC J2248.7-4431*, MNRAS, 438, 1417

Morgan, W. W. 1958: *A Preliminary Classification of the Forms of Galaxies According to Their Stellar Population*, PASP, 70, 364

Mukhanov, V. 2005, Physical Foundations of Cosmology

Naab, T. & Burkert, A. 2003: *Statistical Properties of Collisionless Equal- and Unequal-Mass Merger Remnants of Disk Galaxies*, ApJ, 597, 893

Oyaizu, H., Lima, M., Cunha, C. E., Lin, H., Frieman, J., & Sheldon, E. S. 2008: *A Galaxy Photometric Redshift Catalog for the Sloan Digital Sky Survey Data Release 6*, ApJ, 674, 768

Padmanabhan, N., Budavári, T., Schlegel, D. J., Bridges, T., Brinkmann, J., Cannon, R., Connolly, A. J., Croom, S. M., Csabai, I., Drinkwater, M., Eisenstein, D. J., et al. 2005: *Calibrating photometric redshifts of luminous red galaxies*, MNRAS, 359, 237

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., et al. 2011: *Scikit-learn: Machine Learning in Python*, Journal of Machine Learning Research, 12, 2825

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., et al. 1999: *Measurements of* Ω *and* Λ *from 42 High-Redshift Supernovae*, ApJ, 517, 565

Petrosian, V. 1976: *Surface brightness and evolution of galaxies*, ApJ, 209, L1

Pforr, J., Maraston, C., & Tonini, C. 2012: *Recovering galaxy stellar population properties from broad-band spectral energy distribution fitting*, MNRAS, 422, 3285

Pickles, A. J. 1998: *A Stellar Spectral Flux Library: 1150-25000 Å*, PASP, 110, 863

Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., & et al. 2014: *Planck 2013 results. XVI. Cosmological parameters*, A&A, 571, A16

Pogson, N. 1856: *Magnitudes of Thirty-six of the Minor Planets for the first day of each month of the year 1857*, MNRAS, 17, 12

Postman, M., Coe, D., Benítez, N., Bradley, L., Broadhurst, T., Donahue, M., Ford, H., Graur, O., Graves, G., Jouvel, S., Koekemoer, A., et al. 2012: *The Cluster Lensing and Supernova Survey with Hubble: An Overview*, ApJS, 199, 25

Puschell, J., Owen, F., & Laing, R. 1982: *Redshift Estimates for Distant Radio Galaxies Based on Broadband Photometry*, in IAU Symposium, Vol. 97, Extragalactic Radio Sources, ed. D. S. Heeschen & C. M. Wade, 423

Rau, M. M., Seitz, S., & et al. in prep.

Ree, C. H., Jeong, H., Oh, K., Chung, C., Lee, J. H., Kim, S. C., & Kyeong, J. 2012: *Ultraviolet Color-Color Relation of Early-type Galaxies at* $0.05 < z < 0.12$, ApJ, 744, L10

Richard, J., Kneib, J.-P., Ebeling, H., Stark, D. P., Egami, E., & Fiedler, A. K. 2011: *Discovery of a possibly old galaxy at z= 6.027, multiply imaged by the massive cluster Abell 383*, MNRAS, 414, L31

Richards, J. W., Freeman, P. E., Lee, A. B., & Schafer, C. M. 2009: *Exploiting Low-Dimensional Structure in Astronomical Spectra*, ApJ, 691, 32

Roberts, M. S. & Haynes, M. P. 1994: *Physical Parameters along the Hubble Sequence*, ARA&A, 32, 115

Saglia, R. P., Snigula, J., Senger, R., & Bender, R. 2013: *Implementation of PhotoZ under Astro-WISE. A photometric redshift code for large datasets*, Experimental Astronomy, 35, 337

Saglia, R. P., Tonry, J. L., Bender, R., Greisel, N., Seitz, S., Senger, R., Snigula, J., Phleps, S., Wilman, D., Bailer-Jones, C. A. L., Klement, R. J., et al. 2012: *The Photometric Classification Server for Pan-STARRS1*, ApJ, 746, 128

Salpeter, E. E. 1955: *The Luminosity Function and Stellar Evolution.*, ApJ, 121, 161

Sánchez, C., Carrasco Kind, M., Lin, H., Miquel, R., Abdalla, F. B., Amara, A., Banerji, M., Bonnett, C., Brunner, R., Capozzi, D., Carnero, A., et al. 2014: *Photometric redshift analysis in the Dark Energy Survey Science Verification data*, MNRAS, 445, 1482

Sánchez-Blázquez, P., Peletier, R. F., Jiménez-Vicente, J., Cardiel, N., Cenarro, A. J., Falcón-Barroso, J., Gorgas, J., Selam, S., & Vazdekis, A. 2006: *Medium-resolution Isaac Newton Telescope library of empirical spectra*, MNRAS, 371, 703

Sánchez-Blázquez, P., Rosales-Ortega, F. F., Méndez-Abreu, J., Pérez, I., Sánchez, S. F., Zibetti, S., Aguerri, J. A. L., Bland-Hawthorn, J., Catalán-Torrecilla, C., Cid Fernandes, R., de Amorim, A., et al. 2014: *Stellar population gradients in galaxy discs from the CALIFA survey. The influence of bars*, A&A, 570, A6

Schawinski, K., Kaviraj, S., Khochfar, S., Yoon, S.-J., Yi, S. K., Deharveng, J.-M., Boselli, A., Barlow, T., Conrow, T., Forster, K., Friedman, P. G., et al. 2007: *The Effect of Environment on the Ultraviolet Color-Magnitude Relation of Early-Type Galaxies*, ApJS, 173, 512

Schechter, P. L. 1980: *Mass-to-light ratios for elliptical galaxies*, AJ, 85, 801

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998: *Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds*, ApJ, 500, 525

Schneider, P., Kochanek, C. S., & Wambsganss, J. 2006: *Gravitational Lensing: Strong, Weak and Micro*, in Saas-Fee Advanced Course 33: Gravitational Lensing: Strong, Weak and Micro

Schombert, J. M. 1986: *The structure of brightest cluster members. I - Surface photometry*, ApJS, 60, 603

Schweizer, F. 1987: *Dust and gas - Overview*, in IAU Symposium, Vol. 127, Structure and Dynamics of Elliptical Galaxies, ed. P. T. de Zeeuw, 109–122

Sérsic, J. L. 1963: *Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy*, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sérsic, J. L. 1968, Atlas de galaxias australes

Shapley, A. E., Steidel, C. C., Erb, D. K., Reddy, N. A., Adelberger, K. L., Pettini, M., Barmby, P., & Huang, J. 2005: *Ultraviolet to Mid-Infrared Observations of Star-forming Galaxies at z˜2: Stellar Masses and Stellar Populations*, ApJ, 626, 698

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998: *Modeling the Effects of Dust on Galactic Spectral Energy Distributions from the Ultraviolet to the Millimeter Band*, ApJ, 509, 103

Smee, S. A., Gunn, J. E., Uomoto, A., Roe, N., Schlegel, D., Rockosi, C. M., Carr, M. A., Leger, F., Dawson, K. S., Olmstead, M. D., Brinkmann, J., et al. 2013: *The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey*, AJ, 146, 32

Smith, J. A., Tucker, D. L., Kent, S., Richmond, M. W., Fukugita, M., Ichikawa, T., Ichikawa, S.-i., Jorgensen, A. M., Uomoto, A., Gunn, J. E., Hamabe, M., et al. 2002: *The u'g'r'i'z' Standard-Star System*, AJ, 123, 2121

Steidel, C. C. 1996: *Normal galaxies at extreme redshifts.*, in Bulletin of the American Astronomical Society, Vol. 28, Bulletin of the American Astronomical Society, 1312

Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996a: *Spectroscopy of Lyman Break Galaxies in the Hubble Deep Field*, AJ, 112, 352

Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., & Adelberger, K. L. 1996b: *Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts $z > 3$*, ApJ, 462, L17

Steinhaus, H. 1957: *Sur la division des corps matériels en parties.*, Bull. Acad. Pol. Sci., Cl. III, 4, 801

Stoughton, C., Lupton, R. H., Bernardi, M., Blanton, M. R., Burles, S., Castander, F. J., Connolly, A. J., Eisenstein, D. J., Frieman, J. A., Hennessy, G. S., Hindsley, R. B., et al. 2002: *Sloan Digital Sky Survey: Early Data Release*, AJ, 123, 485

Tamura, N., Kobayashi, C., Arimoto, N., Kodama, T., & Ohta, K. 2000: *Origin of Color Gradients in Elliptical Galaxies*, AJ, 119, 2134

Tinsley, B. M. 1980: *Evolution of the Stars and Gas in Galaxies*, 5, 287

Tonry, J. L. 1981: *Velocity dispersions of low luminosity ellipticals - L approximately equal to sigma-cubed*, ApJ, 251, L1

Toomre, A. & Toomre, J. 1972: *Galactic Bridges and Tails*, ApJ, 178, 623

Trinchieri, G. & Fabbiano, G. 1985: *A Statistical Analysis of the Einstein Normal Galaxy Sample - Part Two - Elliptical and s0 Galaxies*, ApJ, 296, 447

Tyson, J. A., Wittman, D. M., Hennawi, J. F., & Spergel, D. N. 2003: *LSST: a complementary probe of dark energy*, Nuclear Physics B Proceedings Supplements, 124, 21

van Gorkom, J. 1992: *Cold Gas in Elliptical Galaxies*, in Astrophysics and Space Science Library, Vol. 178, Morphological and Physical Classification of Galaxies, ed. G. Longo, M. Capaccioli, & G. Busarello, 233

Voronoi, G. 1908: *Nouvelles applications des paramtres continus á la théorie des formes quadratiques*, Journal für die Reine und Angewandte Mathematik, 133, 97178

Wadadekar, Y. 2005: *Estimating Photometric Redshifts Using Support Vector Machines*, PASP, 117, 79

Wang, Y., Bahcall, N., & Turner, E. L. 1998: *A Catalog of Color-based Redshift Estimates for $z \lesssim 4$ Galaxies in the Hubble Deep Field*, AJ, 116, 2081

Way, M. J., Foster, L. V., Gazis, P. R., & Srivastava, A. N. 2009: *New Approaches to Photometric Redshift Prediction Via Gaussian Process Regression in the Sloan Digital Sky Survey*, ApJ, 706, 623

Weidner, C., Kroupa, P., Pflamm-Altenburg, J., & Vazdekis, A. 2013: *The galaxy-wide initial mass function of dwarf late-type to massive early-type galaxies*, MNRAS, 436, 3309

Worthey, G., Faber, S. M., & Gonzalez, J. J. 1992: *MG and Fe absorption features in elliptical galaxies*, ApJ, 398, 69

Yasuda, N., Fukugita, M., Narayanan, V. K., Lupton, R. H., Strateva, I., Strauss, M. A., Ivezić, Ž., Kim, R. S. J., Hogg, D. W., Weinberg, D. H., Shimasaku, K., et al. 2001: *Galaxy Number Counts from the Sloan Digital Sky Survey Commissioning Data*, AJ, 122, 1104

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., et al. 2000: *The Sloan Digital Sky Survey: Technical Summary*, AJ, 120, 1579

Zheng, H. & Zhang, Y. 2012: *Support vector machines for photometric redshift measurement of quasars*, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8451, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Zitrin, A., Moustakas, J., Bradley, L., Coe, D., Moustakas, L. A., Postman, M., Shu, X., Zheng, W., Benítez, N., Bouwens, R., Broadhurst, T., et al. 2012: *CLASH: Discovery of a Bright z ˜= 6.2 Dwarf Galaxy Quadruply Lensed by MACS J0329.6-0211*, ApJ, 747, L9

Appendix A

Photometric Redshift Quality Metrics

We summarize the metrics we use in Chapters 5, 6, and 7 for the quality assessment of the photometric redshift estimations. In the text as in the following Δz denotes the real deviation of the photometric redshift from the spectroscopic redshift, $\Delta z = z_{\rm phot} - z_{\rm spec}$, and Δz_{rf} is the error in the rest frame $\Delta z_{rf} = \Delta z/(1+z_{\rm spec})$. The estimated error from any photo-z code is termed δz.

Catastrophic Outlier Rate

The percental fraction η of so-called *catastrophic outliers* is defined by their absolute photometric redshift error in the rest frame, $|\Delta z_{rf}| \geq 0.15$. While the value of 0.15 is somewhat arbitrary, we chose to adopt it in consistency with Ilbert et al. (2006).

Bias

The bias is the mean deviation of $z_{\rm phot}$ from $z_{\rm spec}$,

$$\langle \Delta z \rangle = \frac{1}{N} \sum_{i=1}^{N} \Delta z_i, \text{ or}$$
$$\langle \Delta z_{rf} \rangle = \frac{1}{N} \sum_{i=1}^{N} \Delta z_{rf,\,i}$$

in the rest frame respectively, where i runs over the total number of objects N in the investigated sample.

Mean Errors

The calculation of the mean errors is as for the bias above, but with $|\Delta z_i|$ instead of Δz_i, and $|\Delta z_{rf,\,i}|$ instead of $\Delta z_{rf,\,i}$ respectively.

Square Root of the Sample Variance

The square roots of the sample variance of Δz, and Δz_{rf} in the rest frame are:

$$\sigma(\Delta z) = \left[\frac{1}{N-1}\sum_{i=1}^{N}\left(\Delta z_i - \langle\Delta z\rangle\right)^2\right]^{1/2} \quad \text{or}$$

$$\sigma(\Delta z_{rf}) = \left[\frac{1}{N-1}\sum_{i=1}^{N}\left(\Delta z_{rf,\,i} - \langle\Delta z_{rf}\rangle\right)^2\right]^{1/2}$$

Normalized Median Absolute Deviation

Similar to the root mean squared error, the normalized median absolute deviation (NMAD) for non-outliers (photo-z estimations that do not fulfill the outlier criterion above) gives clues about the Δz distribution, but without taking the flanks (the outliers) of the distribution into account. It is calculated as

$$\sigma_{\mathrm{NMAD}} = 1.48 \cdot \mathrm{median}\left(\Delta z_{rf}\right)_{\mathrm{non-outliers}}.$$

Gaussian Distribution of Δz

To analyze how well the distribution of Δz resembles a Gaussian distribution, we define,

$$P_x \equiv P(\Delta z \leq \Delta z_x) \equiv \frac{1}{N}\sum_{\Delta z}\mathbb{1}_{[-\infty;\Delta z_x]}(\Delta z),$$

where x can hold any percental value of the distribution. $\mathbb{1}_X(x)$ is the indicator function which equals 1 for $x \in X$, and zero otherwise. Let Δz_{50} denote the median of Δz, then $P_{50} = \frac{1}{2}$. We introduce the half of the interval in Δz where 68.27 % of the galaxies of a sample are enclosed, corresponding to a 1σ confidence, by

$$\sigma_{68} = \frac{1}{2}(\Delta z_{86} - \Delta z_{16}).$$

Then, we can check how similar the distribution of Δz is to a Gaussian function by calculating the fractions η_{95} and η_{99}, where $|\Delta z| \leq 2\sigma_{68}$ and $|\Delta z| \leq 3\sigma_{68}$ respectively, and which should read 95.45 % and 99.73 % to coincide with 2σ and 3σ intervals.

Kolmogorov-Smirnov Test

In order to quantify the similarity of the distribution $N(z_{\mathrm{phot}})$ to $N(z_{\mathrm{spec}})$, from which z_{phot} and z_{spec} are drawn, we adopt the Kolmogorov-Smirnov (KS) test. It defines a value that measures the maximum difference between both cumulative distributions:

$$\mathrm{KS} = \max_z\left(\left|F_{\text{photo-}z}(z) - F_{\text{spec-}z}(z)\right|\right),$$

where z runs over the array of both z_{spec} and z_{phot} values, and

$$F_*(z) = \sum_{z_*}\mathbb{1}_{[0;z[}(z_*),$$

with $*$ either being spec-z or photo-z. The KS test can also be performed on the PDFs $P(z)$ of the galaxies. $F_{\text{photo-}z}(z)$ is then the cumulative probability distribution of the stacked PDFs.

Quality of Redshift Error Estimation

Most photo-z codes also estimate the error of their photometric redshift estimate, mostly from an underlying PDF. To measure the quality of the estimated error we introduce a normalized error, $\Delta z' \equiv \Delta z/\delta z$. The distribution of $\Delta z'$ should be nearing a standard normal distribution, if the estimated δz values are correct. Therefore, the mean $\langle \Delta z' \rangle$ and the square root of the sample variance

$$\sigma(\Delta z') = \left[\frac{1}{N-1} \sum_{i=1}^{N} \left(\Delta z' - \langle \Delta z' \rangle \right)^2 \right]^{1/2} ,$$

where N is the number of galaxies, give a measure of the δz accuracy.

Appendix B

SED Fitting the SDSS-II LRG Data

B.1 SED Fitting Results with Different Model Sets

In Section 5 we created templates by fitting model SEDs to the photometry of the SDSS LRGs in order to produce SEDs matching the colors of the data. Here we want to show the fitting results and how well the models are able to cover the observed colors. In Figures 4.9 and 4.11 we saw that both the BC03 and the M09 model SEDs can represent the LRG data, at least for a large part of the investigated redshift range. Both sets have a large variety of ages (plus e-folding timescales and metallicities in the case of BC03 models), thus they offer the perspective to provide a well matching model set for each redshift.
In addition to that, Maraston & Strömbäck (2011, M11) provide a large variety of single stellar populations (SSPs) created from different stellar libraries and with varying metallicities and ages. The M11 models are promising candidates to represent the LRG data in terms of colors, wherefore we analyzed them in this regard and compared to the results with BC03 models. We could only use those M11 model SEDs that are provided for a wavelength range sufficiently large to cover all SDSS filters at all LRG redshifts. These are two flux-calibrated empirical stellar spectral libraries, Pickles (Pickles, 1998) and MILES (Sánchez-Blázquez et al., 2006), and the theoretically derived MARCS (Gustafsson et al., 2008) library. The Pickles models have solar metallicity, either an IMF of Chabrier (2003) or Salpeter, and were extended in the near-IR by M11. A version with a theoretical UV part was additionally created by M11. The MILES SSPs in question were extended in the near-IR and UV, have a revised IR slope, and solar metallicity. The MARCS models have metallicities of $Z = 0.0004, 0.02$, and 0.04, and IMFs of Salpeter and Kroupa (2001). A more thorough description of the SSPs as well as the SEDs can be found in M11 and at `http://www.icg.port.ac.uk/~maraston/M11/`. These models are a subset of all M11 covering a wavelength range wide enough for photometric redshift estimation.
We produced CSPs from the selected M11 models with the software `EzGal`[1] (Mancone & Gonzalez, 2012). We apply exponential SFRs and the same e-folding timescales that were used with BC03 (Section 4.2.2). The whole range of available ages, IMFs, and metallicities were exploited where possible in order to cover the greatest variety within the resulting SEDs.
These CSPs were then fitted to the photometry in the same way as the BC03 CSPs in Section 5 (i.e., superposed by a burst model and with additional extinction) for each stellar

[1] `http://www.baryons.org/ezgal/`

library and IMF separately. The same was also done with M09 models.

The SED fitting results for the different redshift bins and color spaces of M09/11 are shown together with those of the BC03 models in Figures B.1 to B.3. From the M09/11 models in all cases the one with the smallest χ^2 result is chosen for the plot. The upper panel of Figure B.1 ($z \approx 0.02$) shows that the BC03 models cover the LRG colors better in this z range, with the exception of *riiz*, where M09/M11 account for the lower right part that is not covered by BC03. Furthermore, the BC03 color contours trace the LRG distribution better than M09/11 in *grri* and *griz*. In the upper right and middle left panel the colors are shown for the redshift bins at $z \approx 0.1$, and $z \approx 0.2$. Both model sets perform similarly well, except for $z \approx 0.2$, where M09/11 produces a slightly wider spread in *ugri*. Apart from that, the BC03 contours follow well the density of the LRGs in the color spaces *grri* and *griz*. At $z \approx 0.3$ the M09/11 models outperform BC03 in *riiz*, whereas both describe the data colors similarly in the other spaces (middle right panel). Here, the M09/11 SEDs produce two peaks in *riiz* which cannot be observed for the LRGs. Finally, Figure B.3 shows the redshift bin around 0.4, where M09/M11 still produce a wider spread in *riiz*, but are not able to match the bluer part of the LRG's $r-i$ color. In addition to that, neither BC03 nor M09/11 models cover the bluer part of the $i-z$ color.

In Section 5.2 we already mentioned that the spread in the colors of the data cannot be completely accounted for by the photometric errors in this redshift range. The reason for the spread must thus be of physical origin. Unfortunately, not the whole range of $i-z$ of LRGs at $z \approx 0.4$ can be fitted by the investigated models sets while simultaneously matching the $g-r$ and $r-i$ colors (where the latter has smaller errors than $i-z$). This is a feature of BC03 and M09/11 models, and the reason why we could not create SEDs with bluer $i-z$ in Section 5.

The distributions of the χ^2 results of the fitting to the various model sets are shown in Figure B.4. Each analyzed z bin is displayed in a different panel, whereas the χ^2 distributions for all redshift intervals combined is shown in the lower right panel. BC03 outperforms the other models (M09 and M11), regarding the χ^2 values, in every redshift slice except for $z \approx 0.4$, where the M11 Pickles models fit the LRGs better, but only marginally. We would like to add that the M11 Pickles models fit the data best compared to the other M09/M11 SEDs, and outperform the Pickles models with a theoretical UV-part of the spectrum for each fitted object.

Finally, the lower right panel of Figure B.4 shows that the χ^2 values are best also for all z bins combined, as expected from the previous results. Following these results, and because of the previous analysis of the colors of the produced SEDs, we adopted the BC03 models for the creation of photo-z template SEDs.

B.2 Color-Color Relations of the New Models

Here we present the color–color relations of our models and the LRG data. Figures 5.2, and B.5–B.9 show the colors of the best fitting SED models, the preselected models, and the selected models after the removal of redundant SEDs in comparison to the LRG data itself in the *ugri*, *griz*, *grri*, and *riiz* color spaces. We display the LRGs with u band errors below the median and those with errors above with different colors to visualize the difference in the populated loci within the *ugri* space, which becomes evident at $z \approx 0.3$ (Figure B.8, upper

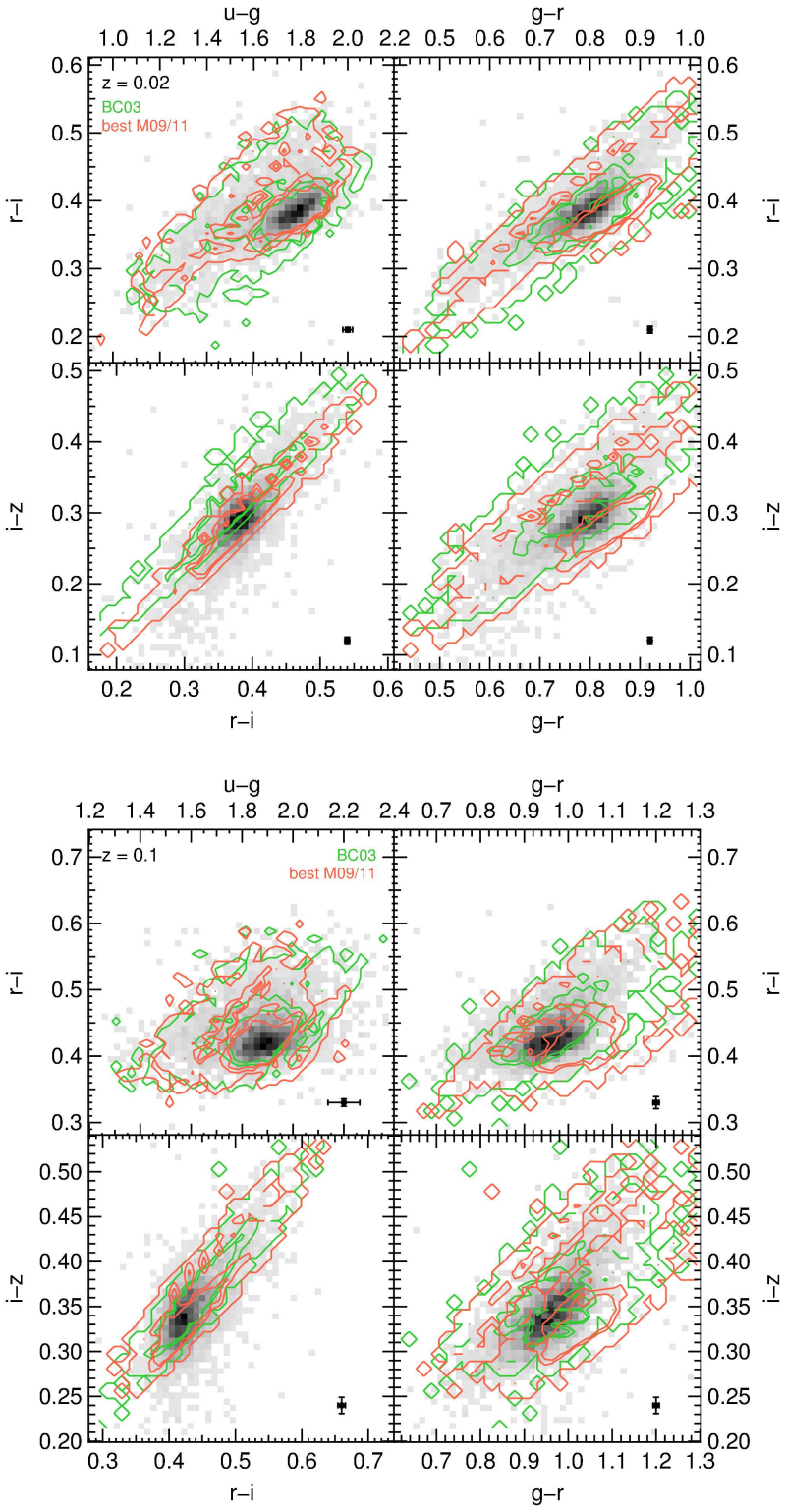

Figure B.1: Color distribution of LRGs (gray) at $z \sim 0.02$ and 0.1. Contours show the SED fitted BC03 (green), and M09/11 (red) models superposed with a burst model to fit the LRGs.

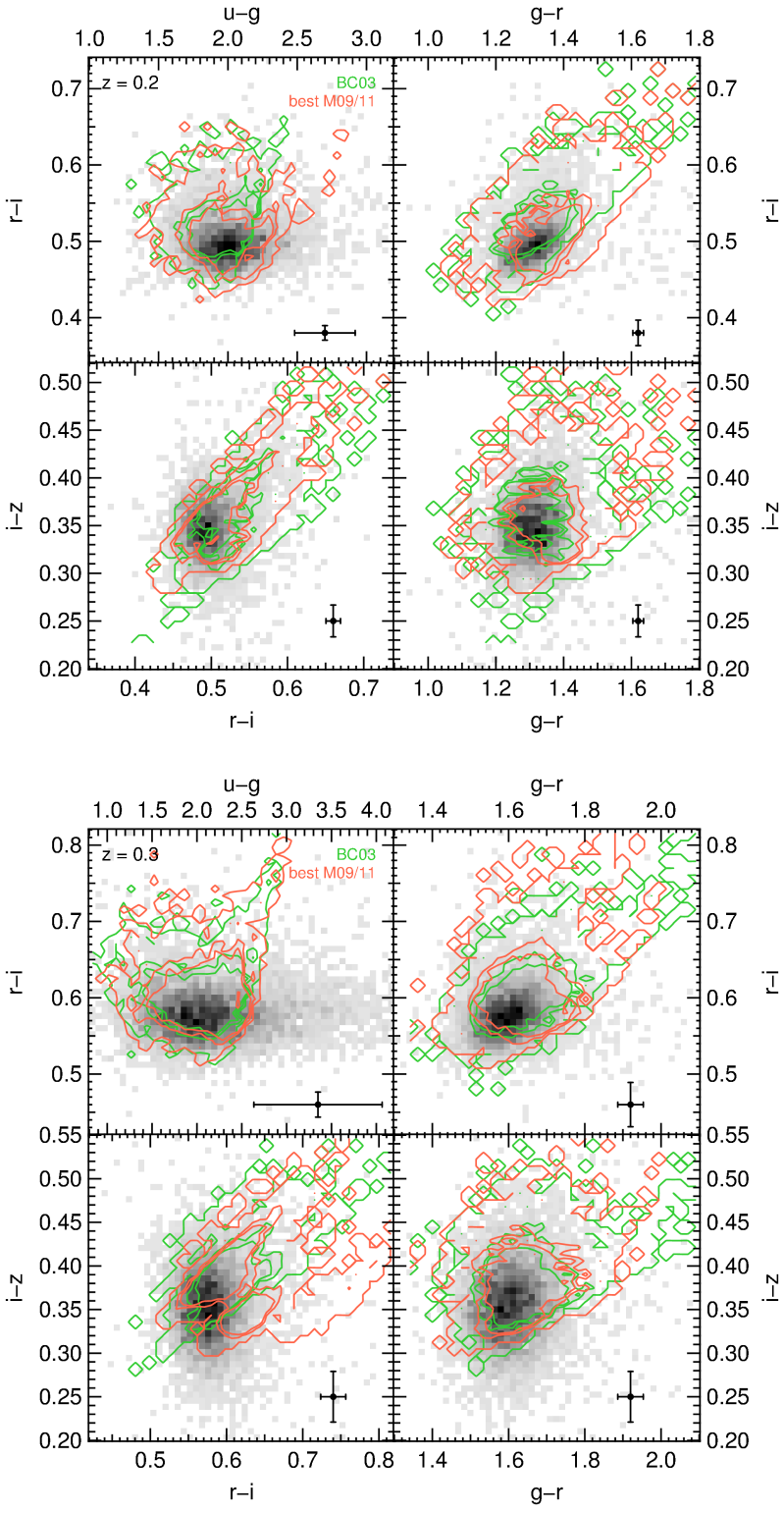

Figure B.2: Color distribution of LRGs (gray) at $z \sim 0.2$ and 0.3. Contours show the SED fitted BC03 (green), and M09/11 (red) models superposed with a burst model to fit the LRGs.

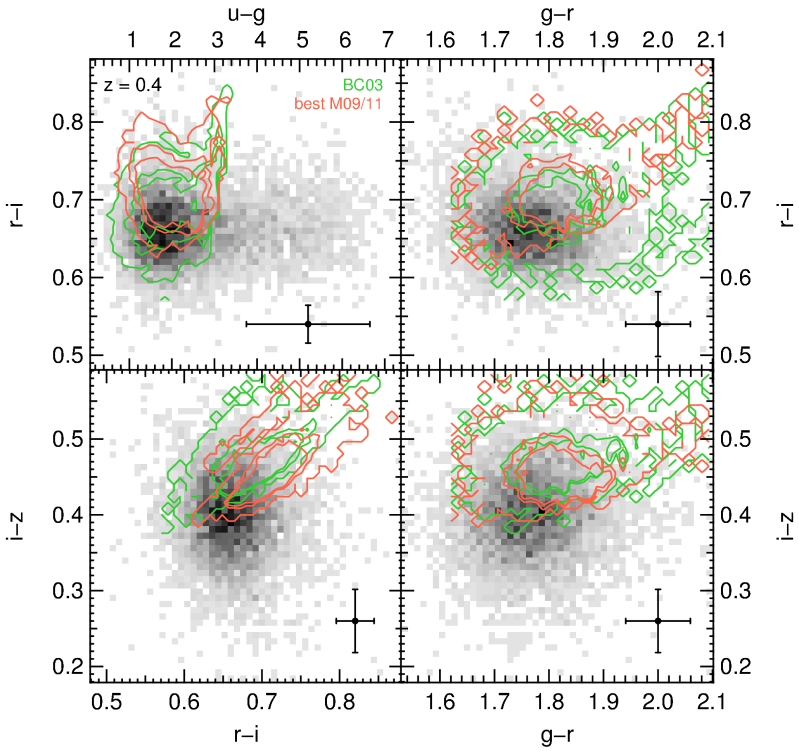

Figure B.3: Color distribution of LRGs (gray) at $z \sim 0.4$. Contours show the SED fitted BC03 (green), and M09/11 (red) models superposed with a burst model to fit the LRGs.

panel) and $z \approx 0.4$ (Figure B.9, upper panel).
The grids we imposed on the *ugri* plane within which we selected templates (see Section 5.2) are displayed in the upper left panels of Figures 5.2 , and B.6–B.9, whereas the color boundaries in $g - r$ and $i - z$ are drawn in the lower panel of Figure 5.2, in Figure B.5 and in the lower left and in the right panels of Figures B.6–B.9.
In order to account for objects with peculiar colors which lie outside the boundaries as well, we select additional SEDs by eye and confirm that they yield a photometric redshift bias lower than 0.01, when used as single template for the subcatalog from the same z range. If they do not meet that requirement, we chose the SED of another object from the color neighborhood of the first one, test for the bias, and so on until we find a model from that color locus which creates only a small $z_{\rm phot}$-bias.
Bright cluster galaxies (BCGs) with very low UV-excess (i.e., in the lower part of the UV red sequence; see Section 5.5.1, Figure 5.8) at $z \sim 0.1$ are not matched by the models already selected for that z range. But we find that two SEDs originating from $z \approx 0.2$ are able to cover the considered region. Therefore, we add these templates to $z \approx 0.1$, i.e., we assign the same redshift priors as for $z \approx 0.1$ models in addition to their former priors (see Section 2.2.2), thereby covering the lower UV red sequence. The colors of these two models are shown by yellow triangles in Figure B.6.

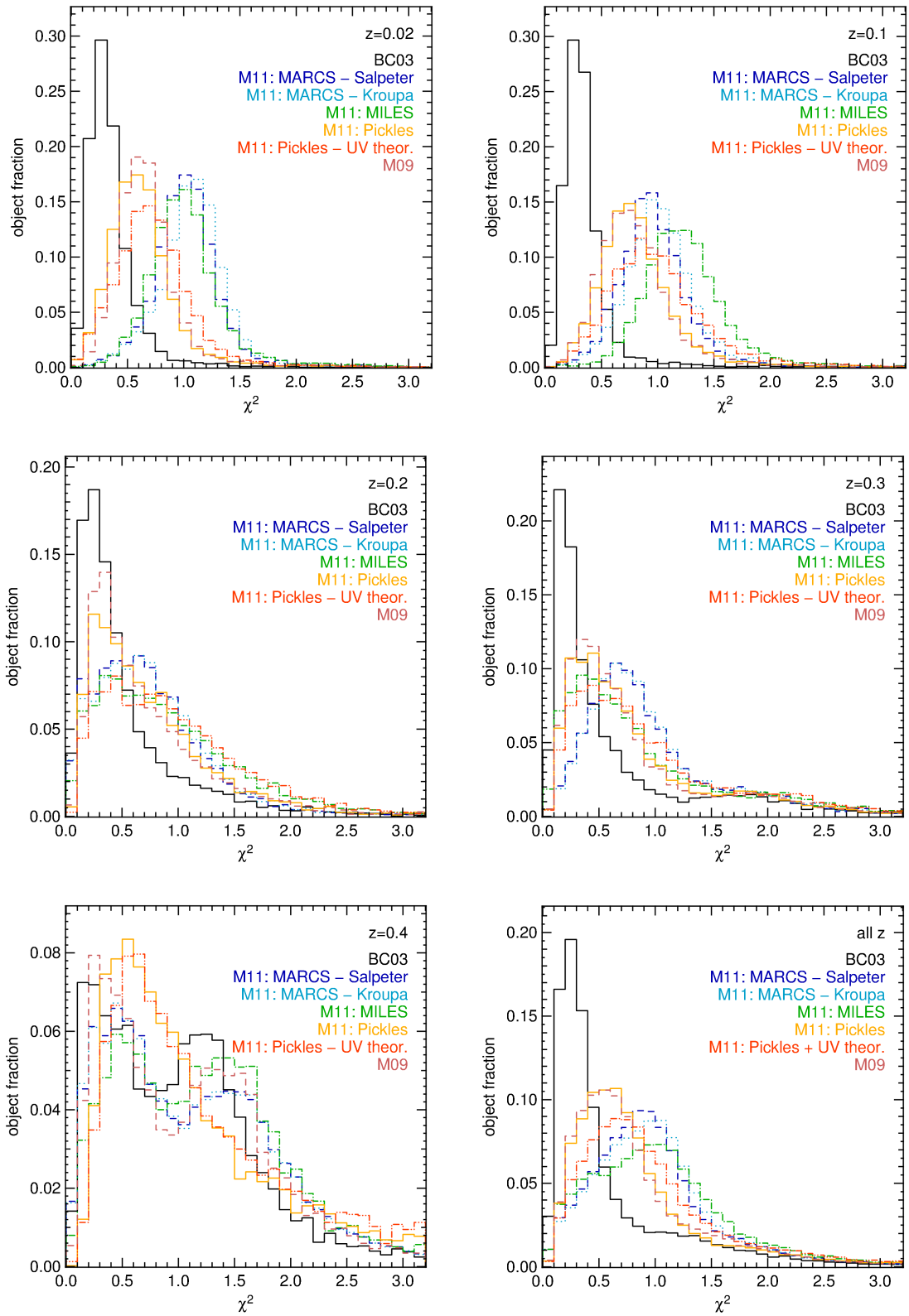

Figure B.4: χ^2 distribution of `SEDfit` results of different model-sets: BC03 (black), M11 MARCS stellar library with Salpeter IMF (blue), MARCS with Kroupa IMF (light blue), M11 MILES (green), M11 Pickles (yellow), Pickles with theoretical UV (red), and M09 (light red). The different panels show the various redshift slices in which the fitting was performed. The lower right panel finally displays the χ^2 results in all redshift intervals combined.

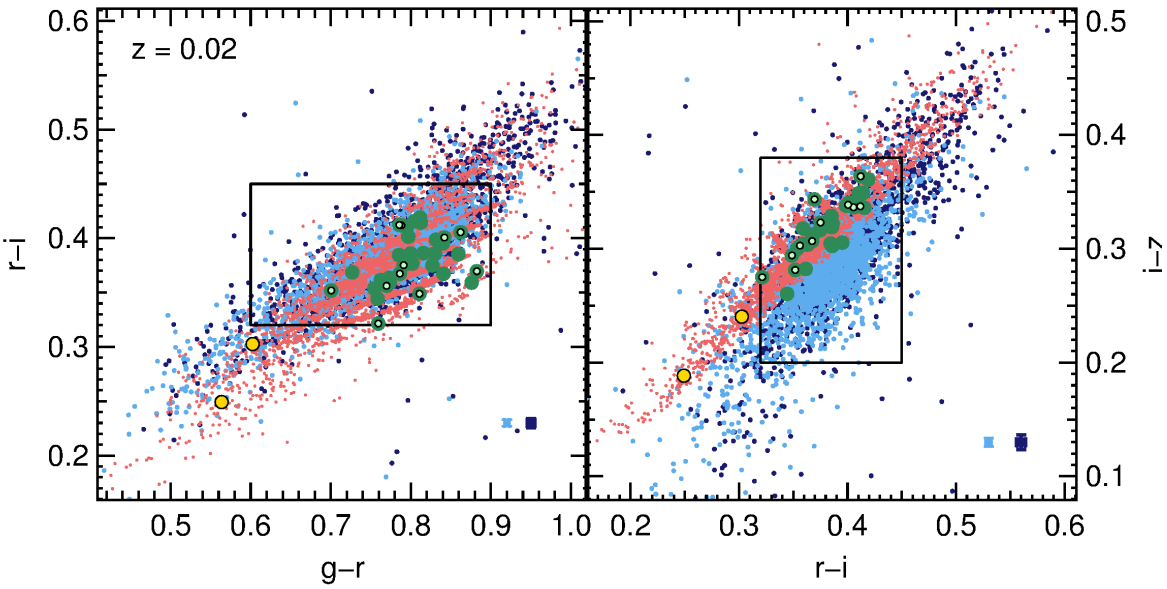

Figure B.5: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.0 \leq z \leq 0.04$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The mean error bars of all objects are shown in the lower right corners in dark blue, and the mean errors of objects with u band errors smaller than the median in light blue (although they cannot be resolved in this figure, since they are very small). The boundaries within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs. The yellow dots are models that shall account for objects outside the selected boundaries in *ugri* and *griz*.

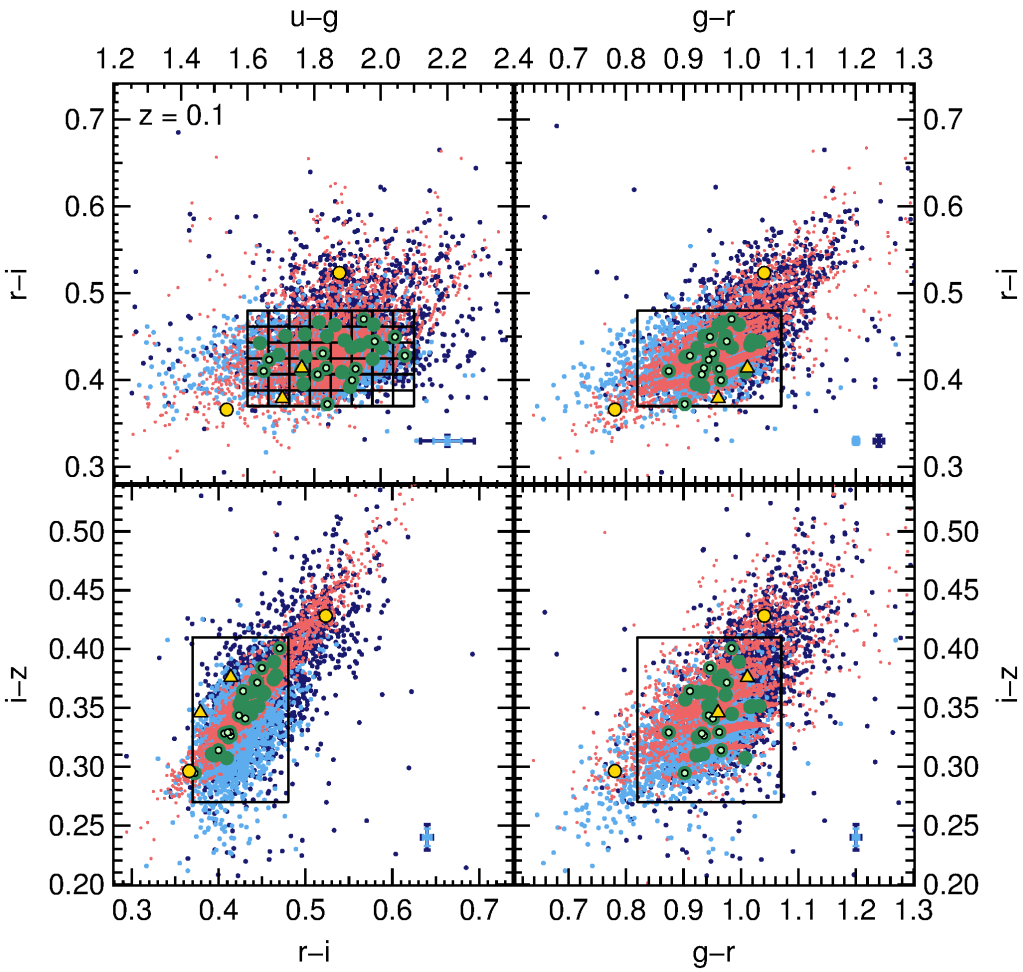

Figure B.6: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.08 \leq z \leq 0.12$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The mean error bars of all objects are shown in the lower right corners in dark blue, and the mean errors of objects with u band errors smaller than the median in light blue (although they cannot be resolved in all panels). The grid (upper panel) and the boundary (lower panel) within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs. The yellow dots are models that shall account for objects outside the selected boundaries in *ugri* and *griz*. Yellow triangles are models from $z \approx 0.02$ that lie in the lower part of the UV red sequence of R12 (see Section 5.5.1, Figure 5.8) and are added additionally in this redshift range.

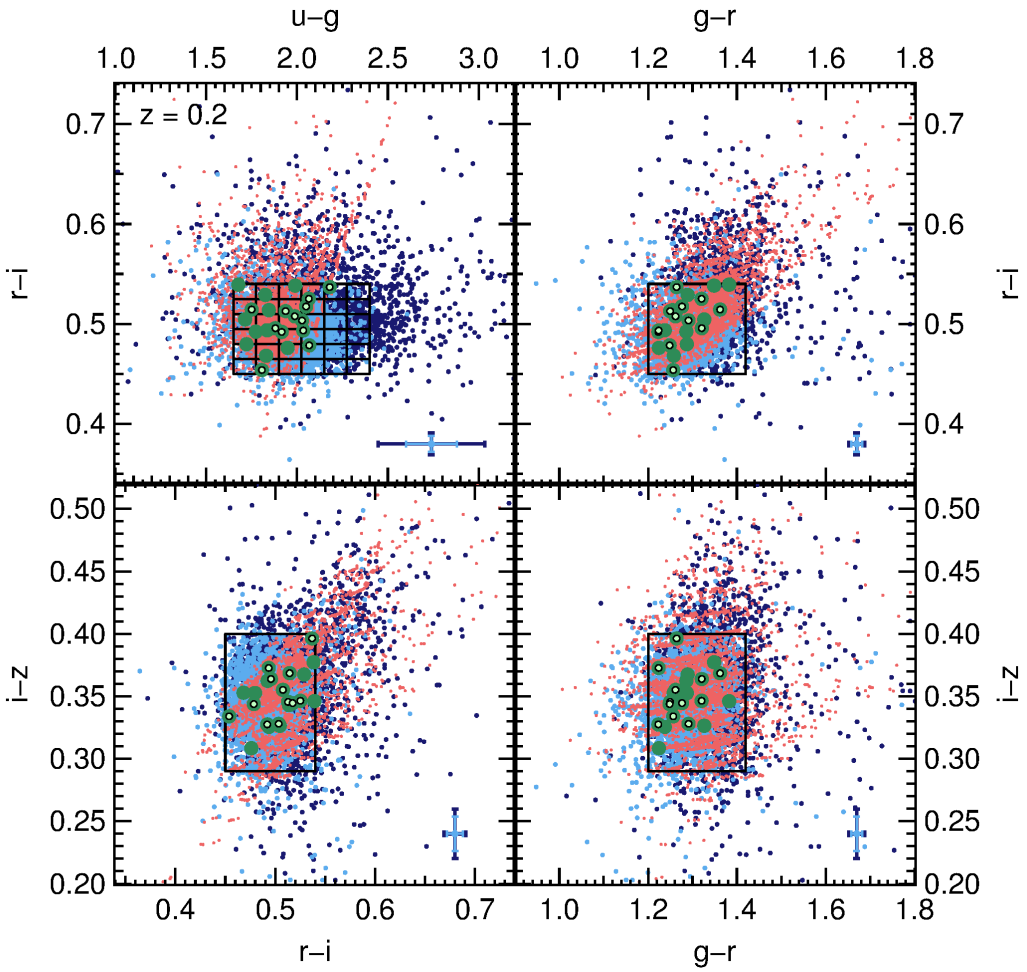

Figure B.7: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.18 \leq z \leq 0.22$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The mean error bars of all objects are shown in the lower right corners in dark blue, and the mean errors of objects with u band errors smaller than the median in light blue. The grid (upper panel) and the boundary (lower panel) within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs.

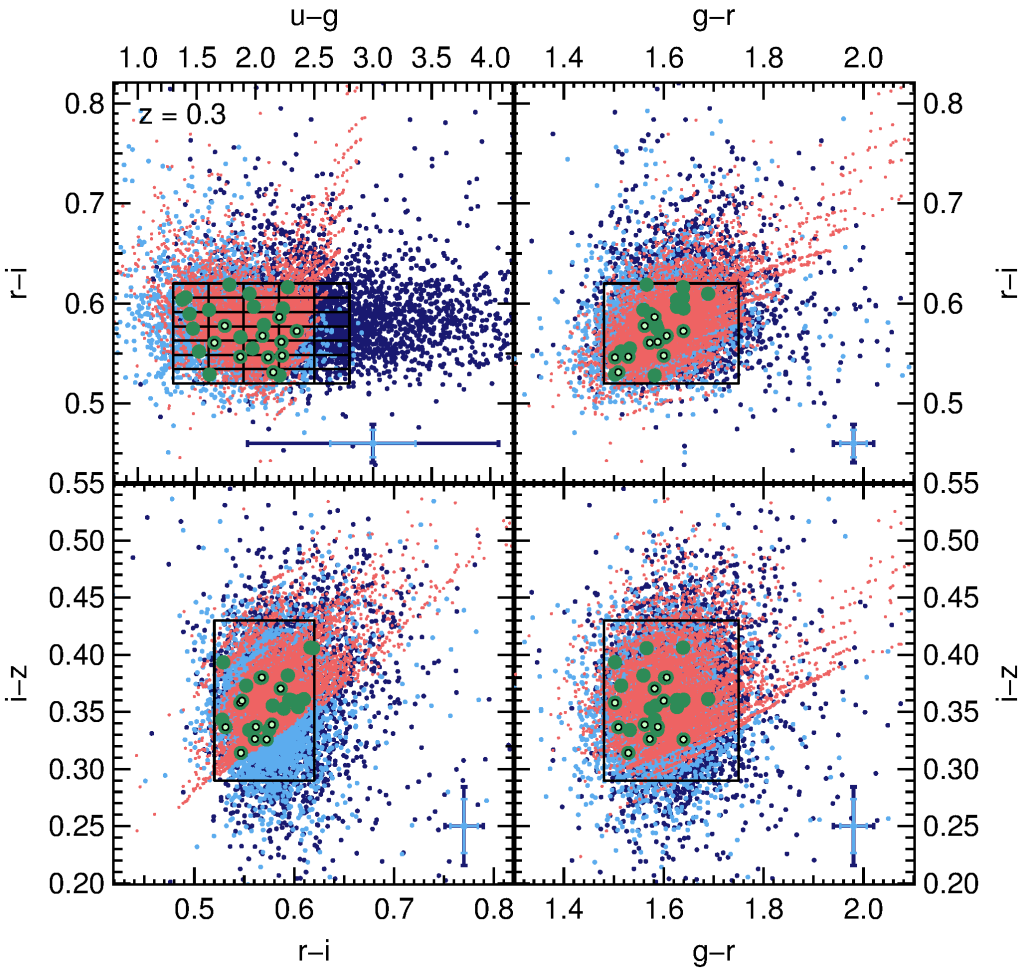

Figure B.8: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.28 \leq z \leq 0.32$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The mean error bars of all objects are shown in the lower right corners in dark blue, and the mean errors of objects with u band errors smaller than the median in light blue. The grid (upper panel) and the boundary (lower panel) within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs.

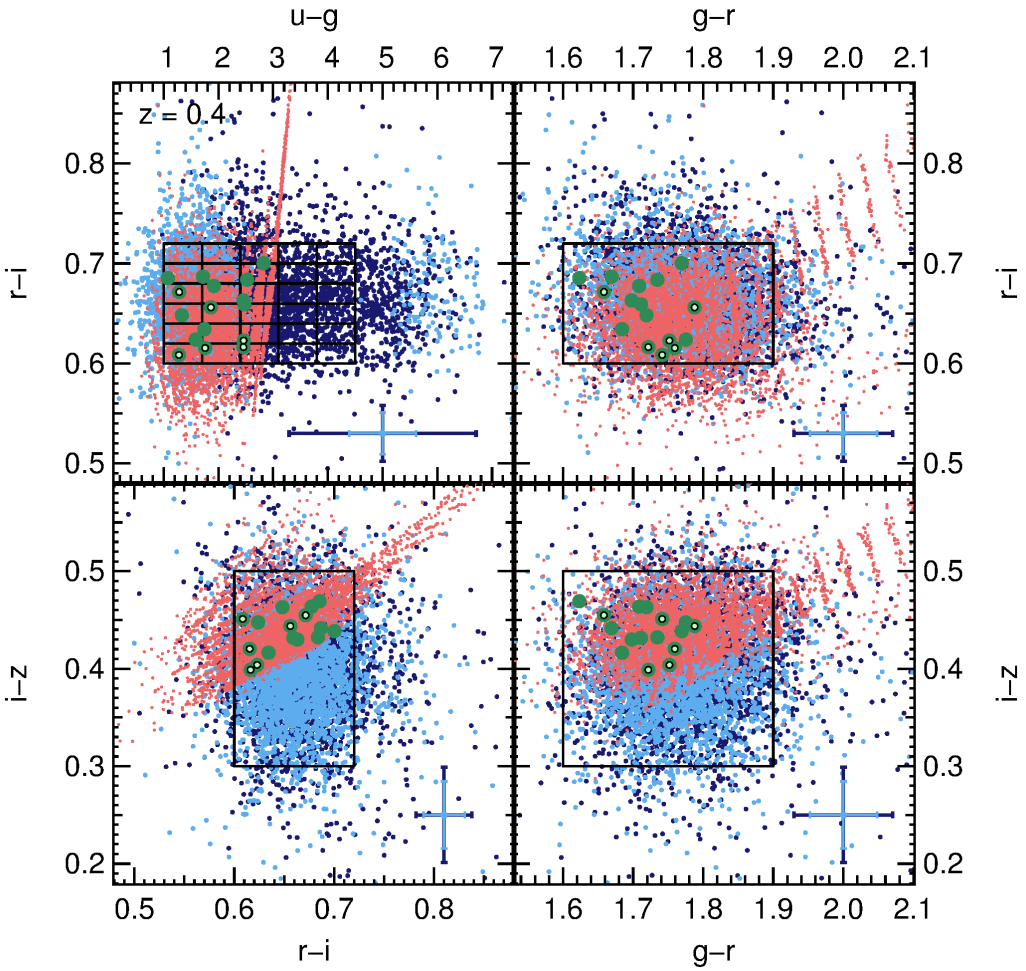

Figure B.9: Color vs. color plots for SDSS data (blue) and their individual best fitting `SEDfit`-SEDs (red) for $0.38 \leq z \leq 0.42$. Objects with u band errors lower than the median are indicated in light blue, whereas those with greater errors are dark blue. The mean error bars of all objects are shown in the lower right corners in dark blue, and the mean errors of objects with u band errors smaller than the median in light blue. The grid (upper panel) and the boundary (lower panel) within which the models are selected are shown in black. The dark green dots are the preselected models, whereas the light green points represent the models that are left over after the removal of redundant SEDs.

Appendix C

SED Fitting Results for the BOSS CMASS Sample

In this paper we create models that should match the BOSS CMASS galaxies in terms of their colors in redshift bins with widths 0.04 centered on $z = 0.5, 0.6, 0.7, 0.8$. We do so by fitting sets of model SEDs with different properties like metallicity, SFHs, and ages to the data (Sec. 6.2.1). In superposition to the SEDs, a burst model and additional dust extinctions for the burst and the main component are further degrees of freedom. In Fig. C.1 we show the distributions of the χ^2 values returned by the SED fitting code, when we fit the BC03, M05, and M11 models in equal setups to the data. The setup of the SED fitting and the code are detailed in Sec. 6.2.1 and are equal for all cases.

In Sec. 6.2.1 we modify the BC03 models, changing their continuum slope by multiplying with $a \cdot \lambda^{-\beta}$ for $\lambda \geq \lambda_i$ with various β and λ_i values. This is done to introduce further degrees of freedom and to create models that match the colors of the data to a better extent. We fit the $\lambda^{-\beta}$ modifications of BC03 SSPs and CSPs with decreasing SFRs to the CMASS galaxy colors with the same setup as mentioned above and as detailed in Sec. 6.2.1. The resulting χ^2 values of the fit are shown together with the previous results of the BC03, M05, and M11 models in Fig. C.1, where they are highlighted by orange histograms.

While the results of the original BC03 models, M05, and M11 produce similar results in terms of χ^2, they are significantly improved by using the BC03 $\lambda^{-\beta}$ models. This is also shown in terms of colors in Figs. 6.9 to 6.12 of Sec. 6.2.1.

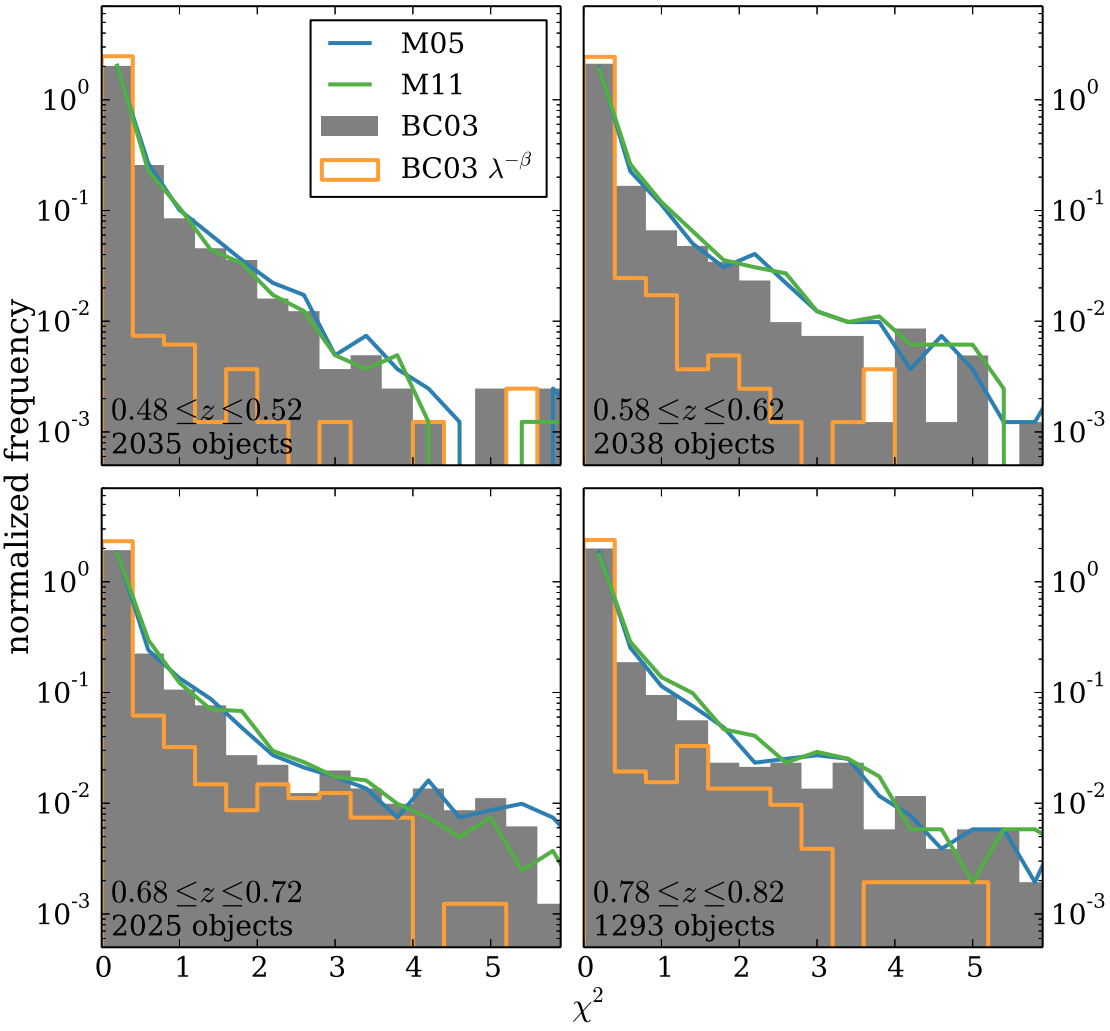

Figure C.1: χ^2 distributions of SED fitting runs (at $z_{\rm spec}$) detailed with BC03 (gray), BC03 $\lambda^{-\beta}$ (orange), M05 (blue), and M11 (green) models for the four considered redshift regions. The frequency is normalized and plotted logarithmically. We can see that although the distributions for BC03, M05, and M11 are similar, the χ^2 values are improved by including the BC03 $\lambda^{-\beta}$ models.

Appendix D

SED Fitting Results for the Quintuply Lensed High-z System in RXC J2248

In Chapter 8 we estimated the physical properties of the high-z quintuply lensed system in RXC J2248 by SED fitting with `SEDfit`. We present here the results of the SED fitting in greater detail. We concentrate again only on ID2&3 which have the cleanest photometry. Figures D.1 to D.4 display the results for exponentially increasing ($\tau < 0$) and decreasing ($\tau > 0$) SFRs. Figures D.5 D.6 show the fitting results with SSP models only, whereas Figures D.7 and D.8 display the results when we perform the SED fit to all models (SSPs and CSPs) combined. In the lower panels of Figures D.1 to D.8 we plot the best fitting SED and the photometry. The panels in the middle and upper rows show the 2-dimensional likelihood distributions of fitting parameters (Z, τ, model age, $A_{\rm V}$), whereas the probability distributions of the mass-to-light ratios (M/L) in the V band are shown in the upper panels. The blue contours in the likelihood distributions outline the 1σ (solid), 2σ (dashed), and 3σ (dotted) confidence levels. The filter bands in which the S/N ratio does not exceed 1 (the dropout filters f225w to f775w) are considered upper limits in the SED fit (lower panels) and are highlighted by empty circles.

Figure D.9 shows furthermore the probability distributions of the model ages, marginalized over the other fitting parameters. The upper limits of the 2σ interval in age is smallest for the fit with SSP models (third panel of Fig. D.9). This is because all stars are assumed to form at the time of formation and therefore the mean stellar age of a SSP is higher than that of a CSP with the same formation redshift observed at the same cosmic time.

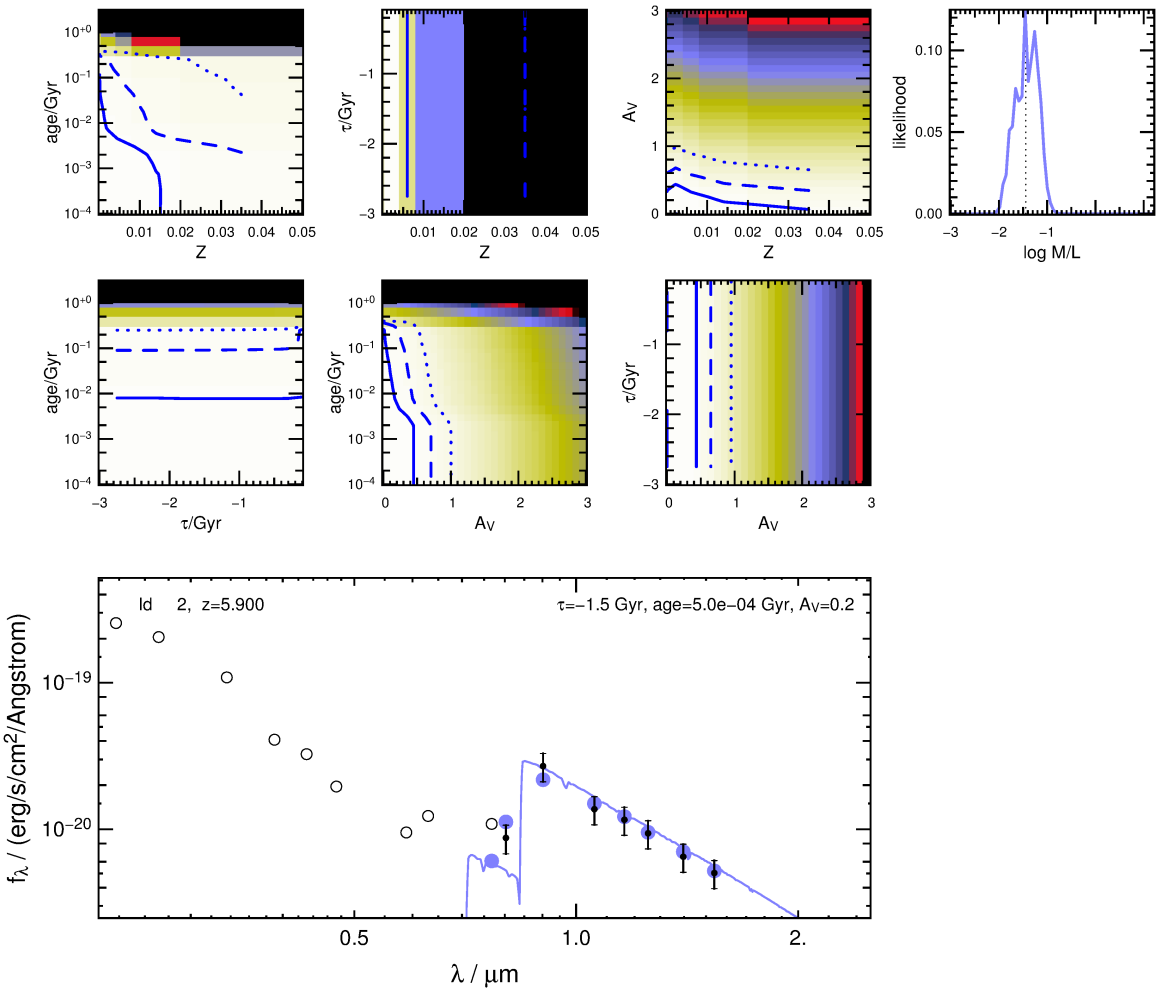

Figure D.1: SED fitting results for ID2 and models with negative τ SFRs. The lower panel shows the input photometry and errors with black points. Empty circles denote the fluxes in the dropout filters where the fluxes are considered upper limits. The best fitting model SED is shown in blue and the convolved fluxes in the detection bands are displayed by filled circles. The density plots in the upper panels show the likelihood distributions of the SED fit in two-dimensional parameter spaces. Blue lines denote the 1σ (solid), 2σ (dashed), and 3σ (dotted) confidence levels. Finally, the likelihood distribution of the mass-to-light ratio in the V band is plotted in the upper right panel. The dotted line in this panel denotes the M/L ratio of the best fitting model.

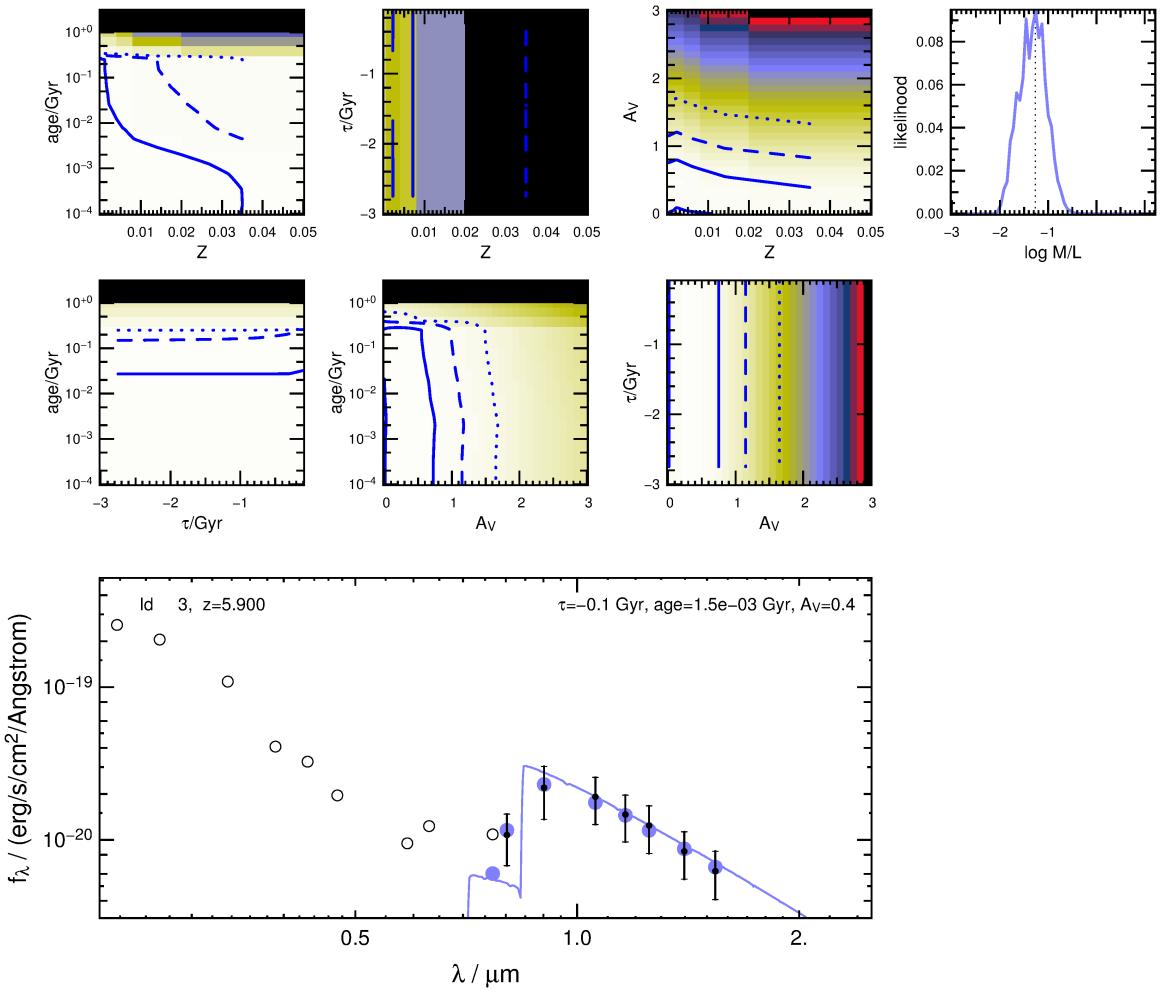

Figure D.2: SED fitting results for ID3 and models with negative τ SFRs. For a detailed description of the plot see Fig. D.1.

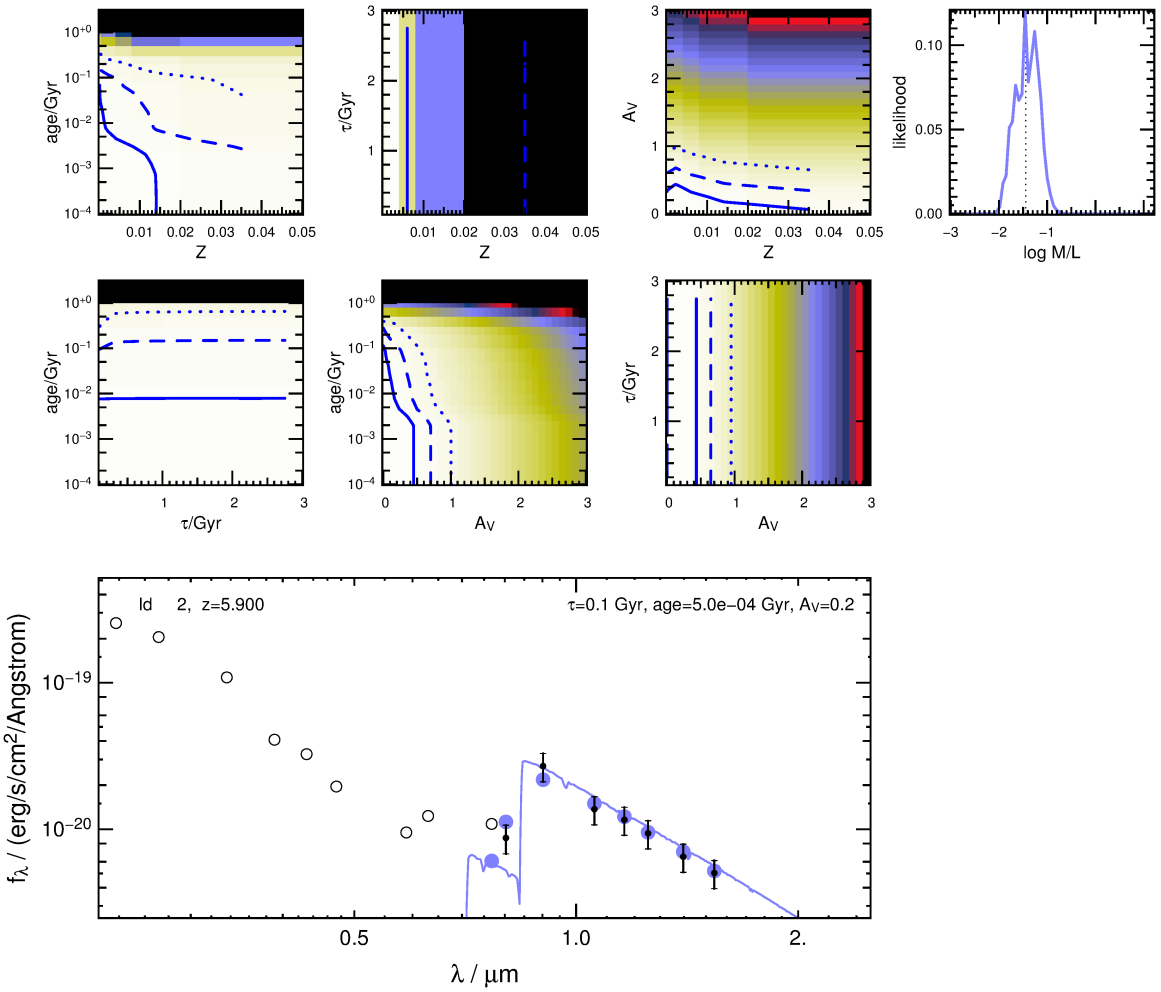

Figure D.3: SED fitting results for ID2 and models with positive τ SFRs. For a detailed description of the plot see Fig. D.1.

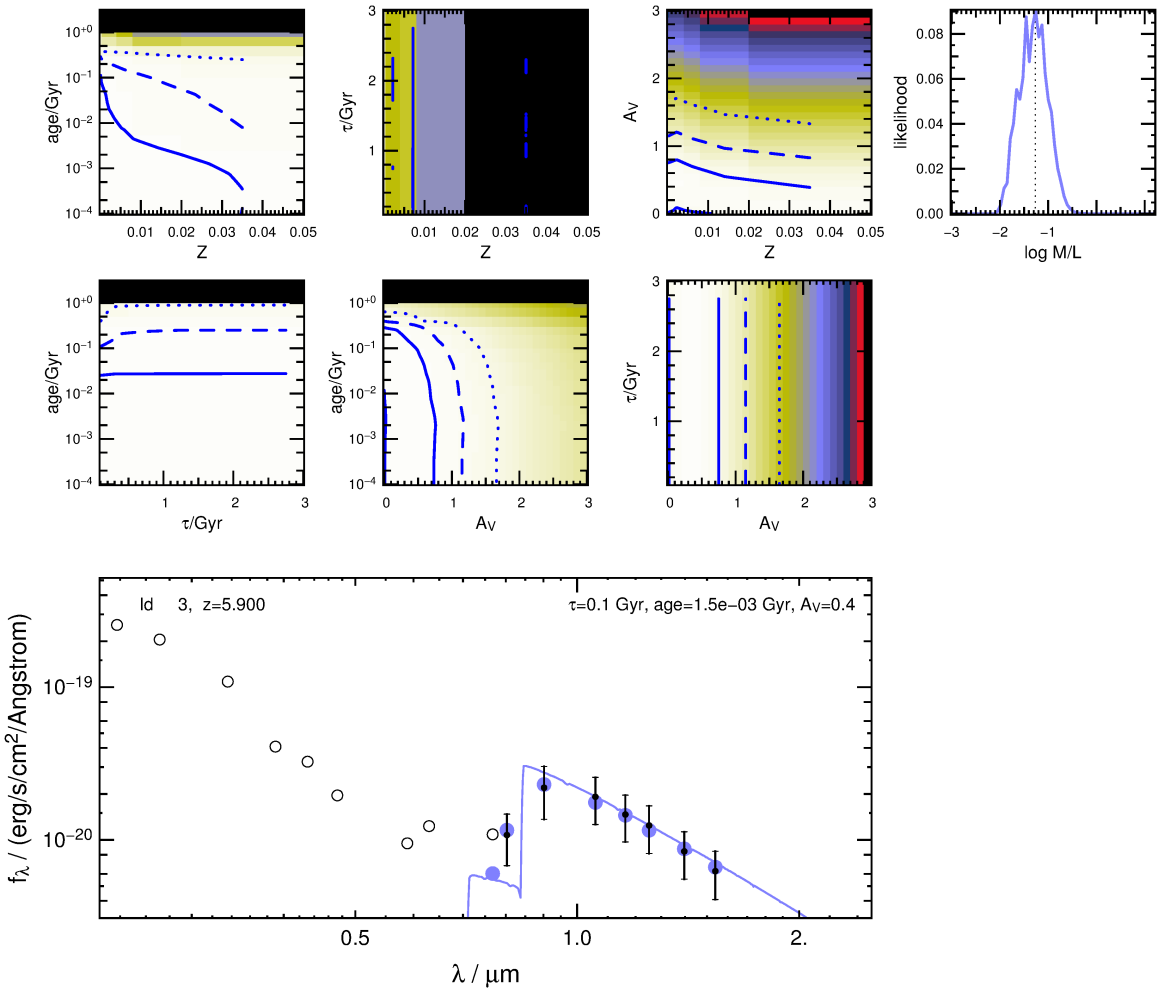

Figure D.4: SED fitting results for ID3 and models with positive τ SFRs. For a detailed description of the plot see Fig. D.1.

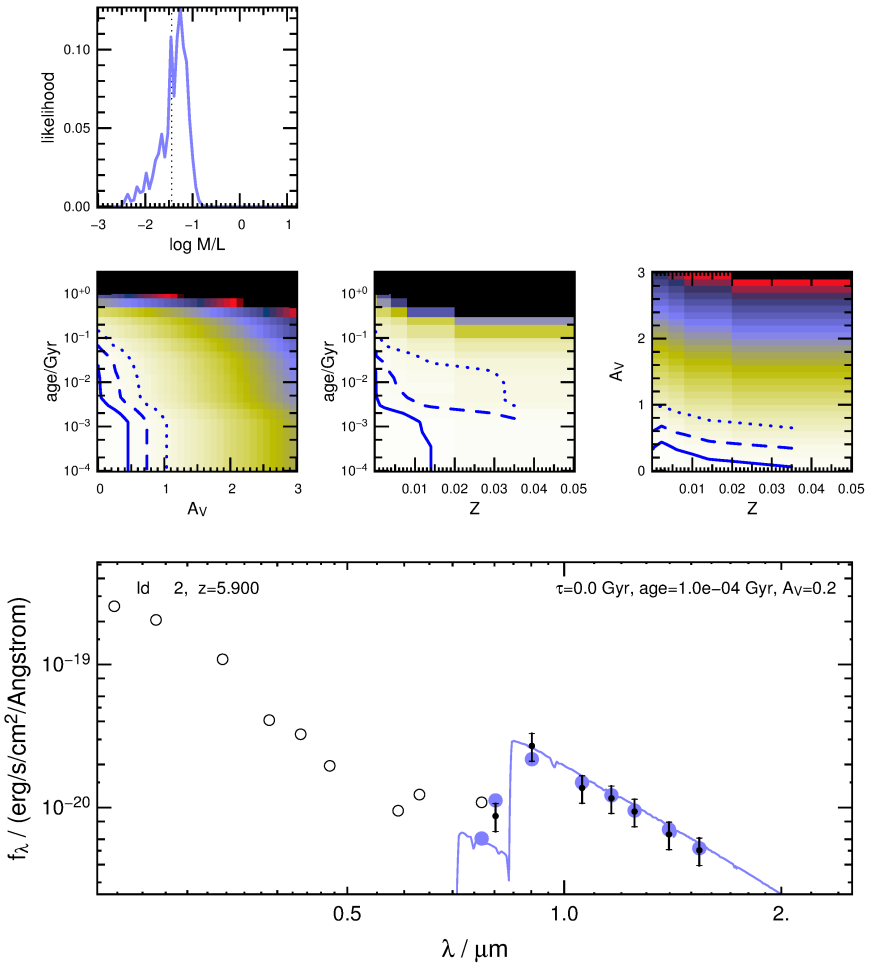

Figure D.5: SED fitting results for ID2 and SSP models. For a detailed description of the plot see Fig. D.1.

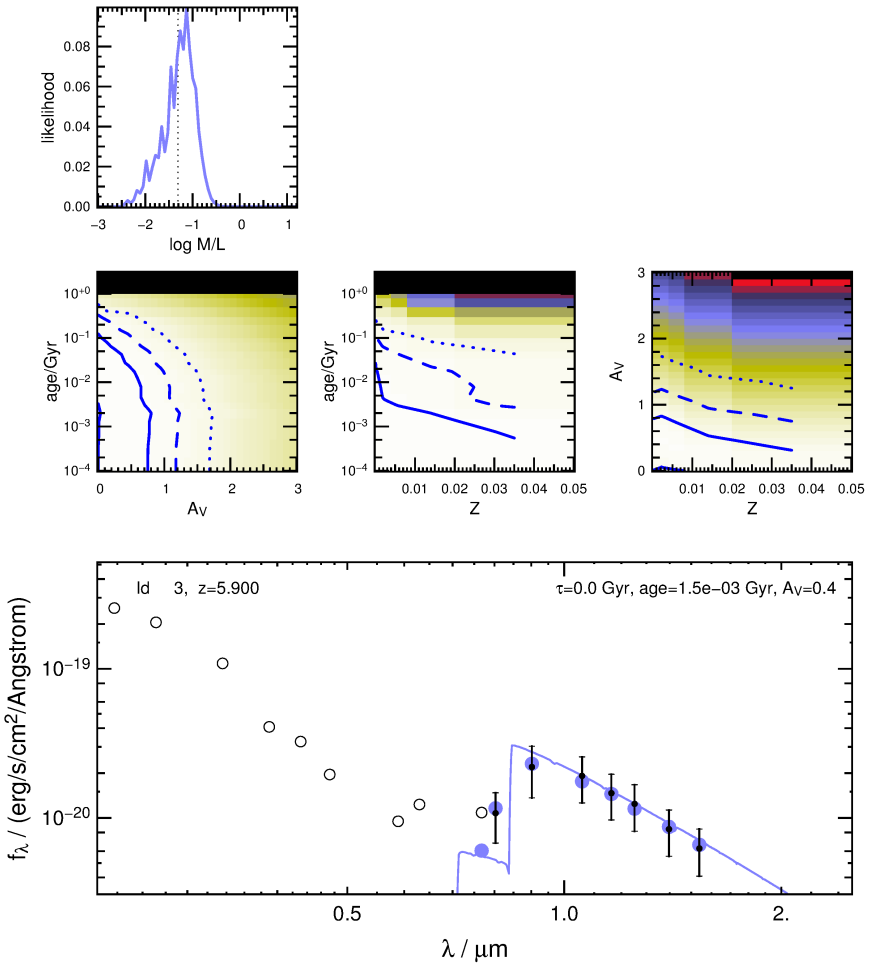

Figure D.6: SED fitting results for ID3 and SSP models. For a detailed description of the plot see Fig. D.1.

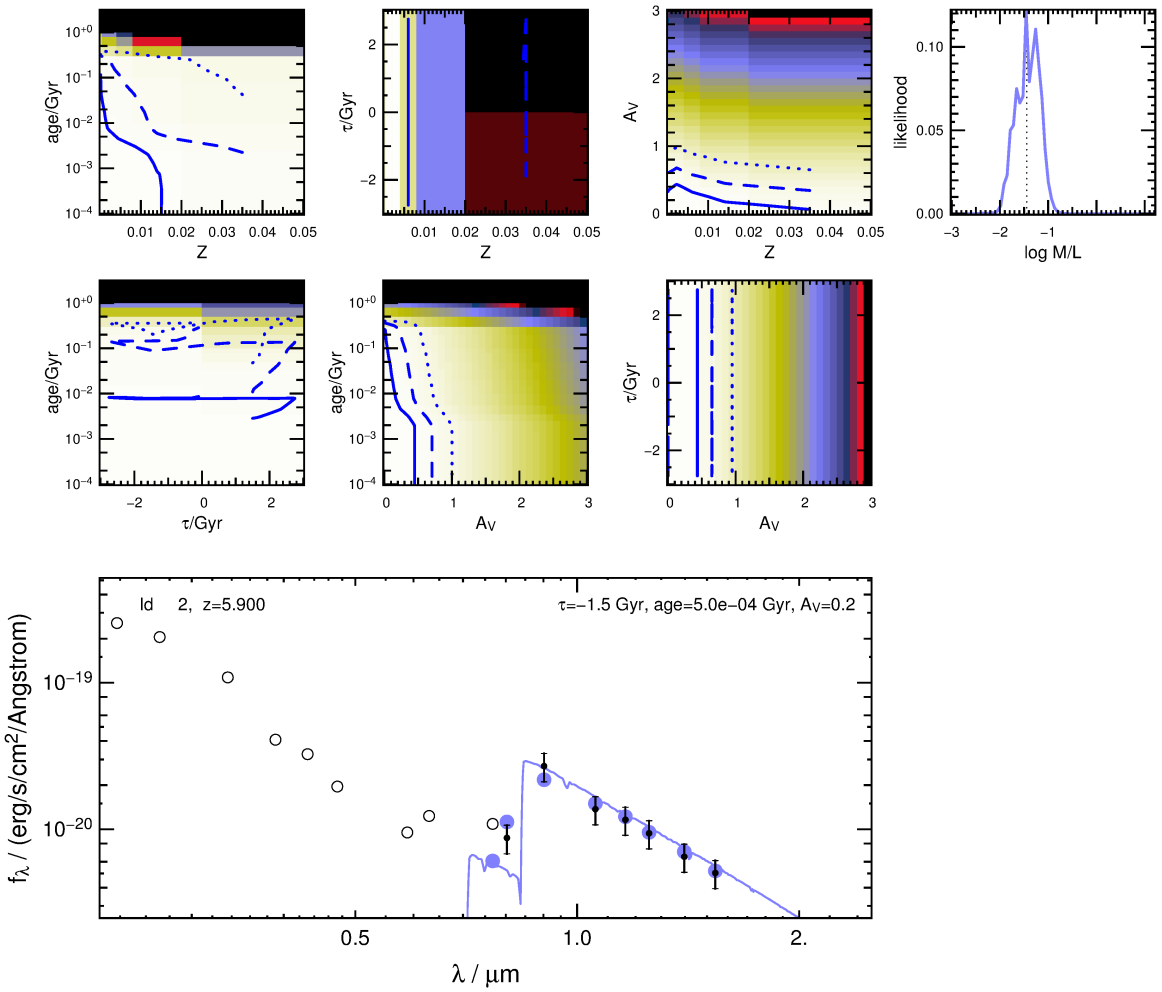

Figure D.7: SED fitting results for ID2 and all models (CSPs and SSPs). For a detailed description of the plot see Fig. D.1.

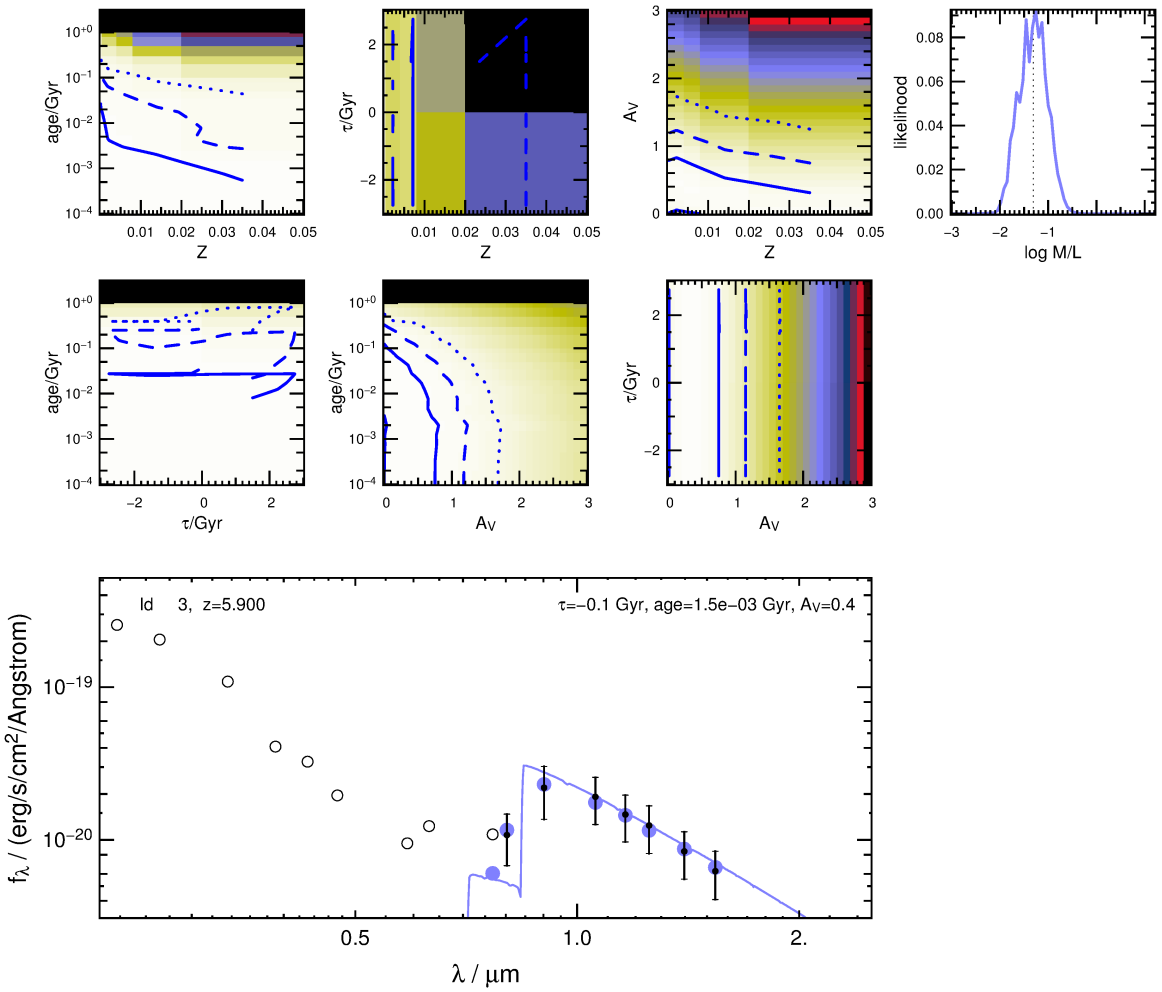

Figure D.8: SED fitting results for ID3 and all models (CSPs and SSPs). For a detailed description of the plot see Fig. D.1.

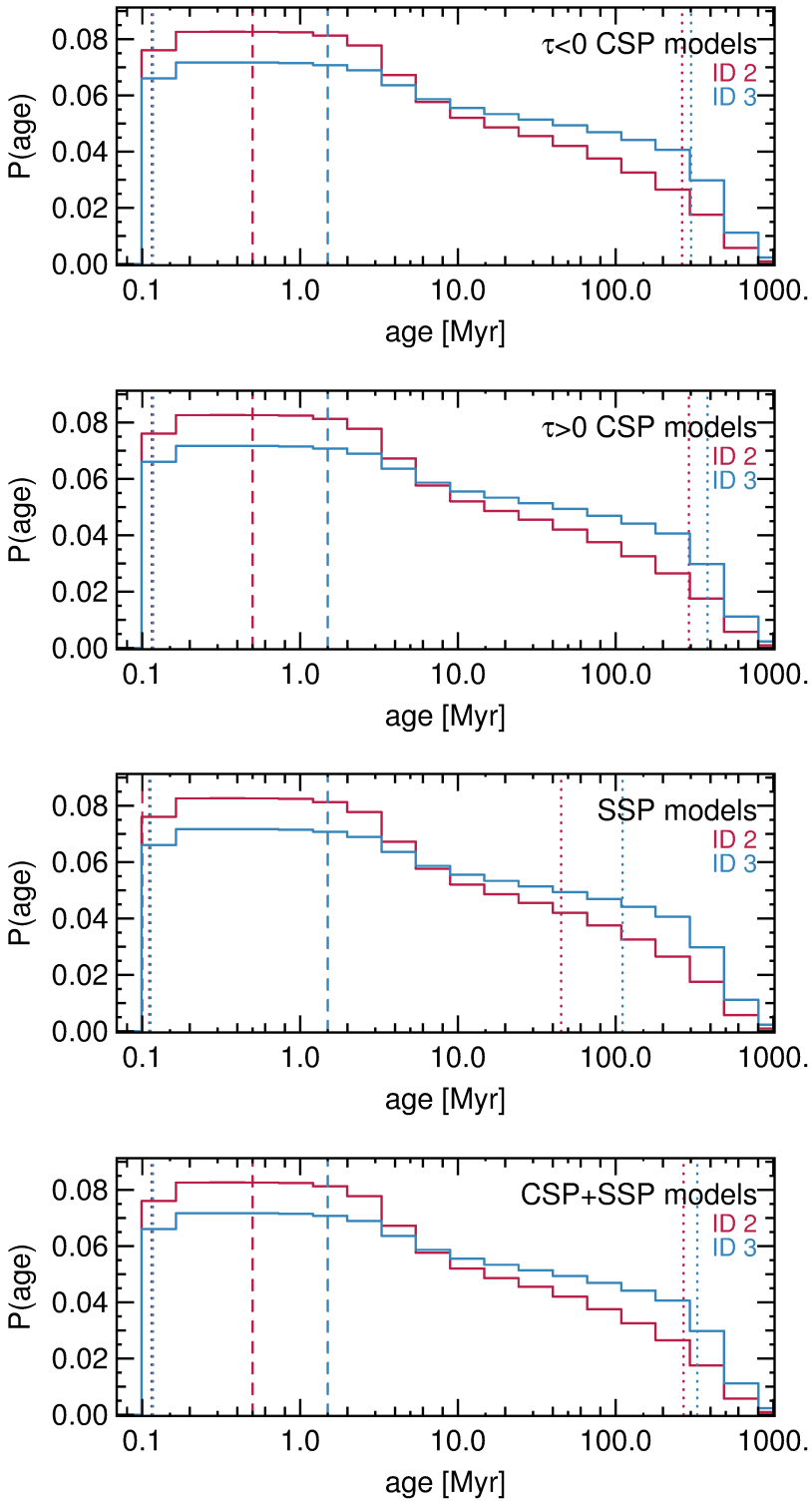

Figure D.9: PDF of the model age for ID2&3, marginalized over the other fitting parameters for a fit with $\tau < 0$, $\tau > 0$ CSPs, SSP models, and for all models combined (from upper to lower panel). The dashed lines are the ages of the best fitting models. Between the dotted lines the total probability is 95.45 % (corresponding to a 2σ confidence interval).

Appendix E

SDSS CasJobs Queries

This chapter contains the queries used for the acquisition of the SDSS data from the SDSS database in order to give the reader more information about the catalogs and to provide the means to duplicate any SDSS catalogs used in this work. SDSS provides the Java program CasJobs on their database website[1] which enables the user to write database queries in the Structured Query Language (SQL). If desired, queries can also be executed on the CasJobs website. These queries will be demonstrated in the following, where we will also give an explanation of the different data values acquired as well as the tables in the SDSS database from which they originate.

E.1 SDSS-II Spectroscopic LRG Sample

In Section 5 we worked with the Spectroscopic LRG sample from SDSS-II Data Release 7 (DR7) presented in greater detail in Section 3.1. Therefore, when executing this query one must set the context to DR7, the latest data release of SDSS-II. For additional information on the parameters and tables we refer to the SDSS-II Schema Browser[2]. When using these codes, <fil> has to be replaced by either u, g, r, i, or z, since <fil> is just a placeholder to shorten the codes.

```
select sp.ObjID, sp.SpecObjID, so.z, so.zErr, so.zConf, so.zStatus,
       sp.RA, sp.Dec,
       sp.ModelMag_<fil>,  sp.ModelMagErr_<fil>,
       sp.Extinction_<fil>, sp.Dered_<fil>,
       sp.PetroMag_<fil>, sp.PetroMagErr_<fil>,
       po.FracDeV_r,
       pz.z as PhotozTF, pz.zErr as PhotozTFErr, pz.absMag_r, pz.KCorr_r,
       pz2.PhotozCC2, pz2.PhotozErrCC2, pz2.PhotozD1, pz2.PhotozErrD1,
       Ha.EW as Ha6565_EW, Ha.EWErr as Ha6565_EWErr,
       Hb.EW as Hb4863_EW, Hb.EWErr as Hb4863_EWErr, sppP.zbSubClass,
       so.PrimTarget
from SpecPhoto as sp
```

[1] http://skyserver.sdss3.org/CasJobs/default.aspx
[2] http://skyserver.sdss2.org/dr7/en/help/browser/browser.asp

```
      join SpecObjAll as so on sp.SpecObjID = so.SpecObjID
      join Photoobjall as po on sp.ObjID = po.ObjID
      join Photoz as pz on sp.ObjID = pz.ObjID
      join Photoz2 as pz2 on sp.ObjID = pz2.ObjID
      join sppParams as sppP on sp.SpecObjID = sppP.SpecObjID
      join SpecLine as Ha on sp.SpecObjID = Ha.SpecObjID
      join SpecLine as Hb on sp.SpecObjID = Hb.SpecObjID
where Ha.LineID = 6565 and Hb.LineID = 4863
       and so.zErr<0.01 and so.zConf>0.99
       and (so.zStatus=4 or so.zStatus=7 or so.zStatus=9)
       and so.zWarning=0x00000000 and so.Specclass=2
       and ((so.PrimTarget & 0x00000020)>0 or (so.PrimTarget & 0x04000000)>0)
```

ObjID, SpecObjID SDSS object IDs. In case of the Photometric ID (`ObjID`) these are unique object identifiers that stem from the camera column, run, and field. `SpecObjID` is an identifier for the spectroscopic sample.

RA, Dec Right ascension (RA) and declination (Dec) are the coordinates in the J2000 equatorial coordinate system.

PrimTarget Primary target category bit mask. Set to select objects from the LRG sample only. That is the object is either selected by *CutI* `TARGET_GALAXY_RED` (`0x00000020`), or *CutII* `TARGET_GALAXY_RED_II` (`0x04000000`), or both. The color and magnitude cuts of the LRG sample are explained in Eisenstein et al. (2001), Padmanabhan et al. (2005), and in Section 3.1.

SpecClass A spectroscopic classification of `2` ensures only galaxies are selected.

ModelMag The model magnitude is the magnitude extracted from determining the better fit between an exponential $I(r) = I_0 \exp(-1.68\, r/r_e)$ and a de Vaucouleurs profile $I(r) = I_0 \exp -7.67[(r/r_e)^{1/4}]$ (de Vaucouleurs, 1948) to the surface brightness of the object.

FracDev_r The `FracDev` value is the weight of the de Vaucouleurs component in the best composite (exponential + de Vaucouleurs) model in the r band.

extinction Galactic extinction inferred from extinction maps of Schlegel et al. (1998).

dered Extinction corrected model magnitudes.

PetroMag The Petrosian magnitude Petrosian (1976), modified as in Blanton et al. (2001) and Yasuda et al. (2001). In order to measure a constant fraction of the light of a galaxies, a radius is set which defines a circular aperture within which the flux is measured. The *Petrosian Ratio* $\mathcal{R}_P$ is the ratio between the azimuthally averaged surface brightness profile $I(r)$ at radius r to the mean $I(r)$ within r. In the SDSS, r_P is determined by $\mathcal{R}_{P,\mathrm{lim}} = 0.2$.

EW, EWErr Equivalent width and error of the desired spectral line.

LineID Identifier for spectral lines. In this case it is the Hα line at $\lambda = 6563\,$Å (`LineID 6565`) and the Hβ line at 4863 Å (`LineID 4863`).

z, zWarning, zStatus, zConf Spectroscopic redshift `z` and associated measurement parameters. `zWarning` set to zero means that no problems were encountered during the determination of the spectroscopic redshift. `zStatus` set to `4 or 7 or 9` ensures that the spectroscopic redshift was measured with high confidence. Additionally, one can set the confidence level `zConf` of the spec-z measurement to a desired value.

z, AbsMag_r from Photoz Photometric redshift and r band absolute magnitude determined by a combination of a polynomial k-d tree nearest neighbor fit and template fitting (Csabai et al., 2000). The results on the LRG sample are presented in Section 5.4, Figure 5.5.

PhotozCC2, PhotozD1 from Photoz2 Photometric redshifts from a neural network code (Oyaizu et al., 2008). Both use concentration indices for the photo-z estimation. The `CC2` version furthermore includes galaxy colors, whereas the `D1` method includes only magnitudes. The results are shown in Section 5.4.

All magnitudes extracted from SDSS are *asinh* magnitudes or sometimes informally called *luptitudes* (Lupton et al., 1999). They were introduced to account for fluxes equal to, or lower than zero which cannot be converted into Pogson's $\log_{10}$ magnitudes (Pogson, 1856), because of the logarithm. The *asinh* magnitudes are defined as

$$m_{\text{asinh}} = -\frac{2.5}{\ln(10)} \cdot \left[\text{asinh}\left(\frac{f/f_0}{2b}\right) + \ln(b)\right], \tag{E.1}$$

where f_0 is the $\log_{10}$ zero point flux. b is a softening parameter, i.e., the typical 1σ noise of the sky in a PSF aperture in $1''$ seeing. It is listed in table E.1 (from Stoughton et al., 2002) together with the asinh magnitude for objects with $f = 0$.

filter	$\boldsymbol{b}\,[\mathbf{10^{-10}}]$	$\mathbf{mag(f = 0)}$
u	1.4	24.63
g	0.9	25.11
r	1.2	24.80
i	1.8	24.36
z	7.4	22.83

Table E.1: Asinh Magnitude Softening Parameters b in the SDSS Filters and the asinh magnitude for an object with $f = 0$ (from Stoughton et al., 2002).

At high signal to noise ratios the asinh are aimed to be virtually identical to the $\log_{10}$ magnitudes[3]. However, we found that the error of $\log_{10}$ magnitudes behaves more realistic for objects with low S/N which is why we prefer the Pogson $\log_{10}$ system for the BOSS CMASS galaxies in Chapter 6.

E.2 SDSS-III BOSS CMASS Galaxy Sample

In chapter 6 we work with data from the BOSS CMASS galaxy sample also discussed in Section 3.2. The information given here can be also accessed in the SDSS DR10 Schema

[3] `https://www.sdss3.org/dr10/algorithms/magnitudes.php#asinh`

Browser[4]. Query requirements for the selection of the LOWZ and CMASS samples are given also on the SDSS-III website[5] and in Eisenstein et al. (2011), where we refer to for more details on the target selection. Again, magnitudes downloaded from the SDSS database are given in units of *luptitudes*, fluxes in *nanomaggies*.

```
select p.ObjID, s.SpecObjID, s.RA, s.Dec, s.z as zSpec, s.zErr as zSpecErr,
       s.ModelMag_<fil>, s.ModelMagErr_<fil>,
       s.CModelMag_<fil>, s.CModelMagErr_<fil>,
       s.PetroMag_<fil>, s.PetroMagErr_<fil>,
       s.Extinction_<fil>, s.Dered_<fil>,
       p.ModelFlux_<fil>, p.ModelFluxIvar_<fil>,
       p.CModelFlux_<fil>, p.CModelFluxIvar_<fil>,
       p.PetroFlux_<fil>, p.PetroFluxIvar_<fil>,
       pz.z as zPhot, pz.zErr as zPhotErr,
       pzrf.z as zPhotRF, pzrf.zErr as zPhotRFErr,
       p.FracDev_r, a.BOSS_Target1, s.Tile
from specPhotoall s, spa a, Photoobjall p, Photoz pz, PhotozRF pzrf
where s.SpecObjID=a.SpecObjID and s.ObjID=p.ObjID and pz.ObjID=p.ObjID and
      pzrf.ObjID=p.ObjID and s.zWarning=0 and s.z>0 and
      (a.BOSS_TARGET1 & 2>0 or (a.BOSS_TARGET1 & 1>0  and
      s.Tile>=10324)) and  a.BOSSPrimary=1 and a.zWarning_NoQSO=0 and
      a.Chunk!='BOSS1' and a.Chunk!='BOSS2' and p.Fiber2Mag_i<21.5 and
      p.Clean=1 and
      (p.CalibStatus_u & 1)!=0 and (p.CalibStatus_g & 1)!=0 and
      (p.CalibStatus_r & 1)!=0 and (p.CalibStatus_i & 1)!=0 and
      (p.CalibStatus_z & 1)!=0
```

Many of the value names in the code above are also in the SDSS-II LRG sample query (App. E.1) and can be looked up there. The rest is explained in the following:

CModelMag, CModelFlux A composite model magnitude/flux derived from the best fitting combination of a exponential and de Vaucouleurs surface brightness profile (cf. `ModelMag` in App. E.1).

***Flux_Ivar** Inverse variance of the flux.

z from Photoz Photometric redshift from a k-d tree nearest neighbor fit (KF). Results are presented in Section 6.3.2 and color predictions in Section 6.3.3.

z from PhotozRF The photometric redshifts based on a random forest code (RF, Carliles et al., 2010). The photo-zs are shown in Section 6.3.2.

BOSS_TARGET1 BOSS target flag that ensures the galaxy is either from the LOWZ (`1`) or CMASS (`2`) sample.

Tile>=10324 Since in the beginning of the BOSS survey LOWZ galaxies were targeted incorrectly, one must select only the higher tiles to get a uniformly targeted sample.

[4] `http://skyserver.sdss3.org/dr10/en/help/browser/browser.aspx`
[5] `http://www.sdss3.org/dr10/algorithms/boss_galaxy_ts.php`

BOSSPrimary Set to `1` gives the best version of the spectrum.

Chunk The `BOSS_TARGET1` flag differs for chunks `boss1` and `boss2`, which are therefore excluded.

Fiber2Mag_i The magnitude in a fiber radius of 2 arcseconds. The cut in i_{fib2} differs between BOSS chunks and is therefore set below the current limit of 21.5.

Clean `clean=1` ensures a "clean" photometry. The flag removes multiple detections on different frames. Also, objects with deblending problems are removed, as well as objects where more than 20% of the PSF flux is interpolated over (that is, only 80% of the flux is actually detected). Furthermore, objects have to be detected in the first pass (unbinned image), are not saturated, and a radial profile could be constructed.

CalibStatus Ensures that one gets an actual photometric observation[6].

As mentioned earlier, the SDSS provides magnitudes only in the asinh system, for which errors seem to be underestimated for low S/N. Therefore, we download also the fluxes and convert them into $\log_{10}$ magnitudes with

$$m = 22.52.5 \log_{10}(f). \tag{E.2}$$

Fluxes themselves are given in *nanomaggies*[7], where one *maggy* is the flux relative to the zero point of the magnitude scale. Magnitudes in the SDSS are given in the AB system, wherefore one nanomaggy corresponds to $3.631 \cdot 10^{-6}$ Jy.

[6] `https://www.sdss3.org/dr10/algorithms/bitmask_calib_status.php`
[7] `https://www.sdss3.org/dr10/algorithms/magnitudes.php#nmgy`

Appendix F

Creation of SSP and CSP Models

In the predominant part of this thesis we analyze, or work with model spectra created from sets of SSPs by stellar population synthesis. We detailed the isochrone synthesis technique in Section 4.2 and give here simplified versions of the codes used to create the models. These are written such that they enable the reader to extend them easily to her/his preferences.

F.1 Generation of BC03 Models

We created BC03 models (Bruzual & Charlot, 2003) in many contexts in this work (Secs. 4.2.2, 5.1, 6.2 and Chap. 8). This can be done with the FORTRAN code `GALAXEV` which is public and can be downloaded at the websites of BC[1]. We present here the C shell (csh or tcsh) scripts used to control `GALAXEV` which are based on the codes written by Niv Drory, but have been heavily modified. They are exemplary but can be easily changed to reflect the specific setup. For the installation of `GALAXEV` and the setup we refer the reader to the `GALAXEV` documentation. In the case of exponentially increasing SFRs, one has to write the SFR as a function of time in a two-column file with t and $\psi(t)$, normalized to $1\,M_\odot$ at $t = 0$. In this way any functional form of the SFR can be provided.

```
# set up to your environment
setenv bc03 <path to galaxev>/src
source $bc03/.bc_cshrc

# define IMF
setenv IMF chabrier
setenv IMFshort chab

# set up model parameters
setenv MET 'm42 m52 m62 m72' # metallicity (see doc)
setenv SFH 1                 # exponentially decreasing SFH
setenv SFH2 6                # user defined SFH
setenv TAU '0.5 1.0'         # tau values in Gyrs
setenv RES 'lr'              # low resolution SEDs
```

[1] `http://www.cida.ve/~bruzual/bc2003` and `http://www.iap.fr/~charlot/bc2003`

```
setenv AGES '0.10,10.0'      # an array of ages in Gyrs (max. 24)

# if bc03 models are only in ASCII format convert them
# with bin_ised routine to binary needed by csp_galaxev
pushd $bc03
./build_filterbin
popd
pushd $bc03/../models/Padova1994/$IMF
foreach i (*_${RES}_*ASCII)
  $bc03/bin_ised $i
end
popd

# create SSPs with the csp_galaxev routine
foreach Z ($MET)
  csp <<END
../models/Padova1994/${IMF}/bc2003_${RES}_${Z}_${IMFshort}_ssp.ised
N
0
bc2003.${RES}.${Z}.${IMFshort}.sspSFH.ised
END

# create CSPs with exponentially decreasing SFR
# with the csp_galaxev routine
foreach Z ($MET)
  foreach tau ($TAU)
      csp <<END
../models/Padova1994/${IMF}/bc2003_${RES}_${Z}_${IMFshort}_ssp.ised
N
$SFH
$tau
N
20.0
bc2003.${RES}.${Z}.${IMFshort}.decSFH.tau${tau}.ised
END
  end
end

# create CSPs with exponentially increasing SFR
# with the csp_galaxev routine
foreach Z ($MET)
  foreach tau ($TAU)
      csp <<END
../models/Padova1994/${IMF}/bc2003_${RES}_${Z}_${IMFshort}_ssp.ised
N
$SFH2
<file with SFR(t)>
```

```
bc2003.${RES}.${Z}.${IMFshort}.incSFH.tau${tau}.ised
END
  end
end

# extract spectra at the specified ages with gpl routine
foreach i (*ised)
  gpl <<EOF
$i
91,30000
$AGES
`basename $i .ised`.sed
EOF
end
```

F.2 Generation of Maraston CSPs

In Appendices C and E.1 we show the SED fitting results using the extensive model sets of Maraston & Strömbäck (2011, M11). They are provided only as SSPs, wherefore we have to create CSPs in order to be able to make a meaningful comparison with the CSP models of BC03. Furthermore, in Sections 4.2.2 and 6.1.1 we also investigate the colors of the model by Maraston et al. (2009, M09) passively evolving with redshift. To create the M11 CSPs as well as the SEDs of the passively evolving M09 model we use the public stellar population synthesis code `EzGal` (Mancone & Gonzalez, 2012) written in python. We refer to Mancone & Gonzalez (2012), the website[2], and the documentation for details on the `EzGal` code. In the following we give only a brief introduction to `EzGal` and the code used to control it and to create the M11 CSP models, as well as the passively evolving M09 model.
As inputs, the `EzGal` software needs a file where the SSP model SEDs are given as a function of age. This can be an ASCII file, but `EzGal` can deal with different file types as well. In the case of ASCII, the first column of the file has to include the wavelength steps of the SSP SEDs, whereas the following columns are the fluxes with the corresponding age values in the first row. The entry in the first column of the first row is a placeholder which can hold any value. Given such a model file, one can create a CSP following any arbitrary user defined SFH. In the next paragraph we show a skeletal example on how to create a passively evolving SSP and CSPs with exponentially decreasing SFRs, both sampled at two ages. The code is written in python with dependency on the `numpy` package.

```
import numpy as np
import ezgal

modfile = '<model filename>'      # string, SSP model file in EzGal format
taulist = np.array([0.5, 1.0])    # array of taus in Gyrs for an exp. SFH
agelist = np.array([0.01, 12.0]) # array of ages in Gyrs to be sampled at

# function returning an exponential SFH
```

[2] downloadable at `http://www.baryons.org/ezgal/`

```
def exponential_sfh(t, tau):
  return np.exp(-1.*t/tau )

# MAIN function creating a passively evolving SSP and CSPs
def mkmodel(modfile, taulist, agelist):

  # set up EzGal model
  ezmodel = ezgal.model(modfile)
  # get model wavelengths
  ll = ezmodel.ls

  # set up SED arrays for SSP and CSP
  sedssp = np.zeros((ll.shape[0], agelist.shape[0]+1), dtype='float')
  sedssp[:,0] = ll
  sedcsp = np.copy(sedssp)

  # MAKE SSP FILES at given ages (passive evolution)
  # loop through ages and get respective SSP SED
  for i in range(agelist.shape[0]):
    sedssp[:,i+1] = ezmodel.get_sed(agelist[i],age_units='gyrs',
                                    units='fl')[::-1]
  # write to SSP file with ages from agelist
  np.savetxt('<SSP output filename>',
             sedssp,
             fmt='%5.1f'+(agelist.shape[0])*'\t%1.3e',
             header='ages = '+str(agelist)+' Gyr\nlambda[A] fluxes(age)')

  # MAKE CSP FILES at given ages
  # loop through tau values in taulist
  for tau in taulist:
    cspmodel = ezmodel.make_csp(exponential_sfh, args=(tau,)) # create CSP
    # the age array for an EzGal model is in units of years - convert to Gyrs
    cspmod_ages = ezgal.utils.convert_time(cspmodel.ages, incoming='years',
                                           outgoing='gyrs')
    # loop through ages and get respective CSP SED
    for i in range(agelist.shape[0]):
      sedcsp[:,i+1] = cspmodel.get_sed(agelist[i], age_units='gyrs',
                                       units='fl')[::-1]
    # write to CSP file with ages from agelist
    np.savetxt('<CSP output filename>.tau'+str(tau),
               sed,
               fmt='%5.1f'+(agelist.shape[0])*'\t%1.3e',
               header='ages = '+str(agelist)+' Gyr\nlambda[A] fluxes(age)')

# EXECUTE MAIN FUNCTION mkmodel
mkmodel(modfile, taulist, agelist)
```

Acknowledgments

First of all, I want to thank Ralf Bender for making this thesis possible and for a fast and reliable correspondence especially in the last phase of the thesis.

I want to give my special thanks to Stella Seitz, my advisor during these years, for her invaluable support and guidance. She was always enthusiastic about discussing science, resourceful in providing new ideas, and never left questions half-answered.

I want to thank my office mates and group colleagues for the numerous discussions and exchanges, and their friendship. My gratitude goes to Fabrice Brimioulle, Anna Monna, Daniel Grün, Ralf Kosyra, Mihael Kodrić, Julia Young, Markus Michael Rau, Oliver Friedrich, Thomas Eichner, Annalisa Mana, and Patricia Spinelli. A special thanks goes to Fabrice Brimioulle, Rebekka Grellmann, Daniel Grün and Julia Young, but also Anna Monna, Markus Michael Rau, and Oliver Friedrich who helped improve this thesis by their valuable comments and suggestions.

I greatly valued the many suggestions and inputs I received from Roberto Saglia and Niv Drory and I would like to thank them for that as well as for their interest in and support for my work. I would like to express my gratitude furthermore to Jan Snigula, who always made time for helping me with the `PhotoZ` code.

Furthermore, I would like to thank also Tadziu Hoffmann and Keith Butler for their help with technical problems.

There are many people at USM and MPE that created the friendly and enjoyable work atmosphere and I want to thank all of them for that. However, I would like to mention some people that made it particularly so and became friends. Therefore, my thanks goes to Anna Monna, Rebekka Grellmann, Fabrice Brimioulle, Christian Obermeier, Mihael Kodrić, Ralf Kosyra, Max Brunner, Julia Young, Daniel Grün, Markus Michael Rau, Oliver Friedrich, Thomas Eichner, Stephanie Pekruhl, Wolfgang von Glasow, Matthias Bierschenk, Carsten Strübig, Yi-Hao Chen, Chien-Hsiu Lee, and many more. I hope we don't lose sight of each other.

Finally, but in no way in the least, I owe a huge thanks to Marcel Lepoittevin for his support. I would like to thank him for his encouragement, the time he spent on listening to my ideas, and that he provided me with valuable input concerning coding and nutrition. Furthermore, I would like to thank my parents for their support, not only during the past few years, but throughout all my life. I thank them for believing in me, and for being there whenever I need them.

Funding

This work was supported by SFB-Transregio 33 (TR33) "The Dark Universe" by the Deutsche Forschungsgemeinschaft (DFG) and the DFG cluster of excellence "Origin and Structure of the Universe".

SDSS-I/II

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is `http://www.sdss.org/`.
The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

SDSS-III

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is `http://www.sdss3.org/`.
SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofisica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

Dark Energy Survey

Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of

Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, Financiadora de Estudos e Projetos, Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Cientfico e Tecnológico and the Ministério da Ciência e Tecnologia, the Deutsche Forschungsgemeinschaft and the Collaborating Institutions in the Dark Energy Survey.
The Collaborating Institutions are Argonne National Laboratories, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the Eidgenössische Technische Hochschule (ETH) Zürich, Fermi National Accelerator Laboratory, the University of Edinburgh, the University of Illinois at Urbana-Champaign, the Institut de Ciencies de l'Espai (IEEC/CSIC), the Institut de Fisica d'Altes Energies, the Lawrence Berkeley National Laboratory, the Ludwig-Maximilians Universität and the associated Excellence Cluster Universe, the University of Michigan, the National Optical Astronomy Observatory, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

CLASH

The CLASH Multi-Cycle Treasury Program (GO-12065) is based on observations made with the NASA/ESA Hubble Space Telescope. The Space Telescope Science Institute is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. The Dark Cosmology Centre is funded by the DNRF.