Galaxy Cluster Identification With eROSITA in the Era of Big Data

Jacob Ider Chitham

München 2021

Galaxy Cluster Identification With eROSITA in the Era of Big Data

Jacob Ider Chitham

Dissertation
an der Fakultät der Physik
der Ludwig–Maximilians–Universität
München

vorgelegt von
Jacob Ider Chitham
aus Barcelona, Spain

München, den 02/07/2021

Erstgutachter: Prof. Dr. Kirpal Nandra
Zweitgutachter: Prof. Dr. Jochen Weller
Tag der mündlichen Prüfung: 15/09/2021

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

Thanks to my mum and Chris for all your love and support. I cannot wait to give you a proper hug after this crazy Brexit/Covid era. Thanks to Sebo, Ricky, Tosca, Oxana, Fruzsi, Damien, Felix and Adam for keeping me sane over these past few years. I will be forever grateful for all the lunchtime chats, birthday surprises, Oktoberfests, jam sessions, Boulderwelt puzzles and fun times during lockdown...without you guys there is no way that this would have been possible. Thanks to Will, Ollie, Ollie, Matt, Tom, Tom, Thom, Gea, Yousuf, Nicciy, Griff, Seward, Calum and Corey for all the international experiences we've shared while I've been away. You always reminded me of who I am and you were always down for impulsive adventures when I needed them the most. Thanks to Alexis for all the good times in Helsinki and Alpine cosmology hikes. Thanks also to Matthias for taking me under your wing, thanks to Eli for all the support and thanks to Nicolas, Esra, Andrea, Paul and Mara for always looking out for me. And last but definitely not least, a huge thank you to Johan for showing me the ways of the warrior! You are definitely the silent hero behind this PhD and if it wasn't for your patience and all the time you invested in me, I wouldn't be half the Physicist I am today. Thank you for believing in me.

Zusammenfassung

Von den vielen kosmologischen Experimenten, die dazu beitragen können, die mysteriöse Natur der dunklen Materie und der dunklen Energie besser zu verstehen, konzentriert sich diese Arbeit auf einen einzigen kosmologischen Test. Dieser Test ist die Häufigkeit von Galaxienhaufen. Das Ziel ist die Messung einer Untergruppe von kosmologischen Parametern (Ω_{m_0} und σ_8) unter Verwendung der größten röntgenselektierten Stichprobe von spektroskopisch bestätigten Galaxienhaufen, die jemals untersucht wurde. Dies geschieht unter der Annahme einer räumlich flachen Geometrie des Universums und der kosmologischen Konstante. Diese Arbeit ist durch die nächste Generation von spektroskopischen Himmelsdurchmusterungen motiviert, die Untersuchung von Galaxienhaufen mit *eROSITA* ermöglicht. Dies basiert auf der Grundlage der Analyse des gesamten Röntgenkatalogs des Himmels. In Vorbereitung auf das *eROSITA* Experiment zur Kosmologie mit Galaxienhaufen stellt diese Arbeit eine Vielzahl von theoretischen Methoden und praktischen Werkzeugen vor. Diese können zur Identifizierung und Extraktion von kosmologischen Informationen aus Stichproben von röntgenselektierten Galaxienhaufen im Zeitalter von Big Data angewendet werden. Diese Methoden beinhalten die präzise und unverzerrte Messung der Rotverschiebung von Galaxienhaufen über große kosmologische Distanzen. Dieser Prozess ist speziell für die Verwendung von inhomogener photometrischer und spektroskopischer Datensätze und der Charakterisierung der erforderlichen optischen Selektionsfunktion entwickelt. Des Weiteren wird der Rahmen zur Messung kosmologischer Parameter unter Verwendung einer Vielzahl von Eigenschaften der Galaxienhaufen erstellt.

Abstract

Of the many cosmological experiments that can be used to help shed light on the mysterious nature of dark matter and dark energy, this thesis focuses on a single cosmological probe. This probe is the abundance of clusters of galaxies. The end goal is to measure a subset of cosmological parameters (Ω_{m_0} and σ_8) using the largest X-ray selected sample of spectroscopically confirmed clusters ever constructed. This assumes a spatially flat geometry and cosmological constant. This work is motivated by the next generation of spectroscopic surveys that will facilitate *eROSITA* cluster science based on the analysis of its all sky X-ray catalogue. In an effort to prepare for the *eROSITA* cluster cosmology experiment, this thesis presents a variety of theoretical methods and practical tools that can be applied to identify and extract cosmological information from samples of X-ray selected clusters in the era of Big Data. This includes the precise and unbiased measurement of cluster redshifts over large cosmological volumes using inhomogeneous photometric and spectroscopic datasets, the characterisation of the optical selection function and the framework to measure cosmological parameters using various cluster mass-proxies.

Kapitel 1

Introduction: Cosmology

This chapter introduces the scientific problems regarding dark matter and dark energy as well as the fundamental cosmological concepts used throughout this thesis relating to cluster cosmology. This includes redshift, cosmological volumes, contributions to the energy density of the universe, the variance of the matter density field and the halo mass function. The goals of this thesis are then summarised in terms of the subsequent chapters within the context of the *eROSITA* cluster cosmology experiment.

1.1 The Dark Universe

Although dark components make up almost all of the energy density of the Universe, very little is known about their physical origin. The formation of large scale structure is dependent on the presence of dark matter and dark energy and their mysterious nature constitutes one of the most outstanding puzzles in all of physics. This section provides a brief historical overview of the dark matter and dark energy. For a summary of the most cutting-edge theories of gravity, general relativity and cosmology please refer to the reviews of Will (2014); Ishak (2019).

1.1.1 Dark Matter

Dark matter was first postulated as the missing component of the total mass inferred from the dynamics of the galaxies in the Coma (Zwicky, 1933) and Virgo (Smith, 1936) galaxy clusters. The rotation curves of spiral galaxies then contributed significantly to evidence for dark matter. The first case was the Andromeda galaxy (Babcock, 1939), where very high rotational velocities were observed for the stars and gas at large radii from the galactic centre. These observations were then repeated with higher quality data and extended to other nearby galaxies (e.g. Rubin & Ford, 1970; Roberts & Rots, 1973; Rubin et al., 1980) which indicated that for observed rotation profiles to be physically viable, there must be a dark component that constitutes most of their mass. Further evidence was provided by dynamical mass estimates of galaxies (Ostriker & Peebles, 1973;

Ostriker et al., 1974; Einasto et al., 1974) which indicated that galaxies form within weakly interacting (cold) dark matter (CDM) halos (Peebles, 1982; Blumenthal et al., 1984; Davis et al., 1985). The luminous arcs observed due to the strong gravitational lensing of distant background sources by massive clusters in the local universe (e.g. Soucail et al., 1988; Lynds & Petrosian, 1989) led to conclusions that dark matter could be distributed differently from the baryonic matter in clusters (e.g. Bergmann et al., 1990). This was later strongly reinforced by the detection of an extreme merging system with spatially segregated weak lensing and X-ray mass profiles. This "bullet cluster" (Clowe et al., 2006; Randall et al., 2008), shown in Fig. 1.1, proved that dark matter and baryonic matter were physically separated components of the total cluster mass. Further observations of galaxy clusters as well as simulations of large-scale structure made it possible to estimate that dark matter accounted for approximately a quarter of the critical energy density of the universe (Bahcall & Cen, 1992; Bahcall et al., 1995) i.e. about five times more mass resides in non-baryonic matter relative to baryonic matter. Later computational advances enabled scientists to study the evolution of dark matter with gravity from the early universe using N-body simulations (Kravtsov et al., 1997; Navarro et al., 1997; Springel, 2005; Klypin et al., 2011). The origin and composition of dark matter is still uncertain, although some theories suggest that its presence can be explained by modifications to the theory of gravity (e.g. Capozziello & de Laurentis, 2011; Famaey & McGaugh, 2012), exotic particles (e.g. Bertone et al., 2005; Arkani-Hamed et al., 2009; Feng, 2010) such as axions (e.g. Rosenberg, 2015; Marsh, 2016; Dekker et al., 2021), supersymmetric (e.g. Jungman et al., 1996), weakly-interacting massive particles (e.g. Pospelov et al., 2008) or primordial black holes (e.g. Carr et al., 2016; Bird et al., 2016). For a detailed review of the history and theory behind dark-matter please refer to Trimble (1987); Turner (1999); Bahcall (2015); Bertone & Hooper (2018).

1.1.2 Dark Energy

After the equations describing the field theory of General relativity were formulated, they were modified to include a cosmological constant (Λ, Einstein, 1917) in an attempt to provide a solution for a static universe (for a detailed historical account refer to Weinberg, 1989). For decades, the widely accepted cosmological model of the Universe was that it is matter-dominated (Einstein & de Sitter, 1932) and that any contribution from a cosmological constant was negligible. As a result, its expansion rate should decelerate with time due to the effect of gravitational forces. However, luminosity measurements of distant Type Ia supernovae (SNeIa) provided the first observational evidence that the expansion of the Universe was actually accelerating with a positive cosmological constant (Riess et al., 1998; Perlmutter et al., 1999). These iconic SNeIa constraints are shown in Fig. 1.2 along with those from a more recent analysis (Scolnic et al., 2018) to demonstrate twenty years of progress in the field. The evidence for cosmic acceleration has since been supported via complementary cosmological probes such as the cosmic microwave background (CMB, Planck Collaboration et al., 2018), Cepheid variables (Riess et al., 2018; Verde et al., 2019) baryonic acoustic oscillations (Eisenstein et al., 2005; Alam et al., 2017; Ata et al., 2018;

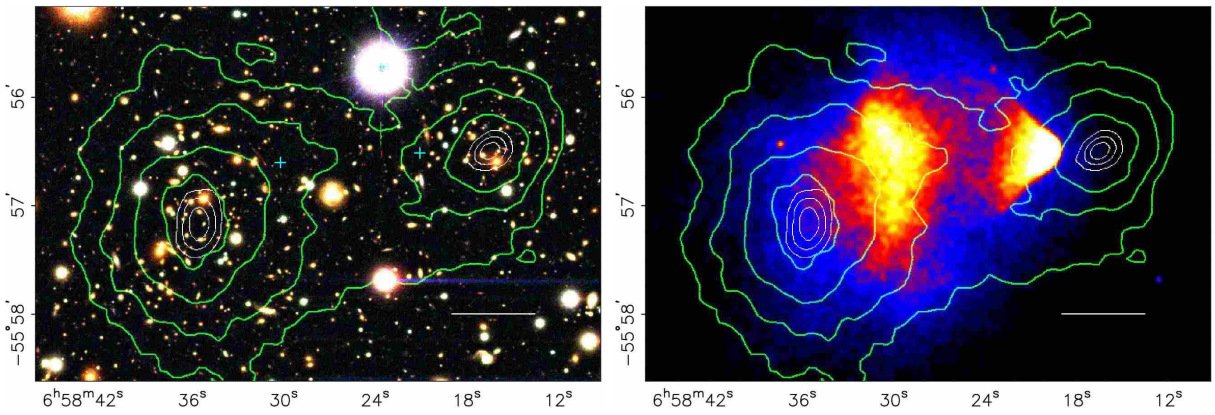

Abbildung 1.1: Adapted from Clowe et al. (2006). X-ray, optical and weak lensing measurements of the iconic "Bullet Cluster". The extreme discrepancy between the distribution of baryonic matter relative to the total mass profile provides direct evidence for dark matter. The white bar represents 200 kpc. Green contours are the weak lensing reconstruction and the white contours represent the positions of the profile peaks (68.3%, 95.5%, and 99.7% confidence intervals). **Left panel**: A colour image of the merging cluster 1E0657-558. at the distance of the cluster produced with Magellan optical data. **Right panel:** a 500 ks Chandra X-ray image of the same cluster.

eBOSS Collaboration et al., 2020), redshift space distortions (Pezzotta et al., 2017; Zarrouk et al., 2018) and cosmic shear (Abazajian & Dodelson, 2003; Abbott et al., 2018; Jullo et al., 2019). Remarkably, for the universe to be accelerating, cosmological models require an additional energy component that accounts for approximately 70% of the total energy density of the Universe and behaves as though it exerts negative pressure. This missing component has been termed "dark energy" or "quintessence" (e.g. Zimdahl et al., 2001; Chimento et al., 2003). For a review, please refer to Carroll et al. (1992); Peebles & Ratra (2003); Copeland et al. (2006); Frieman et al. (2008); Weinberg et al. (2013); Frusciante & Perenon (2020). Although many models for dark energy exist, only a subset are mentioned in this thesis where the dark energy equation of state parameter is either $w = -1$, $w \neq -1$ or $w \equiv w(t)$. These three scenarios correspond to 1) a cosmological constant (ΛCDM) with a uniform apparent energy density in time and space (Zel'dovich, 1968). 2) $w \neq -1$ in a time-in-dependent fashion (w_0CDM). 3) a dynamical fluid that exhibits time evolution (wCDM) (Pillepich et al., 2018).

1.2 The concordance cosmological model

1.2.1 The homogeneous universe

The foundation of the concordance cosmological model is the *cosmological principle*. i.e. the principle that the universe is both isotropic and homogeneous on large scales ($\gtrsim$ 100Mpc). This is expressed using the Robertson-Walker metric

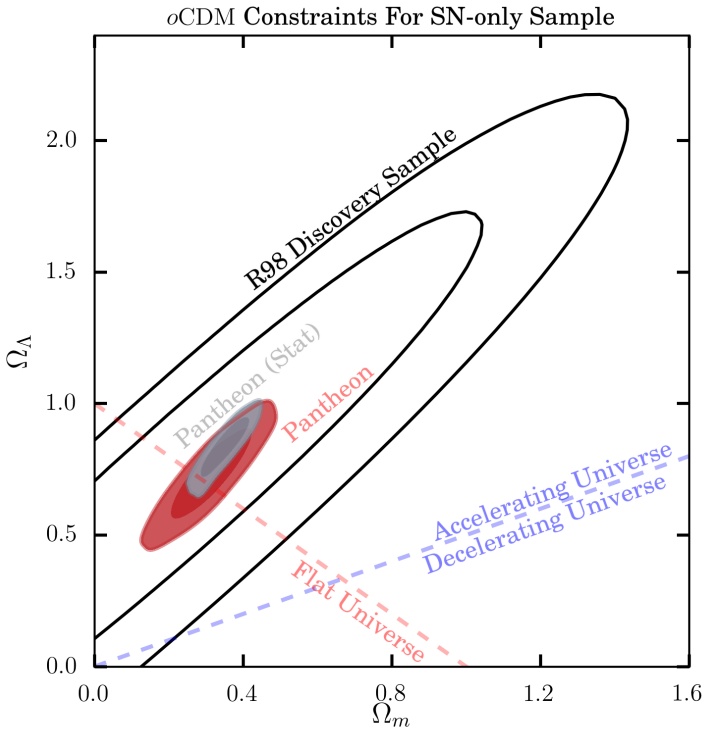

Abbildung 1.2: Evidence for dark energy and the accelerating expansion of the universe from analyses of Type Ia Supernovae. Adapted from Scolnic et al. (2018). Cosmological results assume a oCDM model ($w = -1$, Ω_k varies). Contours represent the 68% and 95% confidence regions in the $\Omega_m - \Omega_\Lambda$ plane. Black contours are derived from the analysis of the R98 Discovery Sample (i.e., the first observational evidence for dark energy based on 16 high-redshift SNeIa, Riess et al., 1998), red contours show results from the more recent Pantheon sample (365 spectroscopically confirmed SNeIa, Scolnic et al., 2018) when accounting for both systematic uncertainties as well as statistical uncertainties (grey contours). The dashed purple line represents the solution for a deceleration parameter of value zero, $q_0 \equiv \Omega_m/2 - \Omega_\Lambda = 0$. This divides the $\Omega_m - \Omega_\Lambda$ plane into regions with an accelerating or decelerating expansion at the present time. The dashed red black line corresponds to solutions with a flat universe with no additional contributions to the energy density $\Omega_m + \Omega_\Lambda = 1$ (i.e., Eq. 1.9 with $k = \Omega_k = 0$ and $i \in \{m\}$).

$$ds^2 = -dt^2 + a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2\left(d\theta^2 + \sin^2\theta\, d\phi^2\right)\right]. \tag{1.1}$$

Here, k describes the curvature of space (the discrete values $k = \{0, +1, -1\}$ correspond to flat, closed and open geometries respectively), t is time, (r, θ, ϕ) are polar coordinates and $a(t)$ is a dimensionless scale-factor which relates *comoving*[1] and *physical* coordinate systems.

[1]The comoving distance between two sources moving with the Hubble flow is constant, the phy-

Cosmological Redshift

Electromagnetic radiation emitted from a source receding with the Hubble flow at time t_e (in the past) is *redshifted* by the time it reaches an observer at t_0 (the present epoch). This redshift is due to the conservation of energy of the emitted photon, whereby its frequency decreases (becomes redder) as its wavelength increases (due to the Hubble expansion) as the speed of light is finite and constant. This process defines the *cosmological* redshift[2].

$$z \equiv \frac{\nu_{\rm e}}{\nu_{\rm o}} - 1 = \frac{\lambda_{\rm o}}{\lambda_{\rm e}} - 1. \tag{1.2}$$

Here, ν and λ represent frequency and wavelength, and subscripts e and o designate emitted and observed quantities. For example, the 4000Å Balmer break present in the spectrum of a passively evolving galaxy at a redshift $z = 0$ (the present epoch) would be redshifted to 8000Å if the same spectral feature originated from a source at $z = 1$ (7.715 Gyr years ago). This cosmological redshift can also conveniently be expressed in terms of the aforementioned scale factor

$$1 + z = \frac{a(t_0)}{a(t_{\rm e})} = \frac{1}{a(t_{\rm e})}. \tag{1.3}$$

Cosmological Expansion

The expansion of the Universe was first discovered by measuring the distance to galaxies in the local universe which indicated that more distant galaxies had larger recession velocities (Hubble, 1929). The constant of proportionality between recession velocity of an extragalactic sources, v, and their distance from the observer, D, is an approximation for the local expansion rate. This defines the Hubble constant as

$$H_0 \equiv \ 100h\ {\rm km\,s^{-1}\,Mpc^{-1}} \approx \frac{v}{D}\,. \tag{1.4}$$

Here, h is simply the dimensionless equivalent of H_0. Observationally, distance-ladder estimates (e.g. 0.732 ± 0.017, Riess et al., 2016) tend to favour higher values of h relative to those derived from measurements of the CMB (e.g. 0.674 ± 0.005, Planck Collaboration et al., 2020).

A more general way to describe the expansion of the Universe on larger scales is to include time evolution i.e., via the Hubble parameter (Eqs. 1.5 and 1.11). This equates to the Hubble constant when evaluated at the present epoch, $H(t = t_0 = 0) = H_0$.

$$H^2(t) \equiv \left(\frac{\dot{a(t)}}{a(t)}\right)^2 = \frac{8\pi}{3}\,\rho(t) - \frac{k}{a^2(t)} \qquad \text{Friedmann equation.} \tag{1.5}$$

$$\dot{\rho(t)} + 3H(t)\,[\rho(t) + p(t)] = 0 \qquad \text{Fluid equation.} \tag{1.6}$$

sical distance between sources is a function of time. This time evolution defines the scale factor by physical distance $= a(t) \times$ comoving distance.

[2]The observed redshift $\tilde{z}$ differs from the cosmological redshift, z, by the additional contribution to an object's radial motion from its peculiar velocity (the local velocity relative to the Hubble flow).

Here, $\rho(t)$ is the energy density, $p(t)$ is pressure and a dot over a variable indicates its derivative with respect to time. In the case of a flat universe, ($k = 0$) and Eq. 1.5 reduces to an expression for the *critical* density

$$\rho_{\rm c}(t) = \frac{3H^2}{8\pi}\,. \tag{1.7}$$

It is often convenient to express the energy density in units of $\rho_{\rm c}$ to define the *total* density parameter

$$\Omega_{\rm tot} \equiv \frac{\rho}{\rho_{\rm c}}, \tag{1.8}$$

which then allows us one to rewrite Eq. 1.5 in terms of its various components

$$\sum_i \Omega_i + \Omega_\Lambda - 1 = \frac{k}{a^2H^2}. \tag{1.9}$$

Here, the sum is over contributions from matter - Ω_m (baryons and cold dark matter), radiation - Ω_γ (photons), relativistic particles - Ω_ν (neutrinos) and a cosmological constant - Ω_Λ (dark energy without time evolution). For simplicity, any subsequent equations in this thesis will only consider the dominant contributions to the total density parameter in the case of a flat universe. It is reasonable to assume the contribution due to radiation is negligible at redshift zero as it is dominated by the CMB (Fixsen et al., 1996; Mather et al., 1999; Fixsen, 2009). This[3] corresponds to an energy density of

$$\Omega_\gamma = 2.47 \times 10^{-5} h^{-2} << \Omega_i \qquad\qquad \forall i \in \{m, \Lambda, \nu\}. \tag{1.10}$$

Here, the magnitude the relevant parameters are $h \approx 0.7$, $\Omega_m \approx 0.3$, $\Omega_m \approx 0.7$, $\Omega_\nu < 0.003$ (Riess et al., 2016; Planck Collaboration et al., 2020; Lahav & Liddle, 2019; Particle Data Group et al., 2020).

Cosmological Distances

The Hubble parameter introduced in Eq. 1.5 can equivalently be expressed in terms of a redshift dependency, which depends on the relevant components of the energy density found in Eq. 1.9 as

$$E(z) \equiv \frac{H(z)}{H_0} = \sqrt{\Omega_\gamma\,(1+z)^4 + \Omega_{\rm M}\,(1+z)^3 + \Omega_k\,(1+z)^2 + \Omega_\Lambda} \tag{1.11}$$

Integrating this expression (Eq. 1.11) provides the *line-of-sight comoving distance*. This is defined to be constant for two objects moving with the Hubble flow.

$$D_{\rm C} = D_{\rm H} \int_0^z \frac{dz'}{E(z')}, \tag{1.12}$$

[3]The current best estimate is 2.7255 ± 0.0006K (Fixsen, 2009) which when transformed into an energy density gives Eq. 1.10.

where, $D_{\rm H}$ is the Hubble distance

$$D_{\rm H} \equiv \frac{c}{H_0} = 3000\,h^{-1}\ {\rm Mpc} = 9.26\times10^{25}\,h^{-1}\ {\rm m}. \tag{1.13}$$

In the case of a flat universe, the line-of-sight comoving distance is equivalent to the *transverse-comoving distance*

$$D_{\rm M} = D_{\rm C}(k=0). \tag{1.14}$$

The *angular diameter distance*, $D_{\rm A}$, is the ratio of the physical transverse size of an object with respect to its angular size on the sky. This relates to the transverse comoving distance via the redshift of the object

$$D_{\rm A} = \frac{D_{\rm M}}{1+z} \tag{1.15}$$

The *comoving volume*, $V_{\rm C}$, is the volume is the three-dimensional analogue of the comoving distance with differential element equivalent to

$$dV_{\rm C} = D_{\rm H}\,\frac{(1+z)^2\,D_{\rm A}^2}{E(z)}\,d\Omega\,dz. \tag{1.16}$$

Here, $d\Omega$ is a solid-angle element (not to be confused with the density parameter). Integrating this comoving volume element from the current day out to a redshift z gives the total comoving volume. All cosmological volumes, V, mentioned in this thesis refer to the comoving volume, $V_{\rm C}$, unless stated otherwise.

For a sense of scale, the comoving number density of galaxies and clusters over the redshift range $z \in [0,1)$ is $1.58\times10^{-7}{\rm Mpc}^{-3}$ and $3.31\times10^{-7}{\rm Mpc}^{-3}$ respectively. These estimates assume a halo mass of $M_{200} \sim 0.66-3.0\times10^{12}{\rm M}_\odot$ for galaxies and $>10^{14}{\rm M}_\odot$ for clusters when integrating Eq. 1.26 with a Tinker et al. (2008) halo mass function (Eq. 1.27) and standard ΛCDM cosmology using the Core Cosmology Library[4] (Chisari et al., 2019). The Hubble volume equates to $(c/H_0)^3 \sim 8.96\times10^{10}{\rm Mpc}^3$, while the mean number density of baryons in the universe is $\sim 0.25{\rm m}^{-3}$ (Planck Collaboration et al., 2020). For a full review of cosmological distances please refer to Hogg (1999).

1.2.2 The in-homogeneous universe

The cosmological principle is an approximation which holds on large scales ($\gtrsim$ 100Mpc), however on smaller scales, the universe is neither isotropic, nor homogeneous. This is apparent from the filamentary distribution of galaxies which make up the cosmic web. This large scale structure is a result of primordial density fluctuations in the early universe which have evolved with the expansion of the universe. This subsection introduces the equations related to large scale structure relevant to galaxy cluster cosmology.

[4]This is based on the Cosmic Linear Anisotropy Solving System (Blas et al., 2011)

Density Fluctuations

Fluctuations in the matter density field $\rho(\boldsymbol{x})$ are described by the density contrast

$$\delta(\boldsymbol{x}) \equiv \frac{\rho(\boldsymbol{x}) - \langle\rho\rangle}{\langle\rho\rangle}, \tag{1.17}$$

where $\langle ...\rangle$ denotes an average over volume and $\boldsymbol{x}$ is a comoving spatial vector. The Fourier transform of $\delta(\boldsymbol{x})$ and its inverse are given by

$$\delta(\boldsymbol{k}) = \int \mathrm{d}^3x\ \delta(\boldsymbol{x}) e^{i\boldsymbol{k}\cdot\boldsymbol{x}}, \qquad \delta(\boldsymbol{x}) = \frac{1}{(2\pi)^3}\int \mathrm{d}^3k\ \delta(\boldsymbol{k}) e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}. \tag{1.18}$$

where k is the wavenumber. The two point correlation function, which given the excess probability of finding two over-densities separated by a given distance $|\boldsymbol{r}|$ (averaged over all spatial coordinates) is then defined by

$$\xi(r) = \langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r})\rangle = \frac{1}{2\pi^2}\int k^3 V^{-1}|\delta(k)|^2\ \frac{\sin(kr)}{kr}\ \mathrm{d}\ln k. \tag{1.19}$$

where V is the volume that ensures $\int|\delta(\boldsymbol{x})|\mathrm{d}^3x$ is finite. This correlation function is the Fourier transform of the power spectrum

$$P(k) \equiv V^{-1}\langle|\delta(k)|^2\rangle, \tag{1.20}$$

which is often expressed in a dimensionless form

$$\Delta^2(k) \equiv k^3 P(k)/2\pi^2. \tag{1.21}$$

The variance of the smoothed matter density field σ is then given by

$$\sigma^2(R) = \int \mathrm{d}\ln k\ \Delta^2(k)|\tilde{W}(kR)|^2. \tag{1.22}$$

where $W(\tilde{k}, R)$ is the Fourier transform of the window (filter) function over which the field is smoothed. This is commonly set a top-hat filter i.e.

$$\tilde{W}(kR)_{\mathrm{tophat}} = \frac{3}{(kR)^3}[\sin(kR) - kR\cos(kR)]. \tag{1.23}$$

The amplitude of the linear power spectrum on the scale of $8h^{-1}$ Mpc is denoted as σ_8. This crucial parameter represents the growth of structure and heavily influences how density fluctuation evolve in the early universe.

The probability of $\delta(\boldsymbol{x}, R)$ in an interval of $\mathrm{d}\delta$ defines the multiplicity function, $f(\sigma)$. In the Press & Schechter (1974) formalism, this is

$$f_{\mathrm{PS}}(\sigma)\mathrm{d}\delta = \frac{1}{\sqrt{2\pi\sigma^2(R)}}\exp[-\delta^2/2\sigma^2(R)]\mathrm{d}\delta. \tag{1.24}$$

The integral of Eq. 1.24 provides the smoothing scale, $F(M)$,

$$F_{\rm PS}(M) = \int_{\delta_{\rm c}}^{\infty} f(\sigma){\rm d}\delta = \frac{1}{2}\,{\rm erfc}\left(\frac{\nu}{\sqrt{2}}\right) \tag{1.25}$$

Here, $\nu \equiv \delta_{\rm c}/\sigma(M)$ is the peak height and $\delta_{\rm c}$ is the critical over-density at which an object (halo) collapses.

The comoving number density, n of halos in the mass interval ${\rm d}M$ is then given by the halo mass function

$$\frac{{\rm d}n}{{\rm d}M}{\rm d}M = \frac{\langle\rho\rangle}{M}\left|\frac{{\rm d}F(M)}{{\rm d}M}\right|{\rm d}M. \tag{1.26}$$

while many other commonly used forms of Eq. 1.24 also exist (e.g. Bond et al., 1991; Sheth & Tormen, 1999; Sheth et al., 2001; Bhattacharya et al., 2011; Despali et al., 2016), the form adopted in this thesis is the Tinker et al. (2008) multiplicity function, which is assumed to be "universal" i.e., independent of cosmology and redshift

$$f_{\rm Ti}(\sigma){\rm d}\delta = A\left[\left(\frac{\sigma}{b}\right)^{-a} + 1\right]e^{-c/\sigma^2}{\rm d}\delta. \tag{1.27}$$

Here, A, a, b, and c are constants calibrated from simulations. Comparisons between the most commonly used halo mass function models are illustrated in Fig. 1.3 relative to the Tinker et al. (2008) model. This comparison indicates that it is sufficiently accurate for cosmological analyses in this work (please refer to Sect. 4.1 for more details).

Updates to these models are necessary for high-precision cosmological cluster experiments where the uncertainties on the theoretical model are non-negligible and cosmological dependencies must be considered. The state of the art implementation of modelling the halo mass function is via emulation (e.g. McClintock et al., 2019b; Nishimichi et al., 2019; Bocquet et al., 2020), which essentially modifies Eq. 1.27 by providing cosmological dependency to the fixed parameters. Emulators are calibrated using Gaussian Process regression (e.g. Ambikasaran et al., 2015) and suites of cosmological simulations that sparsely sample high-dimensional cosmological parameters spaces. This provides an accurate halo mass function model at the sub-percent level for group scales and $\sim 10\%$ for the most massive halos at $z = 0$. The latest analytical mass function models also incorporate a cosmological dependency using simulations by parameterising in terms of growth rate and power spectrum slope to achieve accuracy at the 2-3% level over a wide range of the cosmologies, including dynamical dark energy (Ondaro-Mallea et al., 2021). For the cosmological analyses present in this work, the Poission contribution to the uncertainty on the predicted number counts is dominant and therefore no other contributions are considered (justified in Sect. 4.1). However, in high-precision cosmological experiments with large number of clusters, the sample variance is one the main sources of statistical uncertainty in the covariance matrix. It is therefore necessary to include an analytical model to account for this effect to reduce the biases in the measured cosmological parameters (e.g. Fumagalli et al., 2021).

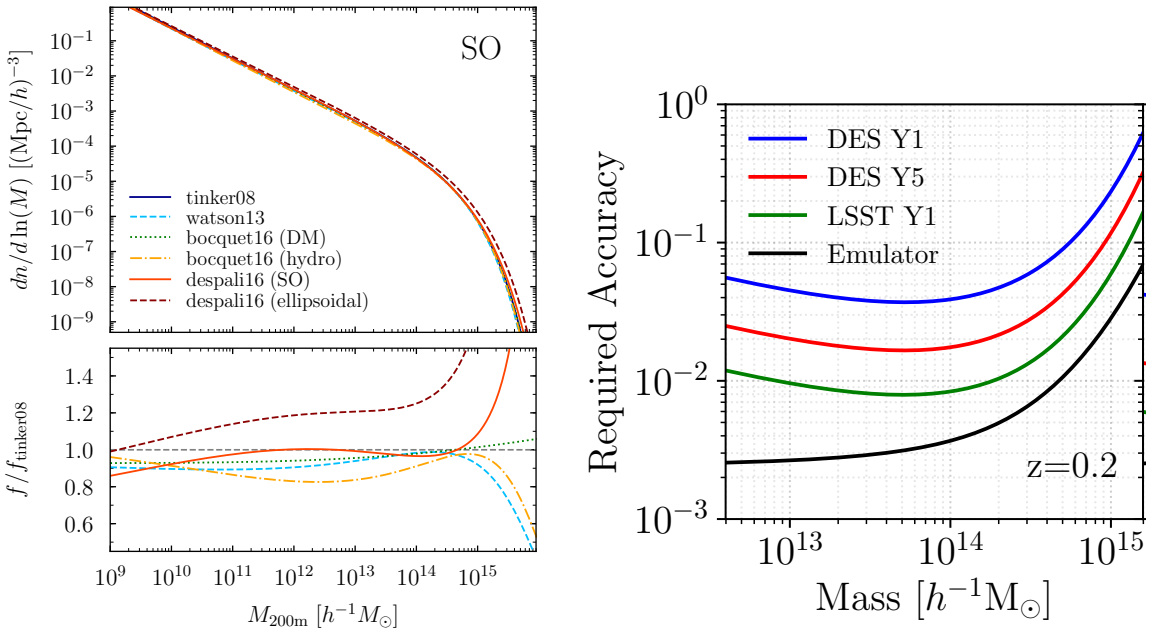

Abbildung 1.3: The difference in the predicted cluster counts due to the halo mass function model is subdominant in comparison to mass calibration and Poission uncertainties in current cosmological analysis with clusters. Usage of the Tinker et al. (2008) model is therefore justified, although future cosmological analysis with clusters detected by *eROSITA* (Merloni et al., 2012), *Euclid* (Laureijs et al., 2011) and the Rubin Observatory (LSST Science Collaboration et al., 2009) will require more accurately calibrated models (emulators) at the percent level. **Left panel**: Adapted from Diemer (2018). Comparisons of halo mass functions (Eq. 1.26) evaluated at $z = 0$. Multiplicity functions are compared relative to the Tinker et al. (2008) model (Eq. 1.27) for Tinker et al. (2008),Despali et al. (2016),Watson et al. (2013), Bocquet et al. (2016). **Right Panel:** Adapted from McClintock et al. (2019b). The halo mass function accuracy of the Aemulus emulator compared to the requirements for Year 1 of the Dark Energy Survey (DES Y1) and future datasets such as DES Y5 and the Rubin Observatory (LSST Y1) at $z = 0.2$. Mass is defined by an overdensity 200 times that of the mean matter density of the universe ($M_{\rm 200b}$).

1.3 Cluster Cosmology

To measure cosmological parameters relating to dark matter and dark energy one needs to conduct cosmological experiments using probes which are sensitive to the parameters of interest. The Dark Energy Task Force (Albrecht et al., 2006) reports that major advancements towards understanding dark energy are possible by increasing the increase the current dark energy figure-of-merit[5] by an order of magnitude. However, no single Stage-IV

[5]The figure-of-merit describes the relative performance in constraining a subset of model parameters. It is inversely proportional to the volume occupied by the 95.4 percent joint credible regions.

experiment[6] is guaranteed to achieve this ambitious goal alone.

It is therefore essential to combine results of independent experiments using a variety of probes. This provides substantially more statistical power and greater robustness to systematic uncertainties, both of which are beneficial when attempting to discriminate between dark energy models. This section (and thesis) focuses on a single cosmological probe: the abundance of clusters of galaxies.

1.3.1 The abundance of galaxy clusters

Galaxy clusters sit at the apex of cosmic hierarchy as virialised objects with masses ranging from approximately $10^{13}M_{\odot}$ for groups, to $10^{15}M_{\odot}$ for the richest clusters. Clusters originate from small amplitude density perturbations which deviated from the Hubble flow during the primordial Universe. These over-densities then relaxed, merged and coalesced, hierarchically to form progressively larger structures through gravitational interactions (Peebles, 1980; Bardeen et al., 1986). The abundance of clusters as a function of mass at different epochs (cluster mass function; Bahcall & Cen, 1993) depends on cosmological parameters and hence can be used to constrain them (for a review see Borgani & Guzzo, 2001; Allen et al., 2011; Kravtsov & Borgani, 2012; Weinberg et al., 2013; Pratt et al., 2019).

The cluster mass function is related to the dark matter halo mass function (HMF) as the largest component of cluster mass is non baryonic (Zwicky, 1933; Navarro et al., 1997). The HMF depends on large scale structure formation (Sheth & Tormen, 2002; Zentner, 2007; Tinker et al., 2008; Diemer & Kravtsov, 2015), and on the cosmological volume over which clusters are observed. These cosmological dependencies make cluster counts particularly sensitive to Ω_m, the matter density of the Universe, and σ_8, the amplitude of the matter power spectrum. The cosmological sensitivity of the cluster mass function is shown in left panel of Fig. 1.4, whereby the number of clusters per redshift interval is modified as each cosmological parameter is independently modified from it fiducial ΛCDM value (Huterer & Shafer, 2018). In order to constrain these parameters simultaneously using observed data, one must measure the number of clusters as a function of redshift and mass and perform a high dimensional fit to the data using the cosmological model and a Monte Carlo Markov Chain sampler. An example is illustrated in the right panel of Fig. 1.4.

Cluster count experiments (e.g. Vikhlinin et al., 2009b; Rozo et al., 2010; Mantz et al., 2015; Zu et al., 2014; Böhringer et al., 2014; Pacaud et al., 2018; Bocquet et al., 2019; Zubeldia & Challinor, 2019; Costanzi et al., 2019b; Kirby et al., 2019; Abbott et al., 2020; Costanzi et al., 2021) provide an independent method to complement cosmological constraints from other probes. The power of combining posteriors from a selection of cosmological probes to increase to break degeneracies across the cosmological parameter space and to increase the overall precision of the constraints is demonstrated in Fig. 1.5. This illustrates individual and combined constrains from cluster abundances (Mantz et al., 2015), the CMB (WMAP Bennett et al., 2013; Hinshaw et al., 2013), (ACT Louis et al., 2017),

[6]These include upcoming Large-Survey Telescopes (e.g. The Rubin Observatory; LSST Science Collaboration et al., 2009) and/or Square Kilometre Arrays (e.g. SKA; Dewdney et al., 2009), and/or Joint Dark Energy (Space) Missions (e.g. *Euclid*; Laureijs et al., 2011).

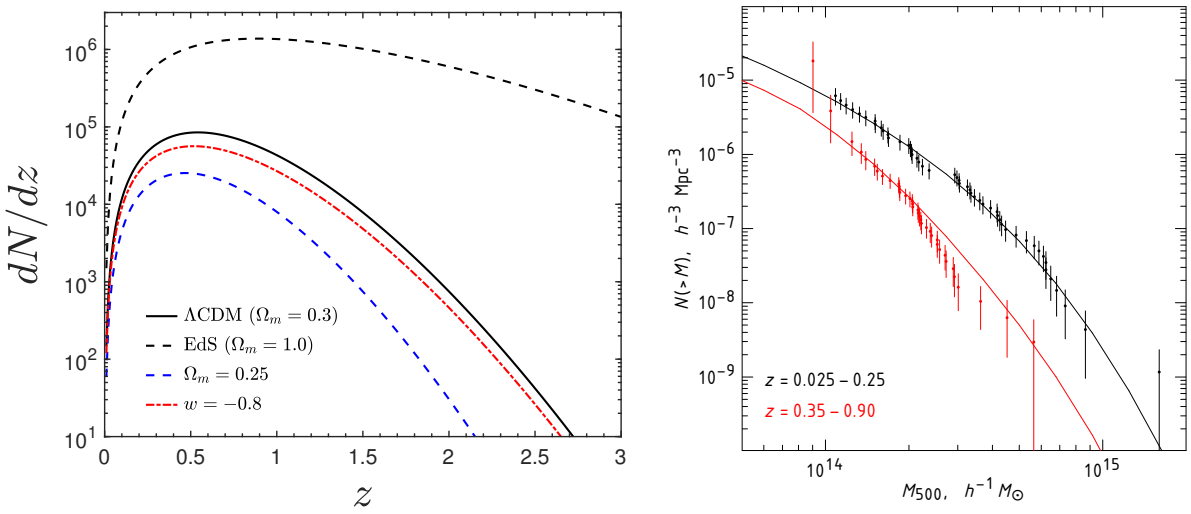

Abbildung 1.4: Adapted from Vikhlinin et al. (2009b); Huterer & Shafer (2018). **Left panel**: Four cosmological models of differential cluster counts as a function of redshift for a 5,000deg^2 survey with a halo mass sensitivity of $> 10^{14}\,h^{-1}M_\odot$. The solid black line shows a fiducial ΛCDM scenario. The dashed black line shows a matter dominated (Einstein & de Sitter, 1932) universe. The blue dashed line shows the impact of reducing Ω_{m_0} from 0.30 to 0.25 relative to the fiducial model. The dashed red line considers a modification to the fiducial model which includes dark energy with $\rho \propto a^{-3(1-w)}$. **Right panel**: The observed cluster mass function for a 400 deg^2 survey of *ROSAT* selected clusters followed up with the *Chandra* X-ray telescope. The points and solid lines indicate the comoving number density of galaxy clusters as a function of mass for observations and the best fit cosmological model respectively. The black subset shows the results for low-redshift bin $z \in [0.025, 0.250)$, while the red subset illustrates the same for a wider high-redshift bin $z \in [0.35, 0.90)$.

(SPT Keisler et al., 2011; Reichardt et al., 2012; Story et al., 2013), Type 1a Supernovae (Suzuki et al., 2012) and BAO (Beutler et al., 2011; Padmanabhan et al., 2012; Anderson et al., 2014).

1.3.2 Multi-wavelength observations of clusters

Clusters exhibit well known observational signatures which enable them to be detected across several different wavelength domains. These include quiescent red sequence galaxies (Gladders & Yee, 2005; Koester et al., 2007b; Szabo et al., 2011), intra cluster light at optical and infrared wavelengths (Conroy et al., 2007; Burke et al., 2015), X-ray emission from the extended intra cluster gas (Kellogg et al., 1975; Böhringer et al., 2000; Vikhlinin et al., 2009a) and the spectral distortion of the cosmic microwave background via the Sunyaev-Zeldovich (SZ) effect (Sunyaev & Zeldovich, 1970, 1972; Birkinshaw, 1999; Vanderlinde et al., 2010; Planck Collaboration et al., 2014).

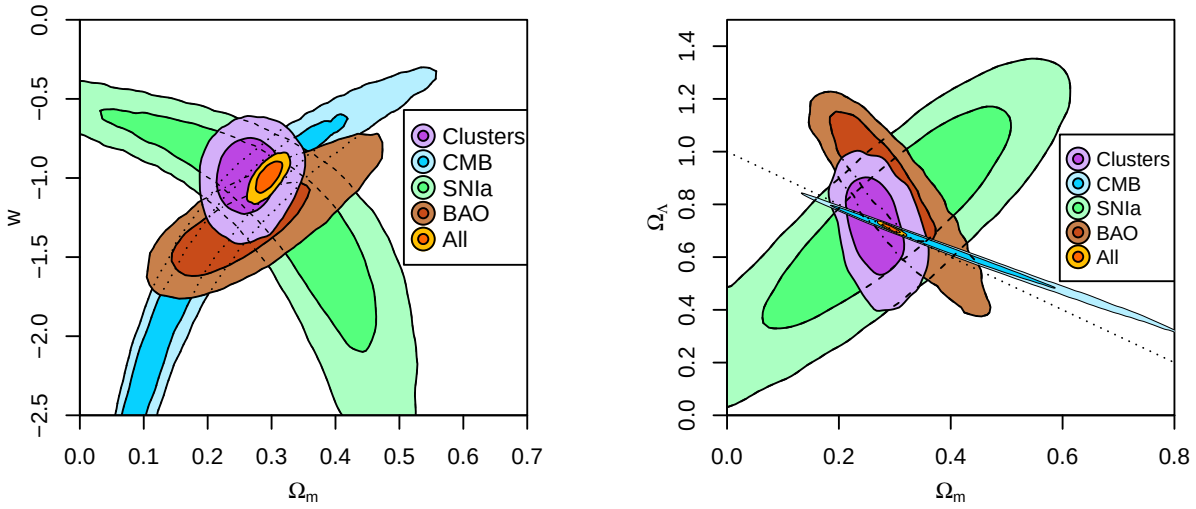

Abbildung 1.5: Adapted from Mantz et al. (2015). Cosmological posteriors from the "Weighing the Giants" cluster sample are shown by the purple contours (68.3 and 95.4 percent confidence levels). Constraints from other cosmological probes such as the CMB (WMAP Bennett et al., 2013; Hinshaw et al., 2013), (ACT Louis et al., 2017), (SPT Keisler et al., 2011; Reichardt et al., 2012; Story et al., 2013), Type 1a SNe (Suzuki et al., 2012) and BAO (Beutler et al., 2011; Padmanabhan et al., 2012; Anderson et al., 2014) are shown in blue, green and brown respectively. When combining posterior distributions from complementary probes it is possible to improve the overall figure of merit and attain more precise cosmological constrains shown in orange. The dotted line in the $\Omega_{m_0} - \Omega_\Lambda$ plane denotes spatially flat ΛCDM models. Cluster and BAO results use standard priors on h and $\Omega_b h^2$ (Riess et al., 2011; Cooke et al., 2014) although these are not included in the combined constraints.

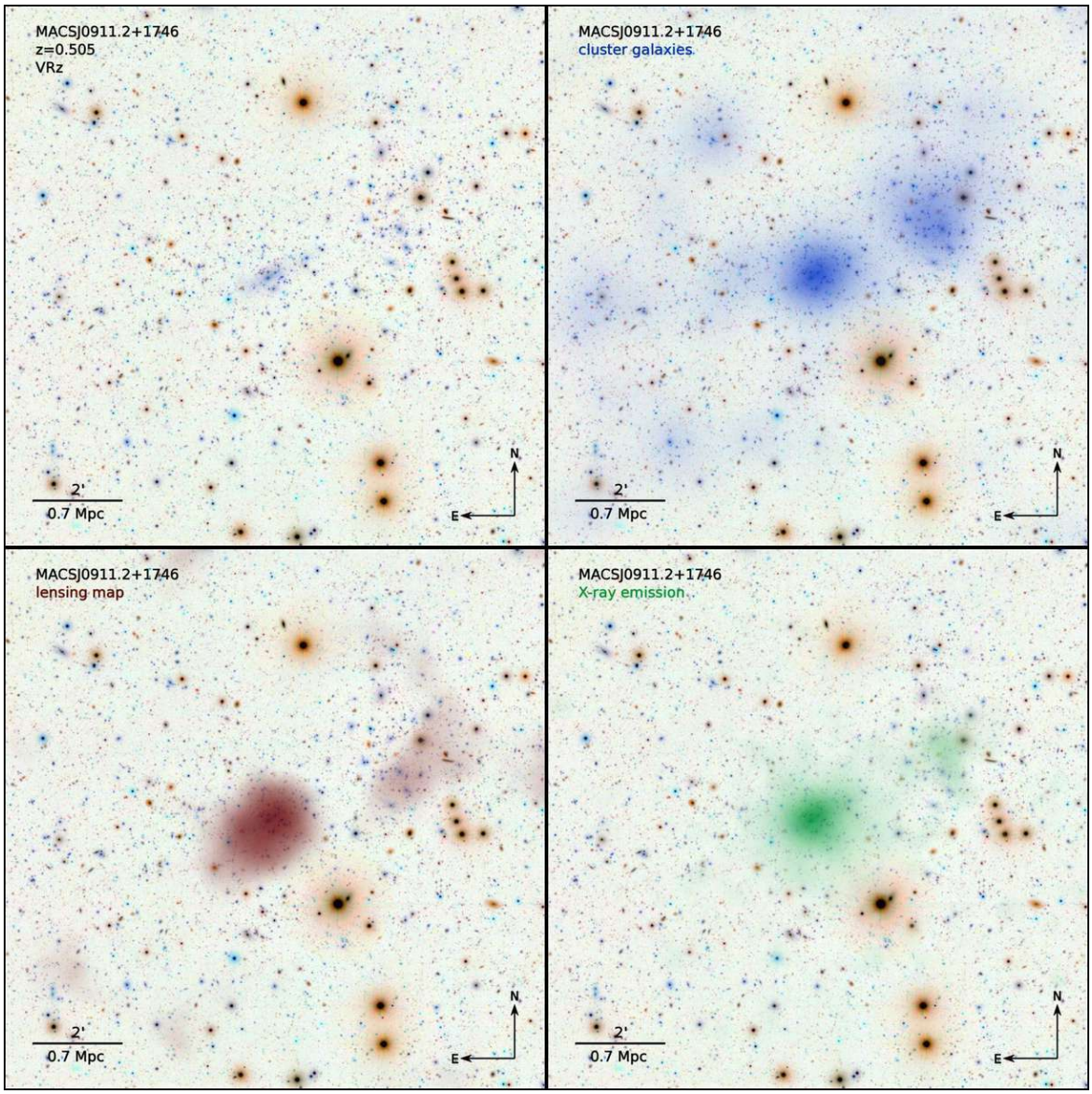

Abbildung 1.6: A multi-wavelength view of MACSJ0911.2+1746, a massive cluster of galaxies at redshift 0.505 (von der Linden et al., 2014). **Upper left panel:** An RGB optical image centred on the cluster, a secondary cluster is also visible approximately 3 arcmin towards the top right. **Upper right panel:** the mass profile inferred from the distribution of cluster galaxies (blue). **Lower left panel:** The total mass distribution deduced from weak lensing in (red). **Lower right panel:** the mass profile traced by the X-ray emission from the intracluster gas (green). https://web.stanford.edu/group/xoc/.

In order to constrain cosmology, a sample assembled using one (or more) of these methods also requires estimates of cluster mass. Arguably the most direct and accurate mass-measurement technique is weak lensing (Murata et al., 2019; McClintock et al., 2019a; Phriksee et al., 2020; Umetsu, 2020). In the case that lensing data are not available, cluster mass-proxies such as core-excised X-ray luminosity (e.g. Mantz et al., 2018), integrated SZ effect signal (Arnaud et al., 2010), velocity dispersion (Munari et al., 2013; Kirkpatrick et al., 2021), optical richness (Ider Chitham et al., 2020) etc. are used in combination with scaling relations to provide a link to the cluster mass and to the HMF.

X-ray luminosity is one of the preferred mass proxies for cluster cosmology (Reiprich & Böhringer, 2002; Vikhlinin et al., 2009b; Mantz et al., 2015; Schellenberger & Reiprich, 2017; Pacaud et al., 2018), however, optical richness is also a theoretically attractive mass proxy due to the small magnitude of the irreducible scatter (Old et al., 2018). Despite this, previous attempts to constrain cosmological parameters using richness (Bahcall et al., 2003; Gladders et al., 2007; Wen et al., 2010; Rozo et al., 2010; Tinker et al., 2012; Mana et al., 2013; Costanzi et al., 2019b) have been consistently dominated by the scatter due to noise in the measurement (Becker et al., 2007; Rykoff et al., 2012; Capasso et al., 2019c).

Each wavelength/mass inference technique brings complementary information towards constraining the cluster mass profile. This is demonstrated in Fig. 1.6, which highlights the difference between the mass profiles obtained for a massive cluster at $z = 0.505$ (MACSJ0911.2+1746, Ebeling et al., 2001) using various mass proxies at different wavelengths. This includes the total mass profile inferred from weak lensing, the gas-mass profile from X-ray data as well as the mass deduced from observations of the cluster galaxies (von der Linden et al., 2014). Each method probes a different astrophysical component of the galaxy cluster and is subject to different sources of systematic uncertainty. As a result, multi-wavelength approaches to cluster surveys are able to improve the overall accuracy of cluster mass-calibration methods (e.g. Grandis et al., 2018, 2020; Salvati et al., 2020) as well as the reliability of cluster redshifts (Clerc et al., 2016, 2020; Kirkpatrick et al., 2021).

1.3.3 The *eROSITA* cluster cosmology experiment

eROSITA (extended ROentgen Survey with an Imaging Telescope Array) is the primary X-ray instrument aboard the SRG (Spectrum-Roentgen-Gamma) observatory (Merloni et al., 2012; Predehl et al., 2021) which was successfully launched into space on the 13th of July 2019. Its main objective is to complete a four-year X-ray all-sky survey (eRASS) consisting of eight independent scans of the celestial sphere (eRASS1-8) to probe the energetic universe. The final cumulative survey (eRASS:8) will surpass the depths of the *ROSAT* All-Sky Survey (RASS; Voges et al. 1999) by a factor of approximately 25 in the soft X-ray band (0.2-2.3 keV) and produce the first survey of its kind in the hard-band (2.3-8 keV). The *eROSITA* project is a collaboration between the Max Planck Institute for Extraterrestrial Physics (MPE) and the Russian Space Research Institute (IKI), with each consortium having access to a single hemisphere. The German consortium is entitled to the scientific exploitation of X-ray data where the galactic longitude exceeds 180 degrees ($180 < l < 360$ defines the *eROSITA*-DE hemisphere).

As of today, the status of the all-sky survey is complete up to eRASS3. A false colour image of eRASS1 sky is shown in Fig. 1.7, including annotations for a selection of the most X-ray bright astronomical features e.g. supernova remnants, nebulae and the Magellanic clouds. In addition to the features that are identifiable by eye in Fig. 1.7, *eROSITA* is forecasted to detect $\sim 10^5$ clusters of galaxies out to $z \sim 1.5$ (Hofmann et al., 2017; Pillepich et al., 2018; Clerc et al., 2018; Comparat et al., 2020a) and $\sim 3 \times 10^6$ active galactic nuclei (AGN, Kolodzig et al., 2013; Comparat et al., 2019) out to $z \sim 6.4$ (Wolf et al., 2021). *eROSITA* will also detect $> 5 \times 10^5$ active stars and $> 15 \times 10^3$ normal galaxies (Basu-Zych et al., 2020). Furthermore, it will provide insights on astrophysical phenomena such as X-ray binaries (Doroshenko et al., 2014), tidal distribution events (Malyali et al., 2019) and the diffuse emission from the Milky way (Predehl et al., 2020). *eROSITA* will also test cosmic isotropy via the directional behaviour of the X-ray luminosity - temperature ($L_X - T$) galaxy cluster scaling relation (e.g. Migkas et al., 2020).

eROSITA is a powerful instrument to conduct cosmological studies, with the potential to constrain cosmological parameters with a variety of probes. *eROSITA*'s primary cosmological probe is cluster counts using X-ray as mass proxy. Additional probes include redshift space distortion in the clustering of clusters and their galaxies (e.g. de la Torre et al., 2013; Pezzotta et al., 2017; Gil-Marín et al., 2020; Bautista et al., 2021) which is expected to yield constraints on Ω_{m_0}, σ_8 and h at the $\sim 10\%$, $\sim 4\%$ and $\sim 15\%$ level for eRASS:8 (Seppi et al. in prep). AGN standard candles (e.g. Risaliti & Lusso, 2019; Lusso et al., 2020) are also expected to independently constrain Ω_{m_0} and Ω_Λ to $\sim 5\%$ and $\sim 10\%$ precision respectively (Comparat et al., 2020b). The combination of eRASS:8 cluster counts and clustering aims to constrain σ_8, Ω_{m_0}, w_0 and w_a to $\sim 1\%$, 2%, 7% and ± 0.25 respectively (Pillepich et al., 2018). Here, w_0 and w_a are the coefficients of time dependent components of the dark energy equation of state when modelled as $w(a) = w_0 + (1-a)w_a$. Fig. 1.8 demonstrates the predicted cosmological constraining power of *eROSITA* when considering optimistic and pessimistic scenarios[7] and how these predicted constraints would be modified when combined with posterior distributions derived from the CMB, BAO and the Joint Light-curve Analysis (JLA) of SNe (Planck Collaboration et al., 2016a). The combination of *eROSITA* and the next generation spectroscopic programmes (Sect. 3.5) i.e., the 4MOST-*eROSITA* cluster redshift survey (Finoguenov et al., 2019b) and SDSS-V (Black Hole Mapper; Kollmeier et al., 2017) are expected to produce preliminary (final) cosmological results in 2023 (2027).

1.4 Thesis structure & goals

As the *eROSITA* cluster cosmology experiment will be carried out after the submission of this thesis, here the goal is to measure cosmological parameters via a mass function experiment using an analogous state of the art X-ray selected cluster sample. This includes

[7] Optimistic (pessimistic) scenarios correspond to having spectroscopic (photometric) redshifts, $L-M$ relation priors which are 4 times better than (equal to) Vikhlinin et al. (2009a), and the inclusion of all clusters across the entire sky with more than 50 X-ray photon counts and $M_{500c} > 1(5) \times 10^{13} M_\odot h^{-1}$.

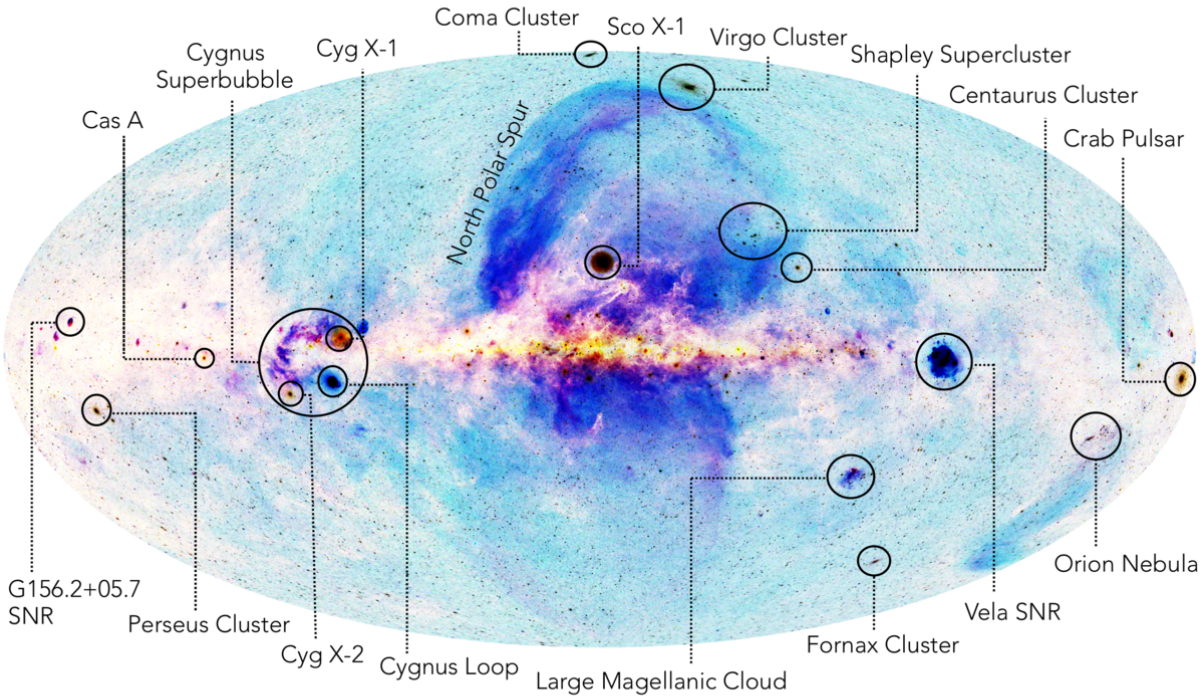

Abbildung 1.7: Adapted from the eRASS1 press-kit (Sanders, Brunner, Merloni, Churazov, Gilfanov). The energetic universe imaged by *eROSITA* in the 0.3-2.3 KeV X-ray band (Blue: 0.3-0.6 keV, pink/purple: 0.6-1 keV, yellow: 1-2.3 keV). The bright strip across the centre of the image is the galactic plane, while the large X-ray bubbles in the Milky Way halo (Predehl et al., 2020) can be seen in purple. The most prominent X-ray astronomical objects that appear as bright X-ray sources are annotated. This includes a selection of iconic low-redshift galaxy clusters, supernova remnants, nebulae and the Magellanic Clouds. https://www.mpe.mpg.de/7461950/erass1-presskit.

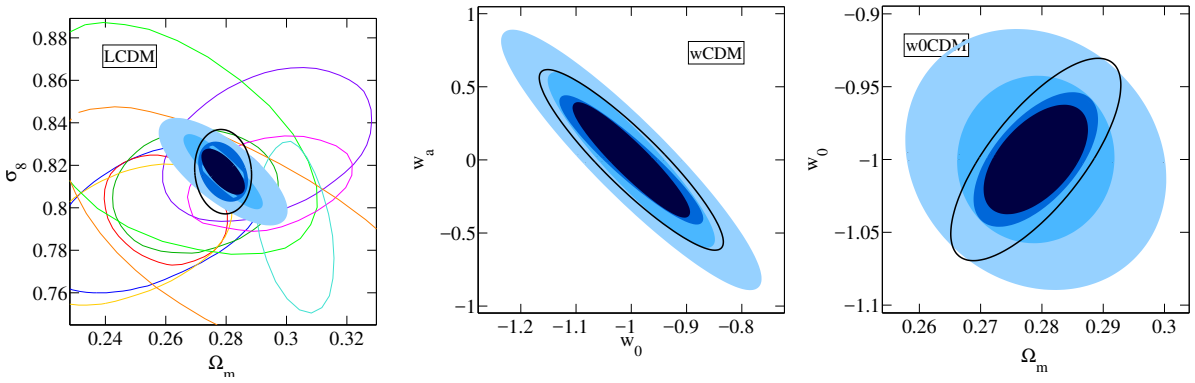

Abbildung 1.8: Adapted from Pillepich et al. (2018). Cosmological forecast of eRASS:8 constraints from the combination of cluster abundances and clustering for various dark energy models (ΛCDM, wCDM and w_0CDM from left to right). Joint 68.3 percent credible regions are depicted by shades of blue for a selection of parameter pairs: pessimistic/optimistic (light-blue pair), *Planck* CMB constraints combined with BAO and JLA constraints (thin black line, Planck Collaboration et al., 2016a), pessimistic/optimistic including CMB + BAO + JLA (dark-blue pair).

significant contributions to the spectroscopic identification of *eROSITA* sources (Sect. 3.1) and the measurement of redshifts from optical photometry and spectroscopy. The scientific objectives of the thesis are summarised below:

- Measure cluster redshifts using optical photometry and spectroscopy for X-ray selected samples of galaxy clusters. Maximise the effective survey area (and cosmological volume) of the *eROSITA*-DE sky with optical coverage using heterogeneous publicly available photometric and spectroscopic data sets. Prove the precision and accuracy of the measured redshifts. This is achieved in Kap. 2 - Theory: measuring galaxy cluster redshifts.

- Create a spectroscopic catalogue of X-ray selected galaxy clusters in preparation for cosmological cluster count analyses. This includes the provision of spectroscopic targets for follow up programmes within SDSS and 4MOST. It also includes methods to create a volume limited sample and a model to calculate its effective area and selection function as a function of redshift. This is achieved in Kap. 3 - Method: creating spectroscopically confirmed cluster catalogues.

- Constrain Ω_{m_0}, σ_8 as well as free parameters of the observable-mass relation(s) during cosmological analyses of spectroscopically confirmed X-ray selected galaxy clusters using optical and dynamical mass proxies. The cosmological pipeline should be generalised such that it can be adapted for the eRASS cluster sample. This is achieved in Kap. 4 - Results: cosmological interpretation of cluster abundances.

Kapitel 2

Theory: measuring galaxy cluster redshifts

As discussed in Kap. 1, the abundance of clusters of galaxies can be used to constrain cosmological parameters. In order to measure the cluster mass function, one needs a sample of galaxy clusters with measurements of distance and mass (or an observable mass-proxy). The distance to each cluster must be measured from cluster redshifts under the assumption of a cosmological model (Sect. 1.2.1). This chapter provides an overview of measuring redshifts for X-ray selected clusters using optical photometry and cluster finding/confirmation algorithms. It also demonstrates how optical spectroscopy is used to refine the photometric redshift measurement by dynamically processing member galaxies associated with the extended X-ray cluster emission. On average, spectroscopic cluster redshifts (Sect. 2.3) are a factor of ten more precise than those based on optical photometry (Sect. 2.2). Spectroscopic redshifts are therefore extremely useful in the context of cluster cosmology as they significantly reduce the uncertainty of cluster observables, cosmological distance measurements.

Work presented in this chapter is based on Ider Chitham et al. in prep., Ider Chitham et al. (2020) and the author's contributions to Finoguenov et al. (2020); Clerc et al. (2020); Kirkpatrick et al. (2021). The theory behind of photometric (Sect. 2.2) and spectroscopic (Sect. 2.3) galaxy cluster redshift algorithms used in this thesis are inspired by the pioneering work of Mancone & Gonzalez (2012); Rykoff et al. (2012); Klein et al. (2018) and Clerc et al. (2016); Capasso et al. (2019b); Ferragamo et al. (2020) respectively.

2.1 Introduction

For extragalactic sources such as clusters and their galaxies, the observable which facilitates the measurement of redshift is their spectrum. The spectrum is usually modelled as a combination of continuum emission as well as emission and/or absorption lines. The rest-frame wavelengths that contribute to the spectrum of a source (at redshift z) are shifted in the red direction by a factor of $1 + z$ when observed (at redshift zero). This allows one

to measure redshifts over a wide range of wavelengths using known (rest-frame) spectral features as points of reference.

Although it is possible to measure redshifts from the emission lines found in the X-ray spectra of the ICM (e.g. Hashimoto et al., 2004; Lamer et al., 2008; Lloyd-Davies et al., 2011; Yu et al., 2011), for *eROSITA* this will only be possible for a subset of the clusters with intermediate masses at $z < 0.45$ at a precision of $\Delta z/(1+z) \leq 10\%$ assuming an exposure time of $t_{\rm exp} = 1.6$ks (Borm et al., 2014). This is due to the fact that redshift precision depends strongly on the signal-to-noise of X-ray spectra, therefore, precise measurements are only feasible for very bright clusters or those which have longer exposure times (e.g. at the poles of the eRASS). As galaxy clusters contain over-densities of early-type galaxies relative to the field (Dressler, 1980), the most accurate and precise cluster redshifts are obtained from an ensemble of galaxy redshifts (e.g. Beers et al., 1990; Böhringer et al., 2004; Clerc et al., 2016; Ider Chitham et al., 2020) measured via optical spectroscopy (e.g. Szokoly et al., 2004; Koulouridis et al., 2016; Clerc et al., 2020).

An example of typical spectra from early-type (elliptical, bulge and lenticular) galaxies is shown in Fig. 2.1 along with those from late-type (spiral and starburst) galaxies. This demonstrates that the rest-frame spectral features of galaxies are a strong function of morphology. The most prominent features present in the spectra of late-type galaxies are a blue continuum and emission lines (e.g. [OII], Hβ, [OIII], Hα and [NII]). This is due to their relatively high rates of star formation i.e. populations of young and hot stars. On the other hand, early-type galaxy spectra tend to contain absorption lines (e.g. CaII H and K, G band, MgI and NaD) with very little emission below the 4000Å break. This is due to their relatively lower rates of star formation i.e. populations of older and redder stars. Fig. 2.2 illustrates how these prominent features are redshifted in the optical spectra of 32,326 member galaxies from 2,740 X-ray selected clusters (most of which contribute to Fig. 2.15 as detailed in Sect. 3.1). In this case, galaxy redshifts are extracted from the observed spectra using archetype template fitting (e.g. Cool et al., 2013; Ross et al., 2020; Raichoor et al., 2021)[1] although principal component analysis (e.g. Bolton et al., 2012) and cross-correlation (e.g. Hinton et al., 2016) techniques are also frequently used in the literature.

Efficiently observing optical spectra for large samples of extragalactic sources is best achieved using ground-based multi object spectrographs (MOS). This simultaneously measures the spectra of many sources as appose to single slit spectroscopy which is limited to measuring a single spectrum at a time. Examples of large spectroscopic surveys of galaxies include 2dFGRS (Colless et al., 2001), 6dFGS (Jones et al., 2009), GAMA (Driver et al., 2011), VVDS (Le Fèvre et al., 2013), DEEP2 (Newman et al., 2013), SDSS-IV DR16 (Ahumada et al., 2020). The next generation of such surveys that are most relevant for the spectroscopic analysis of *eROSITA*-DE clusters (detailed in Sect. 3.5) include 4MOST (de Jong et al., 2019), DESI (DESI Collaboration et al., 2016) and SDSS-V (Black Hole Mapper; Kollmeier et al., 2017). Carrying out such large optical spectroscopic surveys

[1]The DESI (DESI Collaboration et al., 2016) pipeline classifier `redrock` https://github.com/desihub/redrock is used in the cited eBOSS work.

of galaxy clusters is challenging not only due to the need for deep optical photometric imaging and deep spectroscopic exposures. The former is required in order to select the most probable galaxies (Sect. 2.2.1) and the latter needed to confidently observe the desired spectral features at the required signal-to-noise ratio. In MOS the maximum number density of targets which can be simultaneously observed is also limited to avoid cross-talk (spectral contamination due to neighbouring fibres). As a result of these challenges, many spectroscopic surveys are magnitude limited with increasing incompleteness towards high redshift. However, approximations to true cluster redshifts can also be estimated using photometry (Sect. 2.2). Photometric redshift techniques are often complementary to spectroscopic redshift measurements as they can provide redshift information relatively quickly and inexpensively compared to large spectroscopic programmes.

2.2 Photometric cluster redshifts

As galaxy cluster surveys often span vast cosmological volumes it is infeasible to measure the optical spectrum of every observable galaxy within the survey footprint. It is therefore common practice to estimate cluster redshifts using photometry prior to subsequent spectroscopic observations of their member galaxies. Photometric redshifts utilise a set of multi-band filters to sample the spectral energy distribution (SED) of an object in order to approximate its true redshift. In general, the performance of photometric redshift techniques depends on a number of factors e.g. the quality of the photometry, how well the desired spectral features are sampled by the subset of the spectrum encompassed by the photometric filters, the robustness of the calibration methods and how representative the spectroscopic training dataset(s) are. For a general review of photometric redshifts in astronomy please refer to Salvato et al. (2019).

Spectroscopically observing the subset of galaxies, which are most likely to be associated to galaxy clusters, at the photometric cluster redshift has proved to be an extremely efficient targeting method within optical spectroscopic programmes for X-ray selected galaxy clusters (e.g. Clerc et al., 2016, Ider Chitham in prep.). Photometric redshifts are therefore closely tied to the spectroscopic target selection process. This section introduces photometric redshift estimators for galaxy clusters (and their galaxies) and explains how matched filter algorithms are used to optically select the most probable cluster members for spectroscopic follow-up.

2.2.1 Optical Cluster Finding

Manually searching for clusters of galaxies in optical imaging is one of the oldest methods to identify them (Abell, 1958). Over the last few decades, many different algorithms have been developed to decipher optical and NIR multi-band photometry as well as spectroscopic datasets in order to automate this procedure. Most depend on a combination of the geometrical distribution of galaxies in projection and along the line-of-site, their colour (indicative of their stellar population, redshift etc.) as well as their luminosities. They are

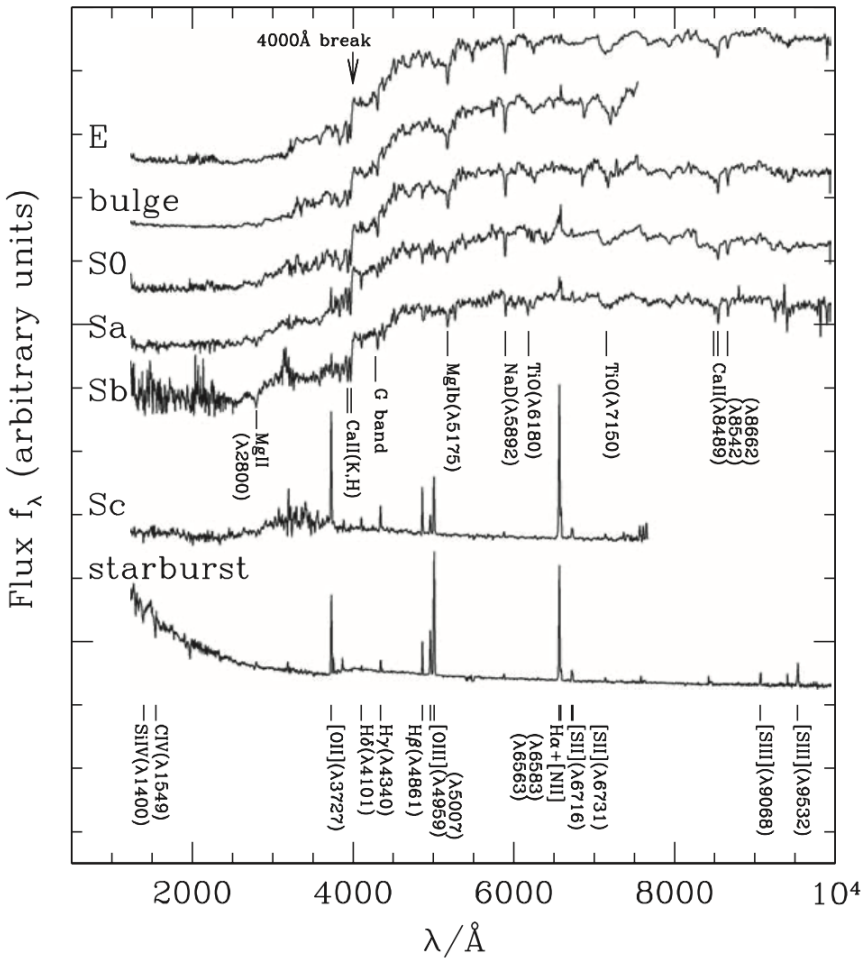

Abbildung 2.1: Adapted from Mo et al. (2010). Galaxy spectra for different morphological types: elliptical (E), lenticular (S0), spiral (Sa, Sb, Sc) and starburst. In late type galaxies, the blue continuum and emission lines are prominent due to the population of young and hot bright blue stars. On the other hand, early type galaxies have an older, redder stellar population due to lower star formation rates. Therefore, their spectrum contains no emission lines with very little flux below 4000Å, which marks the absorption lines of calcium (H and K).

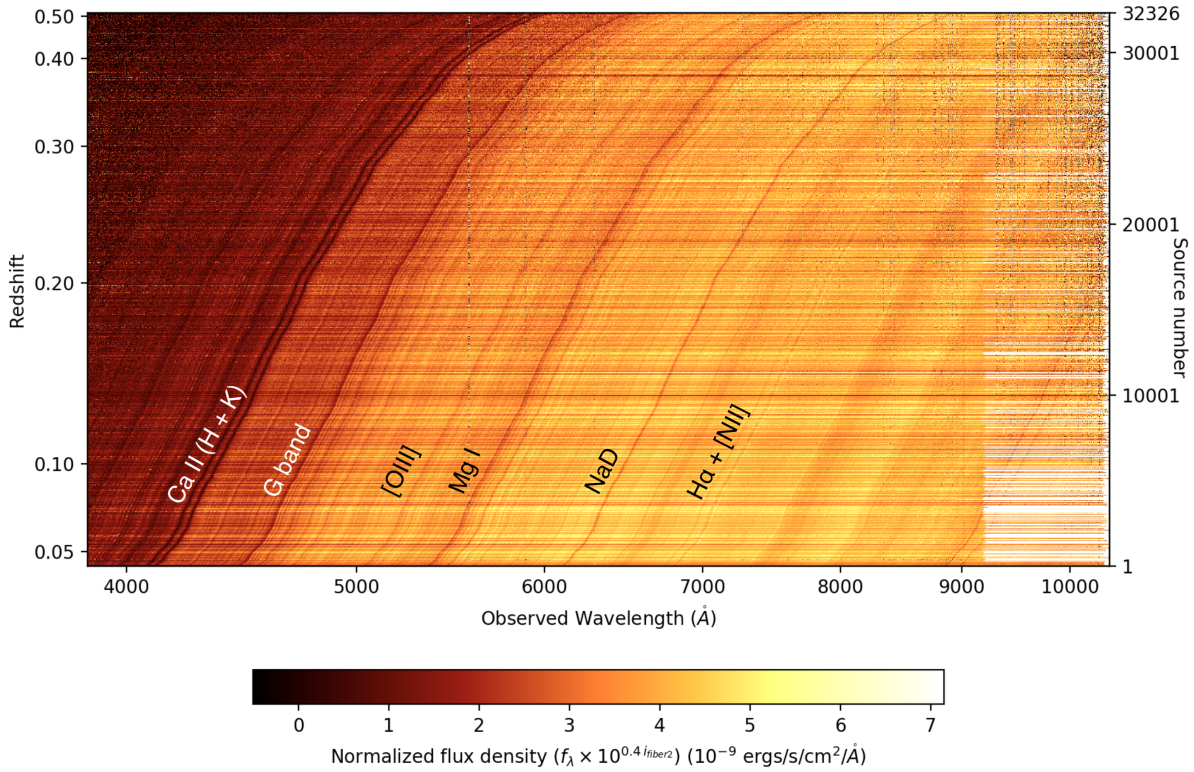

Abbildung 2.2: Adapted from Clerc et al. (2020). Flux-calibrated optical spectra corresponding to 32,326 cluster members in the original visually inspected SPIDERS DR16 galaxy cluster catalogue. The colour represents the flux density in each wavelength pixel of a spectrum, normalised by the spectroscopic fibre flux of each member galaxy. The main absorption features are labelled e.g. Ca II H + K, G-band, Mg I, NaD along with some galaxies that have emission features e.g. [O III] and H α (although these are very faint).

.

often classified by the physical information that is utilised e.g. red-sequence galaxy based techniques (e.g. Gladders & Yee, 2000; Gladders et al., 2007; Oguri, 2014), techniques that search for any colour over-densities (Miller et al., 2005), those based on the presence of brightest cluster galaxies (BCG; e.g. Koester et al., 2007a; Hao et al., 2010). Other classifications include photometric-redshift based methods (e.g. Wen et al., 2012; Tempel et al., 2018; Bellagamba et al., 2018; Aguena et al., 2021) and methods that are designed to be applied directly to spectroscopic galaxy surveys (Duarte & Mamon, 2015; Old et al., 2015, e.g.). Optical cluster finding algorithms are also classified by the statistical methodology that is implemented i.e. matched-filter based algorithms[2] (e.g. Postman et al., 1996; Olsen et al., 2007; Szabo et al., 2011; Klein et al., 2018), Voronoi tessellation methods (e.g. Ramella et al., 2001; Soares-Santos et al., 2011; Murphy et al., 2012), friends-of-friends (FoF; e.g. Wen et al., 2012) and percolation algorithms (e.g. Dalton et al., 1997; Rykoff et al., 2014).

The two "matched-filter" algorithms implemented in this thesis are known as redMaPPer (Red-sequence matched-filter Probabilistic Percolation, Sect. 2.2.2) and MCMF (Multi-Component matched-filter, Sect. 2.2.3). Each algorithm includes a very similar colour filter, galaxy luminosity filter and radial filter. These are briefly introduced in a general context in the subsequent subsections prior to the unique details of each algorithm.

The radial distribution of cluster galaxies

The distribution of galaxies within clusters follows a morphology-density relation (Dressler, 1980) whereby star-forming spiral galaxies ($\sim 20\%$) tend to reside in less dense regions, while lenticular ($\sim 40\%$) and ellipticals ($\sim 30\%$) are more concentrated in inner regions of the cluster (Oemler, 1974). This is due to the fact that the oldest quiescent galaxies, those which have very low star formation rates are have had a longer time to reach the depths of the gravitational potential well of the cluster. A radial cluster galaxy filter $\Sigma(R)$ describes the twp dimensional projected density profile of cluster galaxies. Common profiles include the NFW Navarro et al. (1997) and Einasto (1965) profiles.

Cluster galaxy luminosity function

The luminosity function of galaxies within clusters is well described by a Schechter (1976) function (Eq. 2.1) expressed in terms of apparent magnitude, m (e.g. Hansen et al., 2009).

$$\phi(m) \propto 10^{-0.4(m-m_*)(\alpha+1)} \exp\left(-10^{-0.4(m-m_*)}\right), \tag{2.1}$$

Here m_* is the characteristic magnitude ("knee") of the luminosity function. The faint-end slope is often set to $\alpha = -1$ (independent of redshift). m_* depends strongly on redshift and the photometric band pass used to define magnitudes as described in Sect. 2.2.1. These dependencies are shown in the middle column of Fig. 2.3, whereby redder bands correspond to systematically lower values of $m_*(z)$ which also tend to plateau at lower redshifts.

[2]This is not necessarily specific to optical wavelengths. Analogous techniques have also been successfully implemented at microwave and X-ray wavelengths (e.g. Tarrío et al., 2016, 2018).

Red-sequence colour models

Most optical cluster finders rely on the presence of early-type galaxies which exhibit a well–defined colour-magnitude relation known as the red sequence (Gladders et al., 1998). Models of the apparent colour and luminosity function of red-sequence galaxies are derived from Stellar population synthesis (SPS) models. SPS models describe how the SED of a stellar population changes as a function of age, given a star formation history, metallicity and initial mass function (IMF). In order to compare SPS models with the observed properties of galaxies in in a photometric survey, Eq. 2.2 is used to project a redshifted SED at a given age, $F_\nu[\nu(1+z), t(z, z_f)]$, into models of absolute magnitude, $M_{AB}\,[z, t(z, z_f)]$. This is achieved by integrating over the filter response curve of the imaging telescope $R(\nu)$ (shown for various bands and surveys in the first column of Fig. 2.3).

$$M_{AB}\,[z, t(z, z_f)] = -2.5 \log \left[\frac{\int_{-\infty}^{\infty} \nu^{-1}(1+z) F_\nu[\nu(1+z), t(z, z_f)] R(\nu) d\nu}{\int_{-\infty}^{\infty} \nu^{-1} R(\nu) d\nu} \right] - 48.60 \quad (2.2)$$

For any arbitrary photometric survey filter band, $m_*(z)$ is computed using a Bruzual & Charlot (2003) model. This provides the magnitude evolution of galaxies with a burst of star formation at $z = z_f = 3.0$ assuming they have a solar metallicity and follow a Salpeter IMF (implemented with the `EzGal` Python package Mancone & Gonzalez, 2012). In the case of redMaPPer red-sequence models, $m_*(z)$ is normalised such that $m_{i,\mathrm{SDSS}} = 17.85$ at $z = 0.2$ for an L_* galaxy in the AB system for consistency with other work (Rykoff et al., 2014, 2016).

Models of the expected red-sequence galaxy colour, $\langle c(z) \rangle$, are derived in an analogous way to $m_*(z)$. This is illustrated in third column of Fig. 2.3 for a variety of colours and surveys. As $\langle c(z) \rangle$ represents a theoretical prediction, it systematically differs from the colours of observed galaxies depending on the method of photometric processing used in each optical survey. It is therefore necessary to account for any offsets with respect to the observed data via an empirical calibration process. Optical cluster finding algorithms achieve this by making use of a training set of known clusters (or their members) with accurate spectroscopic redshifts and well defined magnitude measurements and uncertainties. As calibration methods differ for redMaPPer and MCMF, the details of these processes are described in their respective subsections (Sects. 2.2.2 & 2.2.3).

2.2.2 Red-sequence matched-filter Probabilistic Percolation

This section is largely based on the pioneering work of Rykoff et al. (2012, 2014, 2016), in the context of creating the SPIDERS cosmological sample of clusters (Ider Chitham et al., 2020, described in Sect. 3.4). redMaPPer is a matched-filter red-sequence photometric optical cluster finding algorithm with implements the three filters outlined in the previous subsection. redMaPPer is the most widely used cluster finding tool in the literature, due to its high-purity as measured using simulations ($\gtrsim 95\%$ in the SDSS when $\tilde{z} \in [0.1, 0.6)$, Rykoff et al., 2014) and empirically using spectroscopic data ($\gtrsim 90\%$ even at the lowest

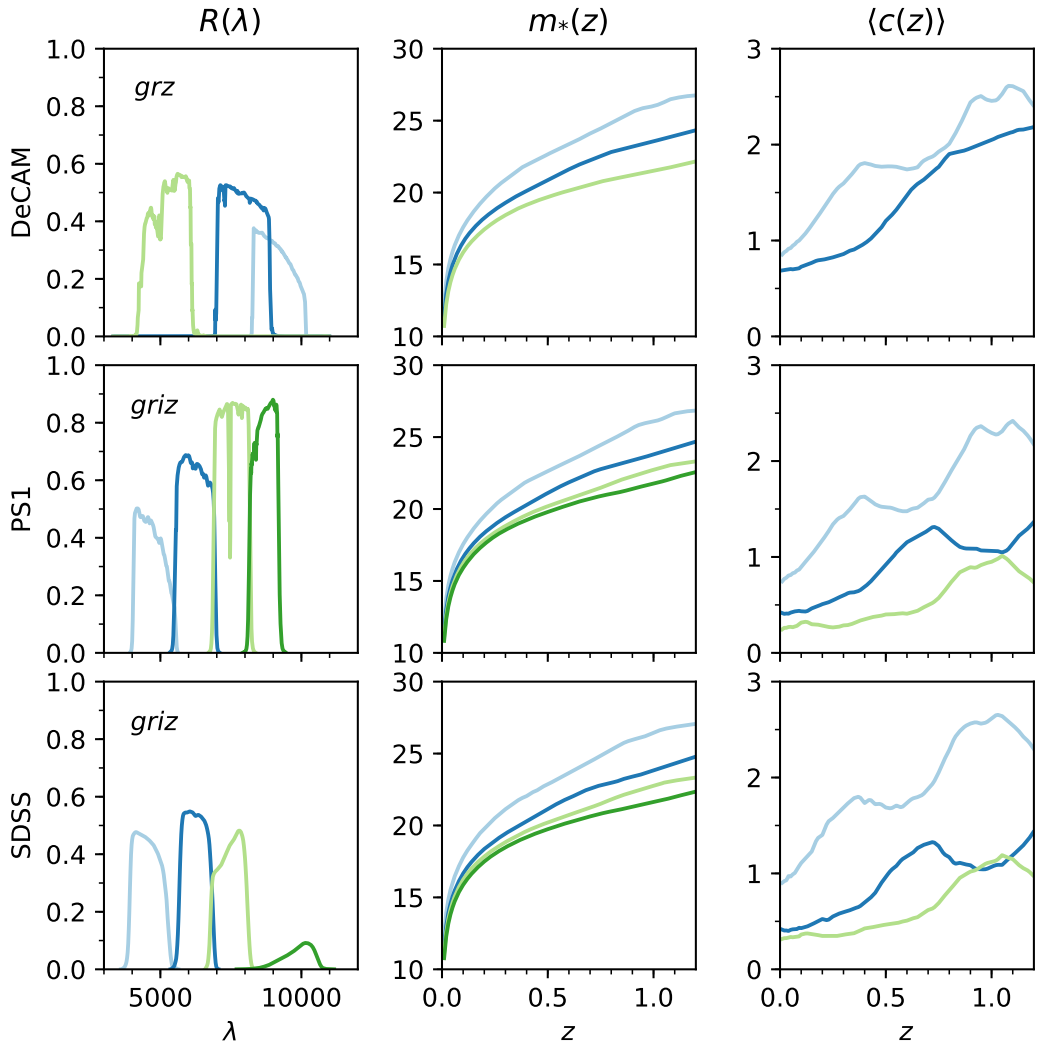

Abbildung 2.3: Adapted from Ider Chitham et al. in prep. Each row corresponds to a different photometric system the Dark Energy Camera (DECam), Pan-STARRS and the Sloan Digital Sky Survey. Relevant filters are described in the top left of each subplot and listed in order of median wavelength. **Left:** Transmission curves, $R(\lambda)$ as a function of wavelength for each optical and NIR filter. These are the inputs to EzGal (Sect. 2.2.1). **Middle:** Models of characteristic cluster galaxy luminosity $m_*(z)$ as a function of redshift colour-coded consistently with the respective transmission curves for each band. **Right:** red-sequence colour model $\langle c(z) \rangle$ as a function of redshift for pair of adjacent bands. For consistency all models are evaluated at the same fiducial cosmology redMaPPer ($\Omega_m = 0.3, \Omega_\Lambda = 0.7, h = 0.7$).

values of optical richness in the HectoMAP redshift survey down to $r = 21.3$, Sohn et al., 2020) and high-completeness, public availability and flexible usage as a cluster finder (e.g. Rykoff et al., 2012, 2016) or follow up tool (e.g. Rozo & Rykoff, 2014; Finoguenov et al., 2020; Bleem et al., 2020). Many observational programmes that optical spectroscopy for galaxies in clusters are facilitated by redMaPPer (e.g. Clerc et al., 2016; Rykoff et al., 2016; Sohn et al., 2018b; Rines et al., 2018; Kirkpatrick et al., 2021; Clerc et al., 2020, Ider Chitham et al. in prep.). It has also been the subject of numerous cosmological experiments (e.g. Costanzi et al., 2019b; Kirby et al., 2019; Costanzi et al., 2021; Ider Chitham et al., 2020), as well as mass calibration analyses (e.g. Saro et al., 2015; Baxter et al., 2016; Farahi et al., 2016; Melchior et al., 2017; Jimeno et al., 2018; Jimeno, Diego, Broadhurst, De Martino & Lazkoz, Mur; Raghunathan et al., 2019; McClintock et al., 2019a; Capasso et al., 2019c; Palmese et al., 2020). The impact of project effects (e.g. Costanzi et al., 2019a; Myles et al., 2020), centring (e.g. Rozo & Rykoff, 2014; Hoshino et al., 2015; Zhang et al., 2019; Hollowood et al., 2019; Hikage et al., 2018) and intrinsic alignment (Huang et al., 2018) have also been studied in detail.

Optical richness

redMaPPer combines the colour, luminosity and radial galaxy filters outlined in Sect. 2.2.1 to form a single filter function

$$u(\mathbf{x}) = [2\pi R\Sigma(R)]\phi(m)\rho_\nu(\chi^2). \tag{2.3}$$

Here, $\mathbf{x}$ represents a vector of galaxy properties (magnitude, colour, position on the sky etc.), $2\pi R\Sigma(R)$ is a projected NFW dimensional density profile, $\phi(m)$ is a Schechter (1976) luminosity filter and $\rho_\nu(\chi^2)$ is a colour filter described by χ^2 distribution (Eq. 2.8). Optical richness, (λ, Eq. 2.4) is defined as the sum of the membership probabilities over all member galaxies within a scale-radius[3] (R_λ, Eq. 2.6) that are above a luminosity threshold ($L_{\rm thresh}$). The probability of membership, $p_{\rm mem}$, depends on the properties of the cluster as well as the galaxy filter, $u(\mathbf{x})$, relative to the background, $b(\mathbf{x})$. It is therefore necessary to solve for these values simultaneously using

$$\lambda = \sum p_{\rm mem} \equiv \sum p(\mathbf{x}|\lambda) = \sum_{R<R_\lambda} \frac{\lambda u(\mathbf{x}|\lambda)}{\lambda u(\mathbf{x}|\lambda) + b(\mathbf{x})}, \tag{2.4}$$

where the statistical uncertainty is given by

$$\mathrm{Var}(\lambda) = \sum p(\mathbf{x}|\lambda)\left[1 - p(\mathbf{x}|\lambda)\right]. \tag{2.5}$$

The scatter in the mass-richness relation is expected to be minimised when $L_{\rm thresh} = 0.2L_*$ ($m_* + 1.75$, Rykoff et al., 2012). The scale-radius is given by

$$R_\lambda = R_0(\lambda/100)^\beta h^{-1}\,\mathrm{Mpc}, \tag{2.6}$$

where $R_0 = 1.0\,h^{-1}\,\mathrm{Mpc}$ and $\beta = 0.2$.

[3]This is not a proxy for any sort of standard over-density radius such as R_{500c} or R_{200m}.

Colour Filter model

The colour filter measures how likely each galaxy in the vicinity of a cluster candidate is to be a member of the red-sequence galaxy population. It is of the following form

$$\chi^2(z) = (\mathbf{c} - \langle \mathbf{c}|z, m\rangle) \left(\mathbf{C}_{\rm int}(z) + \mathbf{C}_{\rm err}\right)^{-1} (\mathbf{c} - \langle \mathbf{c}|z, m\rangle) \tag{2.7}$$

where $\mathbf{c}$ is the observed colour vector, m is the observed reference-band[4] apparent magnitude, $\langle \mathbf{c}|z, m\rangle$ is the colour model, $\mathbf{C}_{\rm int}(z)$ is the redshift-dependent covariance matrix and $\mathbf{C}_{\rm err}$ is the photometric error matrix. For red sequence galaxies, Eq. 2.7 follows a χ^2 distribution with ν degrees of freedom (one for each colour),

$$\rho_\nu(\chi^2) = \frac{(\chi^2)^{(\nu/2-1)} e^{-\chi^2/2}}{2^{\nu/2}\Gamma(\nu/2)}, \tag{2.8}$$

Calibration

In order to create a colour-filter based on real cluster galaxies; $\langle \mathbf{c}|z, m\rangle$, $\mathbf{C}_{\rm int}(z)$ and $\mathbf{C}_{\rm err}$ must be calibrated relative to initial red-sequence models (Sect. 2.2.1) using a set of previously observed training galaxies. In short, this is achieved by matching the subset of spectroscopic galaxies which have been confirmed as clusters members (or the brightest cluster galaxies) with the photometric survey of interest. A 2σ cut is then made about the median initial $\langle \mathbf{c}|z, m\rangle$ values to eliminate outliers. Cuts are applied to each colour, only in the relevant redshift interval where the red-sequence is expected to be populated. In other words, the slope of the red-sequence in the redshift-colour plane is indicative of the sensitivity to measuring photometric redshifts, therefore it is only necessary to make cuts in a single colour per redshift interval depending on which colour provides the most information for photometric redshift determination. This is shown for DR8 of The Legacy Surveys in Fig. 2.4, where a cut is applied in $g - r$ for $z \in [0.05, 0.35)$ and in $r - z$ for $z \in [0.35, 0.85)$. The colour, slope and the width of the red-sequence are then modelled as a function of redshift as shown in Fig. 2.5. This process is repeated for each iteration of the calibration process.

Galaxy Redshift Estimation

redMaPPer is based on a custom photometric redshift estimator $z_{\rm red}$ which is designed to estimate the redshift of red-sequence cluster member *galaxies*. The probability distribution of red-sequence galaxy colour depends on the colour filter (Eq. 2.7)

$$P(\mathbf{c}) \propto \exp\left(-\frac{1}{2}\chi^2\right) \tag{2.9}$$

[4]Ideally, the photometric reference-band magnitude should be chosen to be that which is the most representative of the total magnitude. In the case of forced photometry, the detection band is often chosen e.g.if sources are primarily detected in the i-band then $m = m_i$. However, if a particular redshift regime is of interest it is useful to select the reference-band accordingly e.g. the z-band is recommended for cluster-finding at higher redshift $z > 0.8$ therefore $m = m_z$.

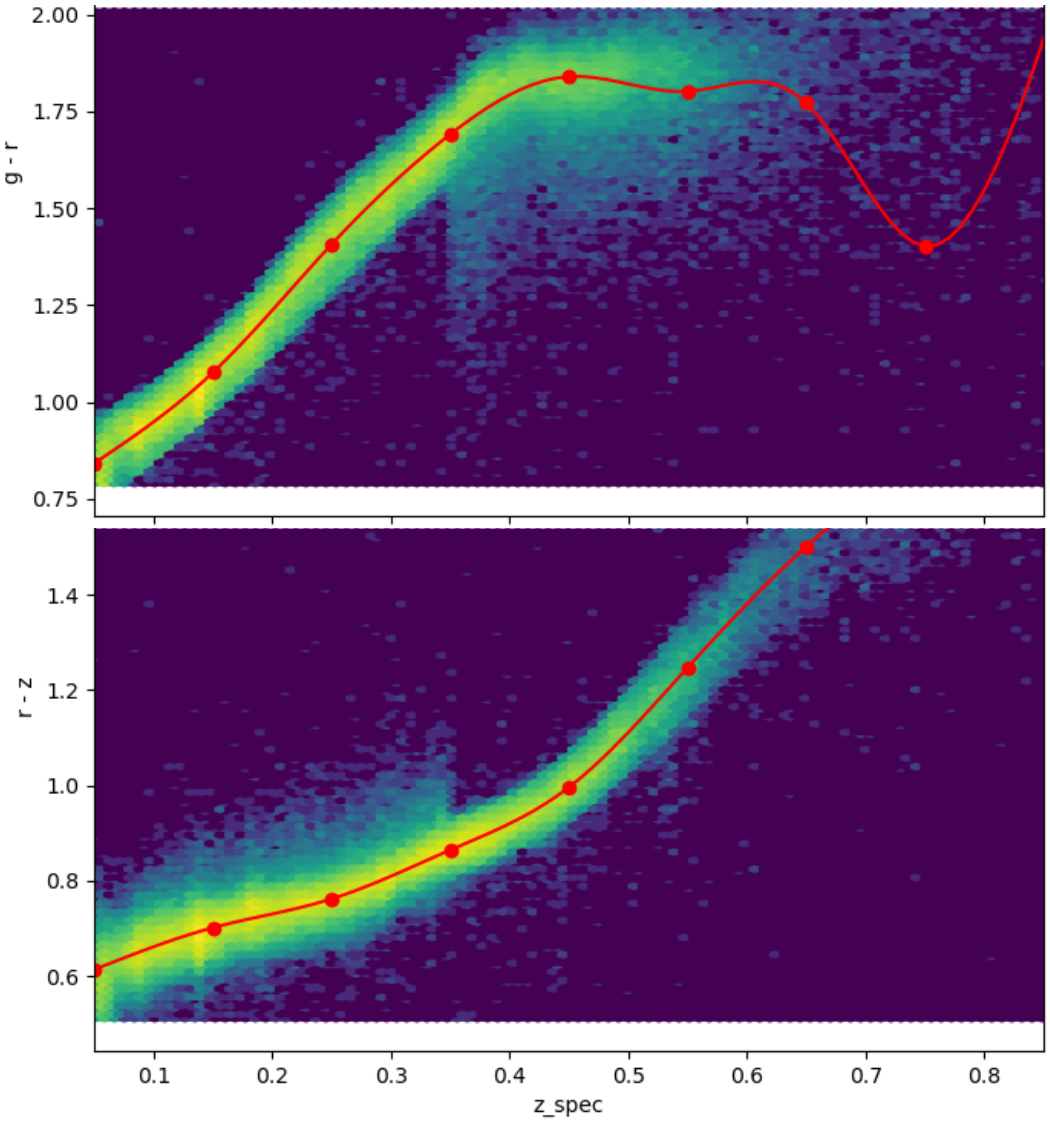

Abbildung 2.4: Adapted from Ider Chitham et al. (2020). Training sample of red spectroscopic galaxies used in the calibration process for DR8 of The Legacy Surveys (Ider Chitham et al., 2020). These galaxies are selected from a subset of confirmed member galaxies present in the archival spectroscopic compilation (Sect. 2.3.2) $2 \times 1.4826 \times \mathrm{MAD(z)}$ of the initial red-sequence colour models illustrated in Fig. 2.3 over the entire survey area (19,300 deg^2, Fig. 3.6). Here, MAD(z) is median absolute deviation of the sample about the median for each redshift bin shown by the red nodes and 1.4826 relates this to the standard deviation of a Normal distribution. The majority of members with $z < 0.6$ come from the original validated SPIDERS DR16 sample Sect. 3.1. **Upper panel:** $g - r$ colour. **Lower panel:** panel $r - z$ colour. Below spectroscopic redshifts of 0.35, the red galaxy selection is done in $g - r$ colour and above this point the selection is made in $r - z$ colour. The motivation behind this is to focus on the colours which provide the greatest redshift sensitivity (largest colour slope).

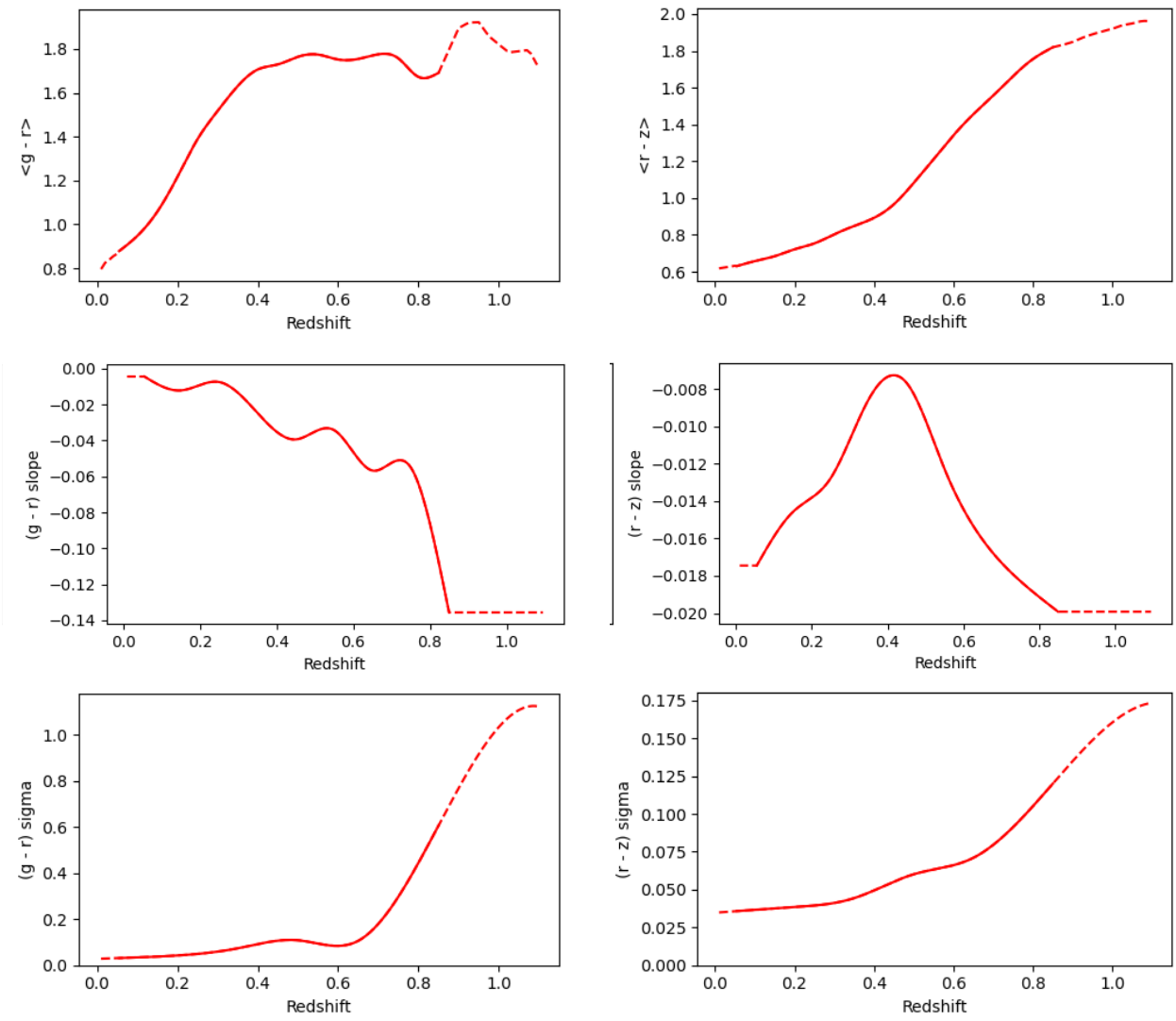

Abbildung 2.5: Adapted from Ider Chitham et al. (2020). Calibrated red-sequence models after a single iteration for $g-r$ (left) and $r-z$ colours (right). This is based on the training set illustrated in Fig. 2.4. With an increasing number of iterations, these distributions become smoother until convergence. **Upper panel**: Colour as a function of redshift $\langle \mathbf{c}|z,m\rangle$ (the magnitude dependency is not shown here). **Middle panel**: The slope *of the width* of the red-sequence. **Lower panel**: The width of the red-sequence.

The probability that a red-sequence galaxy of a given luminosity is at redshift z is proportional to the comoving volume element[5]

$$P_0(z) \propto \frac{dV}{dz} = (1+z)^2 D_A^2(z) c H^{-1}(z). \tag{2.10}$$

Combining Eq. 2.9 and Eq. 2.10 then gives the redMaPPer red-sequence galaxy likelihood

$$\ln \mathcal{L}_{\rm red} = -\frac{\chi^2}{2} + \ln \left| \frac{dV}{dz} \right|. \tag{2.11}$$

The redshift estimator $z_{\rm red}$ is defined as the redshift which maximises the above likelihood. The respective uncertainty on $z_{\rm red}$ is then given by

$$\sigma^2_{z_{\rm red}} = \langle z^2 \rangle - \langle z \rangle^2 = \frac{\int dz \ \mathcal{L}_{\rm red}(z) z^2}{\int dz \ \mathcal{L}_{\rm red}(z)} - \langle z \rangle^2. \tag{2.12}$$

To reduce redshift bias in $z_{\rm red}$ a correction term is computed and applied via

$$z_{\rm red} = z^0_{\rm red} + \langle (z^0_{\rm red} - z_{\rm CG}) | z_{\rm CG} \rangle_{z=z_{\rm red}}. \tag{2.13}$$

This considers the mean redshift offset between $z_{\rm red}$ and the spectroscopic redshift of a central galaxy candidate $z_{\rm CG}$ as a function redshift, evaluated at $z_{\rm red}$ where $z^0_{\rm red}$ is the original, uncorrected redshift estimate[6]. This offset between the corrected and uncorrected galaxy photometric redshifts is illustrated in Fig. 2.6 as a function of spectroscopic (true) redshift for the first iteration of the calibration process of DR8 of The Legacy Surveys (Ider Chitham et al., 2020).

Cluster Redshift Estimation

Measuring the cluster photometric redshift z_λ is analogous to an iterative version of the $z_{\rm red}$ measurement, with i representing the iteration index and $z_{\lambda,i}$ representing the photometric redshift estimate at the $i^{\rm th}$ iteration. For each central galaxy candidate, the following task is performed:

- For the first iteration, $i = 0$ and $z_{\lambda,0} = z_{\rm red}$.
- For subsequent iterations, richness (λ) is calculated and membership probabilities ($p_{\rm mem}$) are measured for each galaxy about the central galaxy candidate. The cluster redshift is then updated as $z_{\rm cluster} = z_{\lambda,i}$.

[5] This assumes that the luminosity function does not evolve over redshift uncertainties.

[6] This equation has been simplified for brevity, however the full correction allows redshift bias to vary with magnitude (see Appendix A.1 of Rykoff et al., 2014).

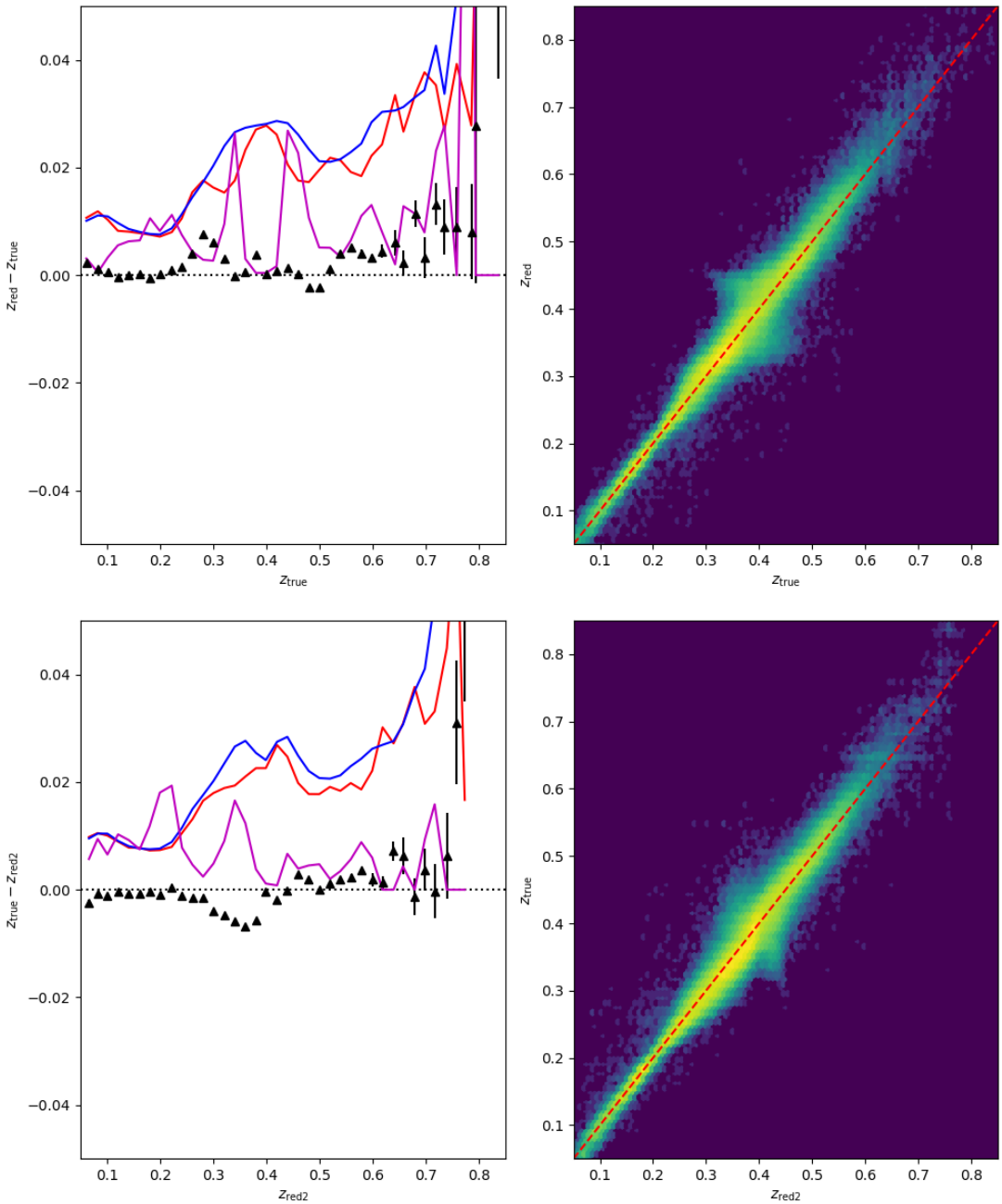

Abbildung 2.6: Adapted from Ider Chitham et al. (2020). **Lower panel:** Uncorrected photometric galaxy redshift ($z^0_{\rm red}$) for member galaxies with $p_{\rm mem} > 0.9$ in The Legacy Surveys DR8 during the first iteration of the redMaPPer calibration, as a function of the central galaxy spectroscopic redshift $z_{\rm CG}$. Here, the notation differs slightly with respect to Eq. 2.13 ($z_{\rm true} = z_{\rm CG}$, $z_{\rm red} = z^0_{\rm red} =$, $z_{\rm red2} = z_{\rm red}$). Each galaxy in the training red-galaxy set (Fig. 2.4) is initially assumed to be a central galaxy of a cluster. Black triangles show the mean redshift offset $z_{\rm red} - z_{\rm CG}$ for each redshift bin. The red line shows the RMS of the redshift offset, the blue line is the mean redshift error and the magenta purple line is the fraction of 4σ outliers as a function of redshift. **Upper panel:** Corrected photometric redshift ($z_{\rm red}$) relative to $z_{\rm CG}$, the RMS, mean redshift uncertainty and outlier fraction are also shown (as in the lower panel) as a function of the corrected redshift.

- The highest probability member galaxies are then used to refine the redshift estimate $z_{\lambda,i+1}$ by maximising the likelihood:

$$\ln\mathcal{L} = \Sigma - \frac{w\,[\chi^2 + \ln|\mathbf{C}|]}{2}. \tag{2.14}$$

  Here, $\ln|\mathbf{C}|$ is the natural logarithm of the determinant of the covariance matrix (see Eq. 2.7) and w is a probability dependent galaxy weight that follows a Fermi-Dirac distribution, varying smoothly from $w(p_{\rm mem} = 1) = 1$ to $w(p_{\rm mem} = 0) = 0$.

- The previous two steps are repeated until the convergence criterion is met: $|z_{\lambda,i+1} - z_{\lambda,i}| < \delta_{z_\lambda}$. In this work $\delta_{z_\lambda} = 0.0002$.

- Once z_λ is estimated, it is then possible to calculate the posterior $P(z_{\rm true}|z_\lambda)$.

$$P(z_{\rm true}|z_\lambda) = \frac{e^{-\frac{\chi^2_{\rm norm}}{2}}\left|\frac{dV}{dz}\right|}{\int d\chi^2_{\rm norm}\left|\frac{dV}{dz}\right|}, \tag{2.15}$$

  where the variable of integration is

$$\chi^2_{\rm norm} = \Sigma w[\chi^2 + \ln|\mathbf{C}|] - \min\left(\Sigma w[\chi^2 + \ln|\mathbf{C}|]\right). \tag{2.16}$$

- In order to ensure that z_λ is unbiased, a correction is applied in an analogous way to what was done for $z_{\rm red}$. To minimise the z_λ bias the following criterion must be achieved $\langle z_{\rm true}|z_\lambda\rangle = z_\lambda$. This correction factor is determined by the bias shown in Fig. 2.7 for each iteration of the calibration process. For further details please refer to Appendix A.2 of Rykoff et al. (2014).

Cluster Ranking

The total cluster likelihood is the sum of the richness likelihood, $\ln\mathcal{L}_\lambda$ and the centring likelihood, $\ln\mathcal{L}_{\rm cen}$

$$\ln\mathcal{L} = \ln\mathcal{L}_\lambda + \ln\mathcal{L}_{\rm cen}. \tag{2.17}$$

Here, the richness likelihood considers the membership probabilities of each cluster galaxy as well as the scale-factor S which accounts for unobserved galaxies. The richness likelihood is evaluated at the cluster photometric redshift z_λ and given by

$$\ln\mathcal{L}_\lambda = -\frac{\lambda}{S} - \sum \ln(1 - p_{\rm mem}), \tag{2.18}$$

The centring likelihood on the other hand considers the luminosity function, $\phi_{\rm cen}$ as well as centring functions $G_{\rm cen}$ and $f_{\rm cen}$ (Rykoff et al., 2012). The centring likelihood is given by

$$\ln\mathcal{L}_{\rm cen} = \ln[\phi_{\rm cen}(m|z_\lambda,\lambda)G_{\rm cen}(z_{\rm red})f_{\rm cen}(w|z_\lambda,\lambda)], \tag{2.19}$$

where the magnitude m, photometric redshift $z_{\rm red}$, and local galaxy density w are combined for each galaxy.

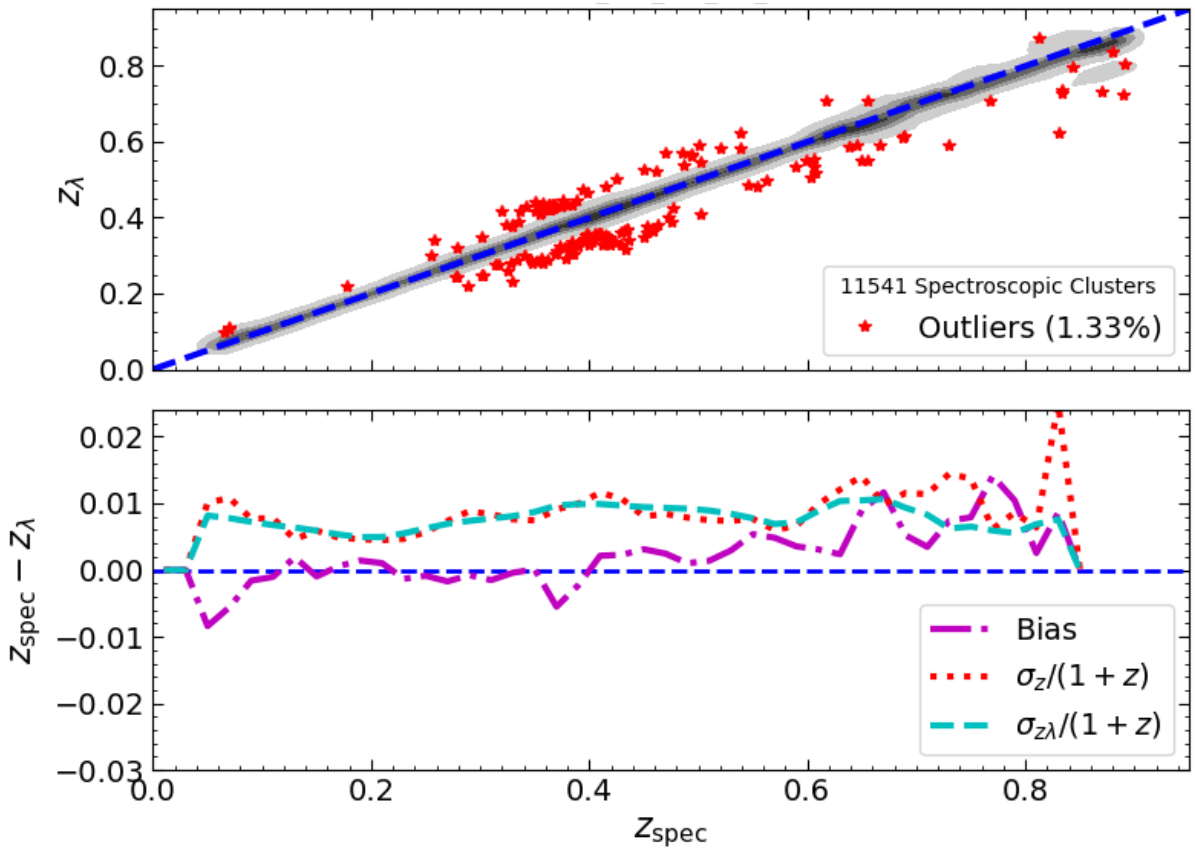

Abbildung 2.7: Adapted from Ider Chitham et al. (2020). **Upper Panel**: z_λ as a function of the spectroscopic redshift of the central cluster galaxy ($z_{\rm CG}$) in The Legacy Surveys DR8 during the first iteration of the redMaPPer calibration (see also Fig. 2.6). Outliers are defined when the discrepancy between spectroscopic and photometric redshifts is more than four times the uncertainty on the photometric redshift ($z_{\rm cg} - z_\lambda > 4\sigma_{z_\lambda}$). **Lower panel:** Median quantities over each redshift bin used in this analysis (bin width of 0.05). The purple line shows the redshift bias: ($z_{CG} - z_\lambda$), the teal line shows the photometric redshift uncertainties: $\sigma_{z_\lambda}/(1 + z_{\rm CG})$ and the red line shows the scatter: $\sigma_z/(1 + z_{\rm CG}) = 1.4862 \times (|z_{\rm CG} - z_\lambda| - (z_{\rm CG} - z_\lambda))/(1 + z_{\rm CG})$.

Scanning-mode

When configured in *scanning-mode*, redMaPPer first estimates the initial photometric redshift estimate of a cluster system given a positional prior. This is carried out by determining the maximum likelihood redshift as illustrated in Fig. 2.8 by evaluating the likelihood at each step on a redshift grid and determining the peak of the profile. When evaluating the relative signal from the redMaPPer optical richness $\lambda(z)$ and redMaPPer optical likelihood $\mathcal{L}(z)$, both distributions are centred around the spectroscopic (and photometric) redshifts. However, the likelihood distribution is narrower about the spectroscopic cluster redshift, indicating that it out performs richness as a way to optically rank clusters. This is to be expected as the total likelihood is a combination of richness *and* centring information (Eq. 4.22). To test the robustness of this cluster likelihood, redMaPPer redshift scans are carried out at the position of each system in a compilation of X-ray, optical and SZ selected archival cluster candidates with known redshifts (Ider Chitham et al. in prep.) shown in Fig. 2.9. The fraction of cluster candidates in bins of the deviation between the literature redshift and the redMaPPer photometric redshift is strong function of the optical cluster likelihood as shown in Fig. 2.10. In other words, the more highly ranked the cluster is, the closer the redMaPPer redshift resembles the literature redshift.

2.2.3 Multi-Component Matched-filter Cluster Confirmation

This subsection describes the application of a modified version of the Multi-Component Matched-filter (MCMF) Cluster Confirmation used to acquire spectroscopic targets for SDSS-IV using Pan-STARRS1 DR1 photometry (Clerc et al. (2020), Ider Chitham et al in prep.). This application of MCMF is discussed in detail in Sect. 3.2, although is introduced here due to its similarity to redMaPPer. The primary difference between the scanning-mode functionality of redMaPPer and MCMF is that redMaPPer determines the maximum likelihood redshift (as a prior for the photometric cluster redshift), while MCMF determines the redshift corresponding to the maximum richness. In the case of multiple richness peaks MCMF chooses the one which corresponds to the minimum contamination fraction (estimate of chance alignment).

The red-sequence calibration process within MCMF utilises models in the similar format to those described in Sect. 2.2.1, however the method of training is different. Given a spectroscopic training set of clusters[7], the initial red-sequence models are calibrated in bins of redshift by stacking the photometric distribution of galaxies in the colour-magnitude plane for galaxies within R_{500} of each cluster centre. The width, slope and intercept is then be measured relative to the input models (after statistical background subtraction, e.g. Hennig et al., 2017). Redshift and magnitude dependent correction factors are then applied to update the models in preparation for the next iteration. This process is repeated until the models converge (after three iterations). In order to measure the richness as a function of redshift, each source is assigned a colour-weight depending on the similarity to

[7]Spectroscopic training clusters do not necessarily need to be optically selected or even correspond to a central galaxy.

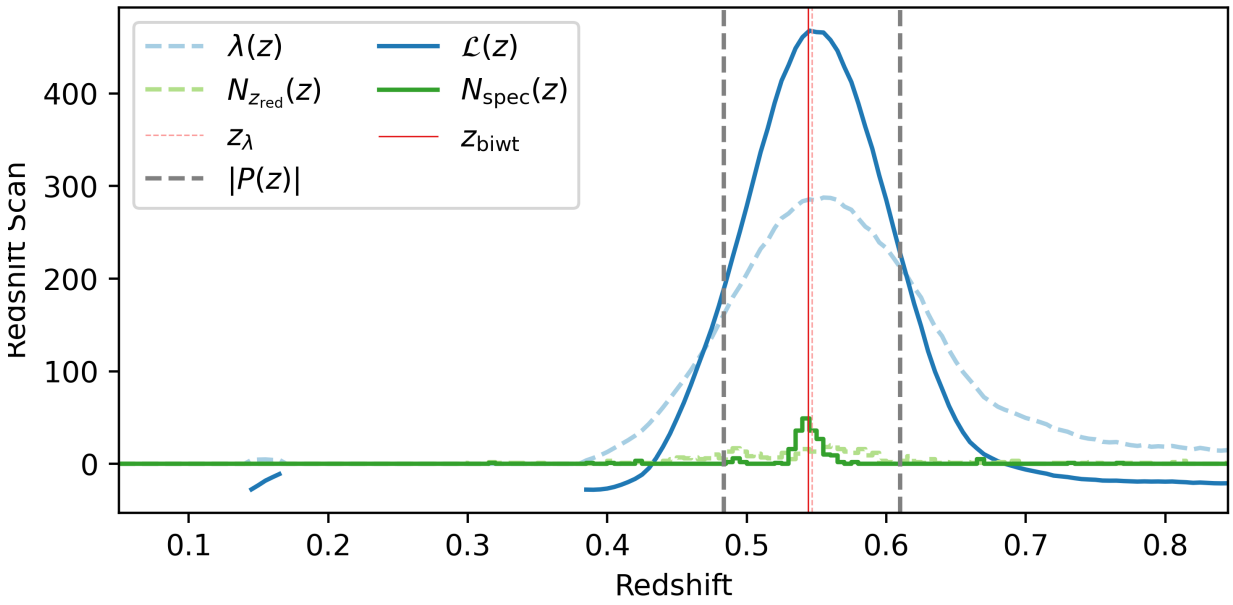

Abbildung 2.8: Adapted from Ider Chitham et al. in prep. Redshift scan from a single X-ray selected cluster produced by redMaPPer configured in *scanning-mode*. The left-hand column of legend corresponds to photometric quantities and the right hand column corresponds to equivalent spectroscopic quantities. The blue lines show redshift-scans of the optical cluster likelihood $\mathcal{L}(z)$ and optical richness $\lambda(z)$. The green lines represent the distribution of photometrically ($N_{z_{\rm red}}$) and spectroscopically ($N_{z_{\rm spec}}$) selected members by redMaPPer (Sect. 2.2.2) and the automated spectroscopic algorithm (Sect. 2.3.1) respectively. Vertical dashed lines show the redshift range over which the photometric cluster redshift probability distribution is considered $|P(z)|$. The red vertical line indicates the bi-weight spectroscopic ($z_{\rm biwt}$) and photometric (z_λ) cluster redshift which are close to (but not exactly the same as) the peak of the redshift scan. Note that the cluster likelihood is occasionally negative when the centring likelihood is negative i.e. ($\mathcal{L}_{cen} < 0$ and $|\mathcal{L}_{cen}| < |\mathcal{L}_\lambda|$, as per Eq. 4.22).

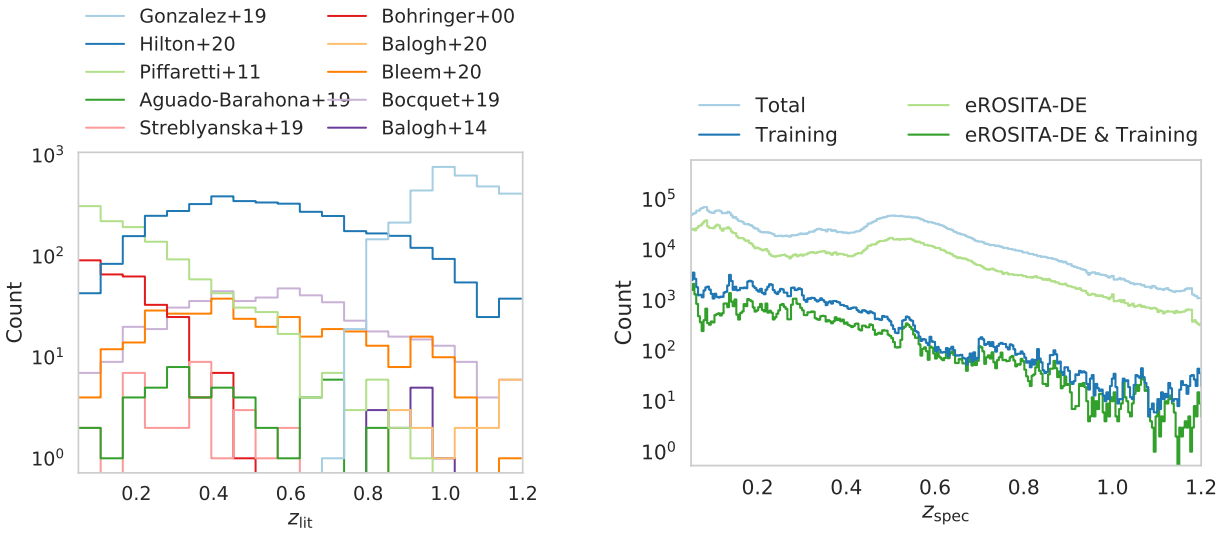

Abbildung 2.9: Adapted from Ider Chitham et al. in prep. **Left:** The literature redshift ($z_{\rm lit}$) distribution of archival cluster candidates over the footprint of The Legacy Surveys (Sect. 3.4). This sample consists of the SPT 2500 deg^2 (Bocquet et al., 2019), SPTPol Extended (Bleem et al., 2020), ACT DR5 (Hilton et al., 2020), MADCOWS (Gonzalez et al., 2019), MCXC (Piffaretti et al., 2011), PLANCK SZ2 (Streblyanska et al., 2019; Aguado-Barahona et al., 2019), GEEC2 (Balogh et al., 2014), GOGREEN + GCLASS (Balogh et al., 2020), NORAS (Böhringer et al., 2000) and cluster samples. **Right:** The distribution of spectroscopic in the archival compilation used in spectroscopic post-processing of SPIDERS cluster redshifts tabulated in Table. 2.1 and illustrated on sky in Fig. 2.13 for sources which constitute the total compilation in the *eROSITA*-DE hemisphere and full sky, as well as the subsets which contribute towards the training set used to calibrate redMaPPer (known spectroscopic members in the literature).

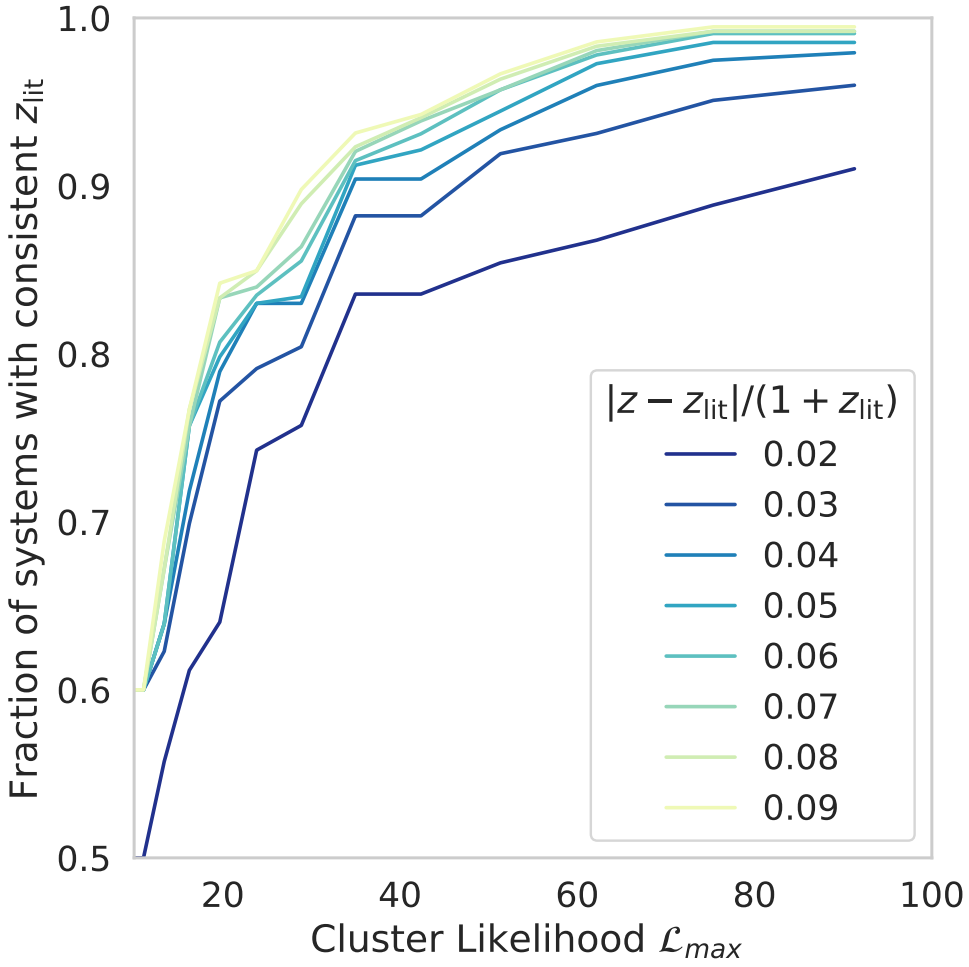

Abbildung 2.10: Adapted from Ider Chitham et al. in prep. The relation between the redMaPPer optical cluster likelihood and the consistency between measured redshifts and their literature value from known cluster candidates in the archive (summarised in Fig. 2.9). Only the $z_{\rm lit} < 0.85$ subset is considered here, although no other cuts or cleaning methods are applied to any of the original catalogues (as they are *candidates*). The solid lines show fraction of literature clusters within bins of absolute deviation between literature and best available redshift in the automated pipeline i.e. $|z - z_{\rm lit}|/(1 + z_{\rm lit})$ as a function of redMaPPer cluster likelihood (Eq. 2.17). The redshift z is the spectroscopic redshift (Sect. 2.3.1) or the redMaPPer scanning-mode photometric redshift (Sect. 2.2.2) if no spectra are available.

the calibrated red-sequence models at each redshift step in the redshift scan (Klein et al., 2018, 2019).

$$w_i(z) = \frac{\prod_{j=1}^{3} G\left(c_{i,j} - \langle c(z)\rangle_j, \sigma_{c_j}(z)\right)}{N(\sigma_{c_1}(z), \sigma_{c_2}(z), \sigma_{c_3}(z))}. \tag{2.20}$$

Here, $c_{i,j}$ is the observed j^{th} colour of the i^{th} galaxy and $\langle c(z)\rangle_j$ is the expectation value of the j^{th} colour according to the calibrated models. G represents a normalised Gaussian function and $\sigma_{c_j}(z)$ denotes the standard deviation. The normalisation factor, $N(\sigma_{c_1}(z), \sigma_{c_2}(z), \sigma_{c_3}(z))$, is an average weight of a population of galaxies that follows the expected distribution of a cluster at the given redshift. The MCMF galaxy luminosity and radial weights are almost identical to those presented in Eq. 2.3. For full details, please refer to Klein et al. (2018, 2019).

2.3 Spectroscopic redshifts

If the number of spectroscopic members (cluster tracer particles) is large, one can then study the dynamical properties of individual clusters in detail (Biviano et al., 2006; Mamon et al., 2013; Old et al., 2018). For the subset of systems where the cluster galaxy population is sufficiently sampled, it becomes feasible to estimate velocity dispersions, identify substructure (Yu et al., 2015) and measure caustic masses (Maughan et al., 2016). This provides complementary mass calibration methods to weak lensing and X-ray gas diagnostics (Gifford et al., 2017). Dynamical mass estimates become increasingly unbiased with larger numbers of spectroscopically confirmed members (Mamon et al., 2013), it is therefore desirable to measure at least ten spectroscopic member redshifts per cluster. However, for the richest, most massive and most optically extended clusters at low-redshift it is feasible to obtain more than a hundred tracers per system. If the number of galaxy clusters with spectroscopic redshifts in a sample is large, it becomes statistically feasible to quantify projection effects[8] (Myles et al., 2020) and to measure gravitational redshifts predicted by General Relativity (Wojtak et al., 2011; Mpetha et al., 2021).

This section introduces the notation relating to spectroscopic cluster redshifts and velocity dispersion measurements used in the cosmological analyses presented in this thesis (Sects. 4.1 & 4.2). The automated algorithm Sect. 2.3.1 has been adapted from the work of Clerc et al. (2016); Capasso et al. (2019b); Ferragamo et al. (2020) and is the basis of automatic spectroscopic redshift pipeline used within the SPIDERS cluster programme (Ider Chitham et al., 2020; Clerc et al., 2020; Kirkpatrick et al., 2021, Ider Chitham in prep.). The algorithm itself is iterative, with the subscript k referring to the k^{th} iteration of the procedure.

[8]Spectroscopic redshifts are used to disentangle substructures and field galaxies that are aligned along the line-of-sight.

2.3.1 Automatic membership

The zeroth iteration of the algorithm ($k = 0$) utilises $N_{\rm spec,0}$ spectroscopic member galaxies to initialise the spectroscopic cluster redshift. This is provided by the biweight location estimate (Beers et al., 1990, Eq. 2.21) prior to the iterative clipping procedure

$$z_{\rm bwt} = \hat{z} + \frac{\sum_{|u_j|<1} (z_j - \hat{z})[1 - u_j^2(z, z_j, 6.0)]^2}{\sum_{|u_j|<1} [1 - u_j^2(z, z_j, 6.0)]^2} \tag{2.21}$$

where z is a vector of galaxy redshifts, $\hat{z}$ is the sample median, j is the index of the spectroscopic member galaxy and u_j is given by

$$u_j(z, z_j, a) = \frac{(z_j - \hat{z})}{a \times {\rm MAD}(z)} \tag{2.22}$$

Here, a is the turning constant and ${\rm MAD}(z)$ is the median absolute deviation of the redshifts of the member galaxies. If $z_{\rm MAD} = 0$, the median redshift is returned. The proper line-of-sight velocity offset (Danese et al., 1980, Eq. 2.23), of all member galaxies are then computed relative to estimate of the cluster redshift

$$\frac{v_j}{c} = \frac{z_j - z_{\rm bwt}}{1 + z_{\rm bwt}} \tag{2.23}$$

Members with $|v_j| > 5000$ km/s are subject to the initial velocity clipping and rejected from the spectroscopic sub sample of member galaxies. This results in $N_{\rm spec,1}$ spectroscopic members for the first iteration ($k = 1$) which are used to recompute the biweight cluster redshift Eq. 2.21. At this stage the velocity dispersion (Beers et al., 1990) is also calculated using the biweight scale estimator (Eq. 2.24, if $N_{\rm spec,1} \geq 15$) or the gapper estimator (Eq. 2.26, if $N_{\rm spec,1} < 15$).

The biweight scale estimator (Tukey, 1958) is defined[9] as

$$\sigma_{\rm bwt}(N_{\rm spec}) = \sqrt{N_{\rm spec}} \frac{\sqrt{\sum_{|u_j|<1} (v_i - \hat{v})^2 (1 - u_j^2)^4}}{|(\sum_{|u_j|<1} (1 - u_j^2)(1 - 5u_j^2))|} \tag{2.24}$$

with an estimated uncertainty of

$$\Delta_{\sigma_{\rm bwt}} = \frac{0.92 \times \sigma_{\rm bwt}}{\sqrt{N_{\rm spec} - 1}} \tag{2.25}$$

where $u_i = u_j(v, v_j, 9.0)$.

The gapper estimator (Wainer & Thissen, 1976) is based on the gaps of an order statistics, $x_j, x_{i+1}, \ldots, x_n$. It is defined as a weighted average of gaps:

$$\sigma_{\rm gap}(N_{\rm spec}) = \frac{\sqrt{\pi}}{N_{\rm spec}\,(N_{\rm spec} - 1)} \sum_{j=1}^{N_{\rm spec}-1} w_j\, g_j, \tag{2.26}$$

[9] Strictly the prefix is $\sqrt{\left(\frac{N_{\rm spec}^2}{N_{\rm spec}-1}\right)}$ although $\sqrt{N_{\rm spec}}$ is an adequate approximation for large $N_{\rm spec}$.

where weights and gaps are given by w_j and g_j respectively

$$w_j = j\,(N_{\rm spec} - 1), \qquad g_j = x_{j+1} - x_j. \tag{2.27}$$

Spectroscopic galaxies with $v_j > 3\sigma$ are rejected as outliers during a further σ clipping process. This procedure is iterated until convergence (no outlier galaxies remain) up to a maximum of ten iterations ($k \leq 10$).

There are several different possible outcomes which can occur for each cluster as a result of automated spectroscopic redshift estimation and member selection. They are as follows:

- $N_{\rm spec,1} = 0$: the initial 5000km/s clipping rejected all members: the procedure cannot proceed and a flag is issued to indicate that convergence failed. This can occur for true systems when there are several distinct structures along the line-of-sight which are far apart in redshift space.
- $0 < N_{{\rm spec},k} < 3$: There are an insufficient number of spectroscopic members left after k steps: it is not possible to estimate the biweight redshift estimate (Eq. 2.21) or the velocity dispersion (Eq. 2.26) and a flag is issued accordingly. These cluster is left for visual inspection if at least one redshift is available.
- $N_k^{zspec} \geq 3$: the process successfully converges, the cluster redshift and velocity dispersion are estimated.

In the cased of the final item, the remaining objects are called spectroscopic members and the final spectroscopic cluster redshift is catalogued with a statistical uncertainty Δ_z determined by the bootstrap resampling of the $N_{\rm spec}$ velocities (Ruel et al., 2014).

$$\Delta_z({\rm standard}) = \frac{1}{c}\frac{\sigma \cdot (1+z)}{\sqrt{N_{\rm mem}}} \tag{2.28}$$

2.3.2 Archival redshifts

In order to maximise the number of spectroscopic cluster redshifts that can be obtained using the SPIDERS pipeline described above, it is necessary to utilise as many spectroscopic galaxies as possible. For this reason, a large archival compilation of spectroscopically observed galaxies has been created to maximise the number of redMaPPer selected cluster galaxies which have readily available spectroscopic redshifts. This is also extremely useful for optimising spectroscopic observational programmes to prioritise fibres for galaxies which do not already have a sufficiently accurate redshifts. This prevents one from unnecessarily observing the same objects multiple times, thereby increasing the overall spectroscopic completeness of the member catalogue. An additional benefit of such a large compilation of spectroscopic galaxies is that selecting previously confirmed cluster members yields a valuable training set which can be used to accurately calibrate the initial models of the red-sequence (Sect. 2.2.1). The minimum required number of spectroscopic cluster redshifts is approximately 50 per redshift bin of width 0.05. For most wide area surveys, this is

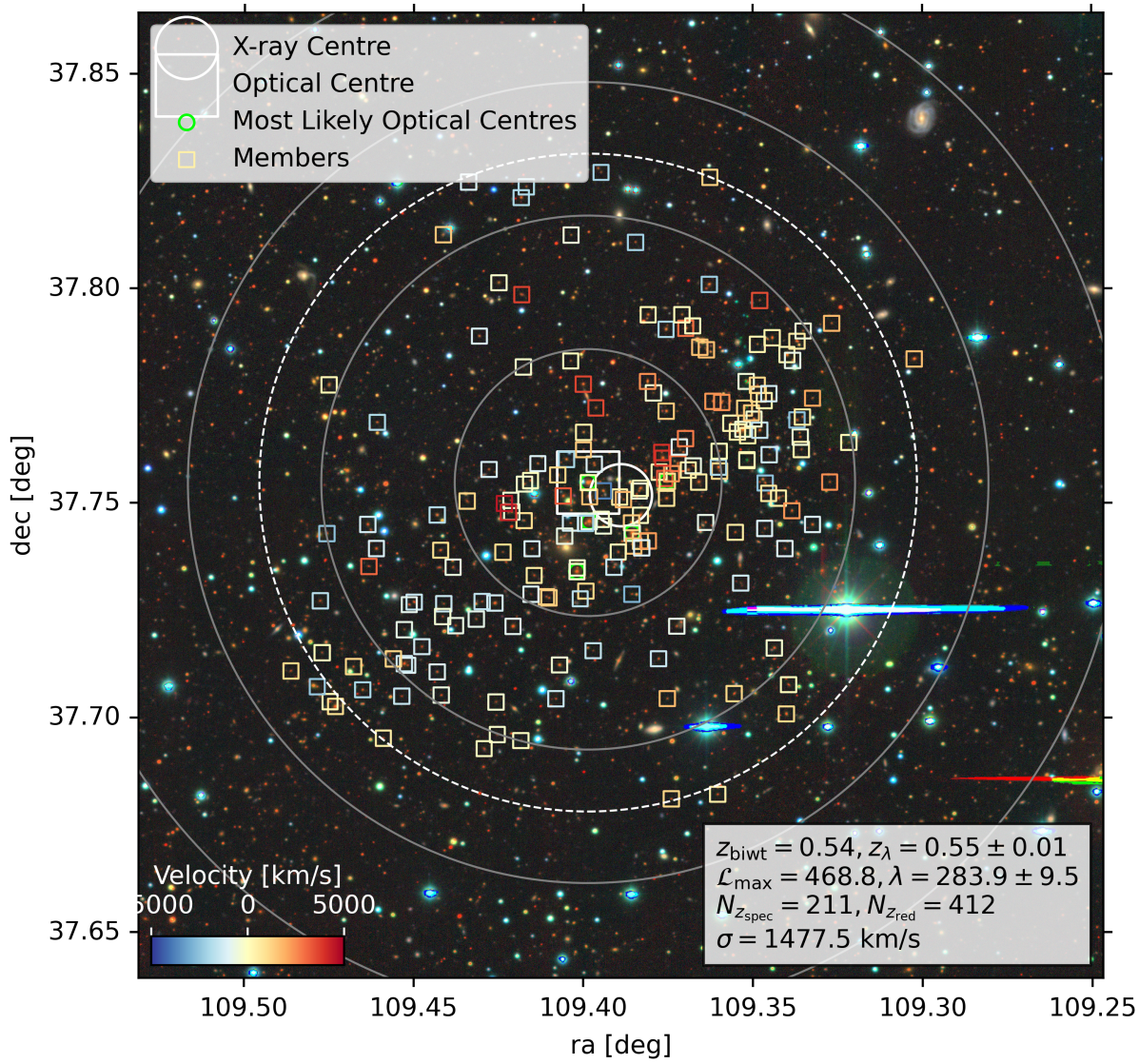

Abbildung 2.11: Adapted from Ider Chitham et al. in prep. The CODEX cluster with the highest number of spectroscopic members at a spectroscopic redshift of $z_{\mathrm{biwt}} = 0.54$, photometric redshift $z_\lambda = 0.55 \pm 0.01$ and richness $\lambda = 283.9 \pm 9.5$. Of the 412 redMaPPer selected members ($N_{z_{\mathrm{red}}}$), 211 have redshifts from archival optical spectroscopy ($N_{z_{\mathrm{spec}}}$, Sect. 2.3.2). 186 are then selected by the automated spectroscopic membership algorithm (N_{mem}, Sect. 2.3.1). These members are indicated by coloured squares, with the hue representing the radial velocity offset relative to the spectroscopic bi-weight cluster redshift (Eq. 2.21), i.e. redshifted or blue-shifted. Concentric grey annuli correspond to intervals of 0.5Mpc, while the white dashed annulus corresponds to the redMaPPer cut-off radius (R_λ, Eq. 2.6). The redMaPPer redshift scan of this system is depicted in Fig. 2.8 along with the redshift distribution of photometrically and spectroscopically selected galaxies. The velocity dispersion (Eq. 2.21) of this system is $\sigma = 1477.5$ km/s and the phase space distribution of its spectroscopic members are shown in Fig. 2.12 using exactly the same radial markers and velocity offset colour scheme.

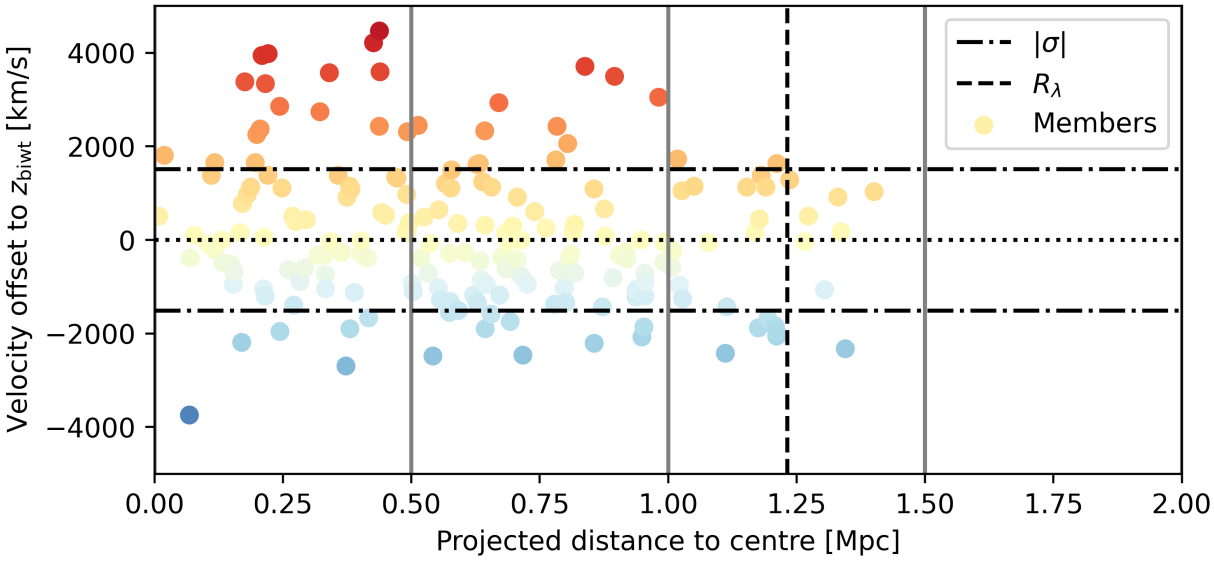

Abbildung 2.12: Adapted from Ider Chitham et al. in prep. Phase space plot for 211 spectroscopic members of the CODEX cluster shown in Fig. 2.11. This illustrates the velocity and radial distributions relate of spectroscopic cluster galaxies after applying the automated selection described in Sect. 2.3.1. The solid and dashed vertical lines represent intervals of 0.5Mpc, about the optical centre of the cluster and the white dashed annulus corresponds to the redMaPPer cut-off radius (R_λ, Eq. 2.6). These correspond to the grey annuli and white annuli in Fig. 2.11 respectively. The colour of the members emphasises how red-shifted or blue-shifted each galaxy is relative to the bi-weight spectroscopic cluster redshift (Eq. 2.21).

achieved with the use of BCGs alone at low-redshifts, however, at above redshifts of 0.7, the number of accurately targeted cluster galaxies with good redshifts diminishes substantially. Therefore, in order to meet to the desired number of spectroscopically confirmed cluster galaxies in this regime, and avoid unwanted biases in the resultant photometric redshifts, it is vital to use as many sources from the literature as possible.

Tabelle 2.1: Adapted from Ider Chitham et al. in prep. Archival redshifts from the SPIDERS spectroscopic compilation. This is a tabular representation of the sky distributions shown in Fig. 2.13. N denotes the number of spectroscopic redshifts from each reference source. Duplicate spectroscopic redshifts are removed by enforcing one unique object per `NSIDE`=2^{18} healpixel with priority given to objects with more precise redshift uncertainties. This corresponds to a resolution of roughly 0.81 arcseconds. $f_{\rm DE}$ indicates the fraction of spectroscopic redshifts which are located in in the *eROSITA*-DE hemisphere and $f_{\rm train}$ indicates the fraction that contribute to the training set of spectroscopic cluster members. Here, the spectroscopic redshift range is truncated at $z_{\rm max} = 1.50$. Please note that the Zou et al. (2019) source is also from compilation of several spectroscopic surveys: 2dFGRS (Colless et al., 2001), 2SLAQ (Cannon et al., 2006), 6dFGS (Jones et al., 2004, 2009), CFRS (Lilly et al., 1995), CNOC2 (Yee et al., 2000), DEEP2 (Davis et al., 2003; Newman et al., 2013), SDSS DR14 (Abolfathi et al., 2018), VIPERS (Garilli et al., 2014; Guzzo et al., 2014), VVDS (Le Fèvre et al., 2005; Garilli et al., 2008), WiggleZ (Drinkwater et al., 2010; Parkinson et al., 2012), and zCOSMOS (Lilly et al., 2007).

Source	N	$z_{\rm min} - z_{\rm max}$	$z_{\rm median}$	$f_{\rm DE}$	$f_{\rm Train}$
Alam et al. (2015)	2716279	0.00 - 1.50	0.35	0.37	0.00
Zou et al. (2019)	1119901	0.00 - 1.50	0.55	0.39	0.00
Colless et al. (2001)	216116	0.00 - 1.46	0.11	0.70	0.00
Coil et al. (2011); Cool et al. (2013)	175644	0.00 - 1.50	0.57	0.49	0.00
Jones et al. (2009)	100472	0.00 - 1.46	0.05	0.78	0.00
Liske et al. (2015)	63958	0.00 - 1.28	0.17	1.00	0.00
Drinkwater et al. (2010)	41212	0.00 - 1.50	0.59	0.57	0.00
Robotham et al. (2011)	33675	0.00 - 0.58	0.19	0.76	1.00
Le Fèvre et al. (2013)	33337	0.00 - 1.50	0.62	0.38	0.00
Garilli et al. (2014)	32803	0.00 - 1.50	0.71	0.00	0.00
Davis et al. (2003); Newman et al. (2013)	31031	0.00 - 1.50	0.86	0.00	0.00
Clerc et al. (2020)	25840	0.01 - 0.60	0.20	0.16	0.50
Huchra et al. (2012)	23088	0.00 - 0.52	0.08	0.40	0.00
Shectman et al. (1996)	22674	0.01 - 0.96	0.29	0.95	0.00
Lilly et al. (2009)	18431	0.00 - 1.49	0.53	1.00	0.00
Childress et al. (2017)	12101	0.00 - 1.50	0.54	0.65	0.00
Mulroy et al. (2019)	10223	0.15 - 0.30	0.23	0.40	1.00
Rykoff et al. (2016)	7720	0.02 - 1.28	0.33	0.41	0.44
Skelton et al. (2014); Momcheva et al. (2016)	6818	0.02 - 1.50	0.89	0.39	0.00

Source	N	$z_{\min} - z_{\max}$	z_{median}	f_{DE}	f_{Train}
Moretti et al. (2017)	6720	0.04 - 0.08	0.06	0.73	1.00
Pimbblet et al. (2006)	6044	0.00 - 1.47	0.40	0.60	0.00
Rykoff et al. (2012)	5827	0.03 - 0.56	0.21	0.41	1.00
Cooper et al. (2012)	5684	0.01 - 1.43	0.62	1.00	0.00
Connelly et al. (2012)	4682	0.01 - 1.46	0.42	0.00	0.00
Golovich et al. (2019)	4286	0.01 - 1.41	0.25	0.35	0.00
Masters et al. (2017, 2019)	3994	0.01 - 1.50	0.81	0.65	0.00
Caminha et al. (2017)	3602	0.03 - 1.50	0.41	1.00	0.00
Cava et al. (2009)	3356	0.03 - 0.07	0.05	0.60	1.00
Wilson et al. (2016)	3052	0.03 - 0.77	0.30	0.64	1.00
Maturi et al. (2019)	3010	0.00 - 1.08	0.21	0.90	1.00
Rines et al. (2018)	2915	0.13 - 0.26	0.21	0.38	1.00
Yang et al. (2018)	2795	0.01 - 1.42	0.10	0.00	0.00
Rines et al. (2016)	2610	0.04 - 0.20	0.14	0.09	1.00
Richard et al. (2021)	2197	0.00 - 1.49	0.44	0.55	0.00
Owers et al. (2011)	2182	0.01 - 0.90	0.16	0.00	1.00
Ruel et al. (2014)	2024	0.00 - 1.48	0.51	1.00	1.00
Khabibullina & Verkhodanov (2009)	1949	0.27 - 1.49	0.44	0.42	0.00
Sohn et al. (2018b)	1915	0.00 - 0.99	0.30	0.00	1.00
Guglielmo et al. (2018)	1849	0.03 - 0.50	0.18	0.00	1.00
Bayliss et al. (2016)	1793	0.07 - 1.24	0.53	1.00	1.00
Lemze et al. (2013)	1791	0.00 - 1.07	0.18	0.66	1.00
Ebeling et al. (2014)	1653	0.07 - 1.35	0.54	1.00	1.00
Koranyi & Geller (2002)	1547	0.01 - 0.88	0.10	0.55	0.00
Hilton et al. (2020)	1422	0.04 - 1.48	0.46	0.53	1.00
Straatman et al. (2018)	1212	0.35 - 1.18	0.74	1.00	0.00
Bradshaw et al. (2013); McLure et al. (2013)	1194	0.00 - 1.50	1.02	0.00	0.00
Bayliss et al. (2017)	1141	0.00 - 1.33	0.42	1.00	1.00
Carrasco et al. (2017)	991	0.21 - 1.02	0.43	0.30	0.00
Sifón et al. (2013)	944	0.27 - 1.08	0.47	1.00	1.00
Crawford et al. (2011)	791	0.04 - 1.48	0.61	0.39	0.00
Psychogyios et al. (2020)	790	0.11 - 0.23	0.17	0.27	1.00
Wen & Han (2018)	764	0.70 - 0.99	0.77	0.24	1.00
Foëx et al. (2017)	751	0.21 - 0.31	0.29	0.59	0.00
Sohn et al. (2019a)	751	0.07 - 0.09	0.08	0.00	1.00
Dawson (2013)	708	0.02 - 1.32	0.53	1.00	1.00
McClintock et al. (2018)	700	0.10 - 0.91	0.49	0.18	0.42
Sohn et al. (2018a)	679	0.03 - 0.38	0.23	0.00	1.00
Balogh et al. (2020)	669	0.85 - 1.47	1.11	0.36	1.00

Source	N	$z_{\rm min} - z_{\rm max}$	$z_{\rm median}$	$f_{\rm DE}$	$f_{\rm Train}$
Caminha et al. (2019)	628	0.10 - 1.45	0.48	0.50	1.00
Smith et al. (2000)	612	0.00 - 0.27	0.07	0.57	0.00
Lilly et al. (1995)	563	0.01 - 1.35	0.57	0.24	0.00
Inami et al. (2017)	518	0.00 - 1.49	0.92	1.00	0.00
Oguri et al. (2018)	494	0.09 - 1.04	0.32	0.80	1.00
Czoske et al. (2001)	440	0.06 - 1.39	0.39	0.00	0.00
Liu et al. (2018)	436	0.08 - 0.10	0.09	0.00	1.00
Dressler et al. (1999)	423	0.36 - 0.58	0.41	0.05	1.00
Amodeo et al. (2018)	409	0.16 - 0.83	0.38	0.18	1.00
Lidman et al. (2020)	402	0.00 - 1.22	0.43	0.60	1.00
Morris et al. (2007)	340	0.01 - 1.41	0.40	1.00	0.00
Agulli et al. (2016)	330	0.05 - 0.06	0.06	0.00	1.00
Meusinger et al. (2020)	318	0.00 - 0.15	0.02	0.00	1.00
Pranger et al. (2014)	305	0.03 - 0.32	0.10	0.00	1.00
Owen et al. (1995)	283	0.02 - 1.05	0.07	0.27	0.00
Olsen et al. (2005)	266	0.08 - 1.13	0.58	1.00	1.00
Garilli et al. (2021)	237	0.43 - 1.49	1.14	0.48	0.00
Rines et al. (2000)	225	0.01 - 0.31	0.12	0.00	0.00
Stanford et al. (2014)	220	0.74 - 1.28	0.99	0.17	1.00
Rudnick et al. (2017)	217	0.42 - 0.81	0.59	1.00	1.00
Jørgensen et al. (2017)	202	0.08 - 1.02	0.20	0.63	0.00
Golovich et al. (2017)	190	0.07 - 1.11	0.11	0.00	0.00
Deshev et al. (2017)	189	0.19 - 0.21	0.20	1.00	1.00
Verdugo et al. (2008)	170	0.04 - 0.30	0.25	0.46	1.00
Coe et al. (2012)	160	0.22 - 0.23	0.23	0.00	1.00
Kluge et al. (2020)	159	0.02 - 0.15	0.06	0.26	1.00
Dawson et al. (2015)	149	0.00 - 0.71	0.19	0.00	0.00
Nilo Castellón et al. (2014)	145	0.18 - 0.79	0.48	0.34	1.00
Hasegawa et al. (2000)	141	0.00 - 0.27	0.09	0.43	0.00
Guennou et al. (2014)	137	0.10 - 1.08	0.79	1.00	0.00
Tran et al. (2007)	135	0.80 - 0.85	0.83	1.00	1.00
Olsen et al. (2003)	127	0.10 - 0.26	0.24	0.72	1.00
Sohn et al. (2019b)	114	0.08 - 0.99	0.41	1.00	1.00
Silverman et al. (2015)	109	0.65 - 1.50	1.46	1.00	0.00
Castagné et al. (2012)	104	0.05 - 0.29	0.14	1.00	0.00
Eisenhardt et al. (2008)	104	1.05 - 1.42	1.11	0.00	1.00
Stott et al. (2008)	102	0.04 - 0.83	0.18	0.25	1.00
Moran et al. (2005)	100	0.38 - 0.41	0.40	0.00	1.00
Halliday et al. (2004)	100	0.08 - 1.17	0.74	1.00	0.00

Source	N	$z_{\min} - z_{\max}$	z_{median}	f_{DE}	f_{Train}
Fossati et al. (2019)	99	0.00 - 1.47	0.88	1.00	1.00
Bezanson et al. (2018)	96	0.60 - 1.02	0.74	1.00	1.00
Tanaka et al. (2009)	91	0.25 - 1.33	1.18	1.00	0.00
Kirk et al. (2015)	90	0.30 - 0.56	0.35	0.58	1.00
Postman et al. (2001)	86	0.75 - 0.94	0.90	0.00	1.00
Hansen et al. (2002)	84	0.11 - 0.36	0.18	1.00	1.00
McDonald et al. (2016)	84	0.28 - 1.22	0.60	0.99	1.00
Deger et al. (2018)	76	0.46 - 0.80	0.60	1.00	1.00
Ferrari et al. (2003)	75	0.28 - 1.10	0.74	1.00	0.00
Pierre et al. (2016)	74	0.04 - 1.05	0.32	0.54	1.00
Ricci et al. (2018)	73	0.03 - 1.06	0.38	0.00	1.00
Zaznobin et al. (2020)	67	0.09 - 0.61	0.27	0.40	1.00
Sluse et al. (2019)	65	0.06 - 0.94	0.48	1.00	1.00
Lee et al. (2019)	51	0.84 - 1.32	0.88	0.00	1.00
McLure et al. (2017)	48	0.66 - 1.43	1.22	0.35	0.00
Khullar et al. (2019)	44	1.25 - 1.49	1.30	1.00	1.00
Willis et al. (2013)	39	0.80 - 1.03	0.86	0.00	1.00
Demarco et al. (2007)	37	1.23 - 1.25	1.24	1.00	1.00
Song et al. (2017)	36	0.02 - 0.27	0.03	0.00	1.00
Schirmer et al. (2011)	35	0.43 - 0.46	0.44	0.00	1.00
Gonzalez et al. (2019)	35	0.76 - 1.46	1.05	0.14	1.00
Muzzin et al. (2009)	35	1.16 - 1.22	1.18	0.00	1.00
Demarco et al. (2010)	35	0.86 - 1.21	1.15	0.00	1.00
Liu et al. (2012)	32	0.13 - 0.36	0.23	0.47	0.00
Runge & Yan (2018)	32	0.08 - 0.49	0.16	0.41	1.00
Rescigno et al. (2020)	29	0.19 - 0.22	0.20	0.00	1.00
Connor et al. (2019)	26	0.25 - 1.20	0.97	1.00	1.00
Streblyanska et al. (2019)	25	0.07 - 0.62	0.37	0.08	1.00
Bayliss et al. (2011)	25	0.26 - 0.64	0.44	0.44	1.00
Mirkazemi et al. (2015)	25	0.04 - 0.62	0.24	0.32	1.00
Jørgensen et al. (2014)	24	1.26 - 1.28	1.27	0.00	1.00
Aguado-Barahona et al. (2019)	24	0.07 - 0.82	0.42	0.38	1.00
Strait et al. (2018)	22	0.73 - 1.28	1.06	0.00	0.00
Foley et al. (2011)	18	1.11 - 1.14	1.13	1.00	1.00
Zeimann et al. (2013)	13	1.06 - 1.46	1.26	0.00	1.00
Wilson et al. (2009)	10	1.32 - 1.35	1.33	1.00	1.00
Castignani et al. (2020)	5	0.21 - 0.57	0.44	0.60	1.00
Stalder et al. (2013)	1	1.32 - 1.32	1.32	1.00	1.00

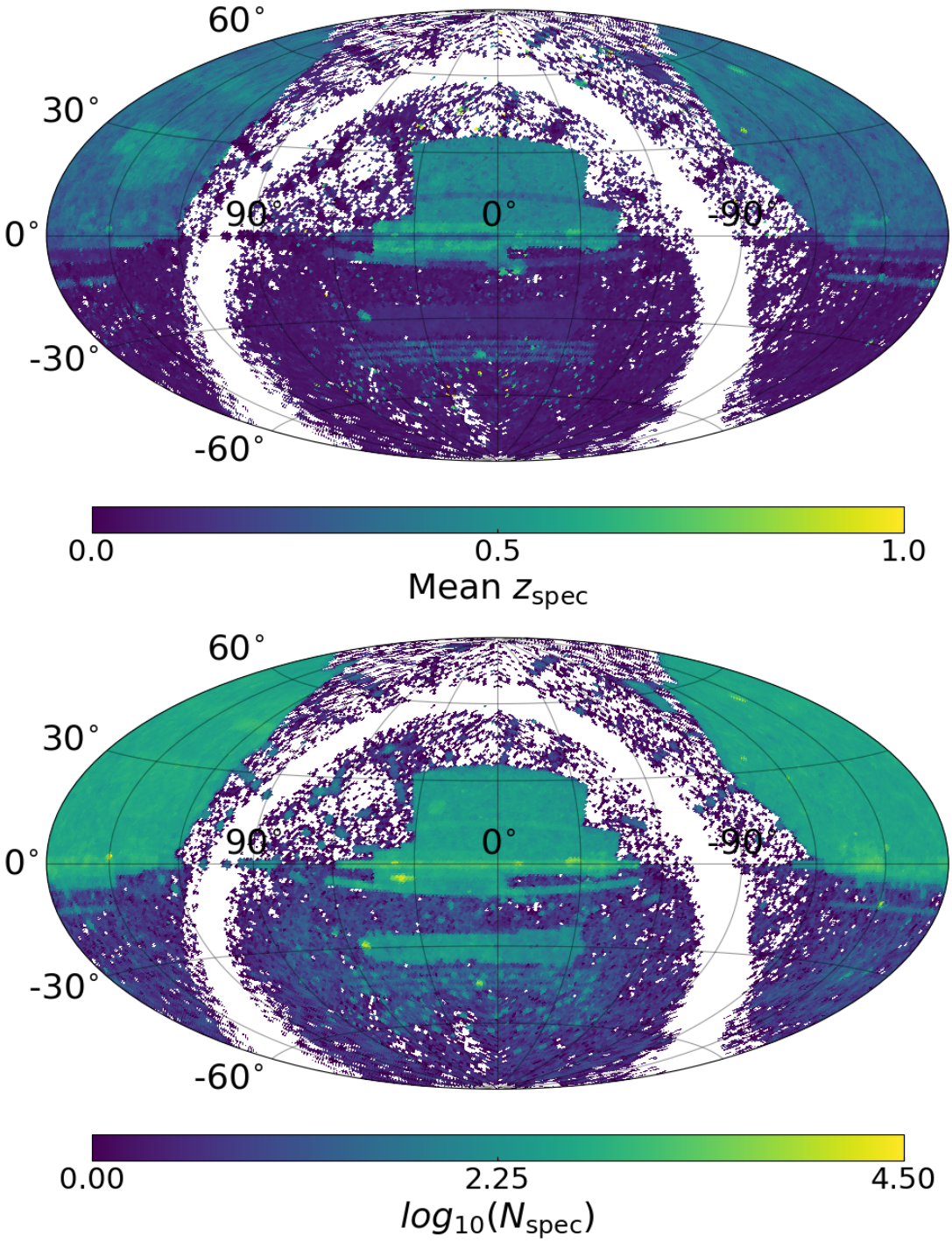

Abbildung 2.13: Adapted from Ider Chitham et al. in prep. The sky distribution of archival redshifts summarised in Table. 2.1. The upper and lower panels illustrate the mean spectroscopic redshift and total number of spectroscopic redshifts per healpixel (a resolution of `NSIDE`=4096 corresponds to a pixel area of 2×10^{-4}deg^2). The corresponding redshift distribution is shown in the right panel of Fig. 2.9 for all galaxies and various subsets; those which contribute to the spectroscopic training set as well as the sub-sample which reside in the *eROSITA*-DE hemisphere.

2.3.3 Results

Combining the archival redshifts outlined in Table. 2.1 (represented graphically in Figs. 2.9 & 2.13), with the automatic spectroscopic selection algorithm presented in Sect. 2.3.1 enables one to improve upon the photometric cluster selection. This is due to a reduction of the statistical redshift uncertainty by a factor of ~ 10 for clusters and ~ 200 for galaxies on average as demonstrated in Fig. 2.15. This is illustrated in terms of distance resolution along the line-of-sight in Fig. 2.16. This figure demonstrates that spectroscopic cluster redshifts are more closely aligned to the large scale structure traced by spectroscopic galaxy surveys relative to photometric redshifts, which are offset by 6.9 Mpc on average in this region. Fig. 2.8 also demonstrates the precision of spectroscopic redshifts relative to photometric redshifts. The redshift precision is graphically represented by the width of the redshift distribution of spectroscopic ($N_{z_{\rm spec}}$) and photometric members ($N_{z_{\rm red}}$) shown in relative to the width of the likelihood scan and the uncertainties associated to photometric cluster redshift. The example dataset presented here is based on X-ray selected clusters from the CODEX (Constraining Dark Energy with X-rays, Finoguenov et al., 2020) source catalogue which have been reanalysed using redMaPPer applied to The Legacy Surveys (for more detail, please refer to Sect. 3.1). The CODEX cluster with the highest number of spectroscopic members is shown in Figs. 2.11. Members that are first photometrically selected by redMaPPer and then subsequently selected by the automated spectroscopic membership algorithm are highlighted by coloured squares which illustrate how redshifted or blueshifted each member is relative to the spectroscopic bi-weight estimate of the cluster redshift. The corresponding phase space distribution is illustrated in Fig. 2.12.

Spectroscopically confirmed CODEX clusters in the SPIDERS survey are visually inspected and diagnostic plots are used to validate spectroscopically selected members on a per cluster basis (Clerc et al., 2016; Kirkpatrick et al., 2021). Analogous diagnostic plots are shown in Fig. 2.14 for spectroscopic members stacked relative to their host properties in bins of redshift for 2,039 SPIDERS clusters with $\lambda > 5$ and $|v| < 3000$km/s from the SPIDERS DR16 sample after a reanalysis with redMaPPer and The Legacy Surveys and re-running the automated spectroscopic member selection (Sect. 2.3.1). The contours in the lower-panel indicate that redMaPPer photometric redshifts are almost unbiased (the central ellipse of each contour set is centred about $z_\lambda = z_{\rm spec}$ independent of redshift bin). It also demonstrates that the probability of spectroscopic selection is highly correlated with the optical membership probability provided by redMaPPer (the density of sources significantly increases as $\ln(p) \to 0.0$, or $p \to 1.0$). The upper-panel of Fig. 2.14 shows the degree to which the radial selection is impacted by fibre collisions in spectroscopic follow up programmes. Although the radial filter implemented in redMaPPer favours galaxies that are closer to the optical cluster centre, projected distances which are approximately a quarter of the redMaPPer scale radius show a higher density source density relative to the central region of the stacked clusters. Although this effect is less prominent at lower redshifts, it is a result of the fact that spectroscopically observing targets in dense environments requires multiple passes to sufficiently sample the entire cluster.

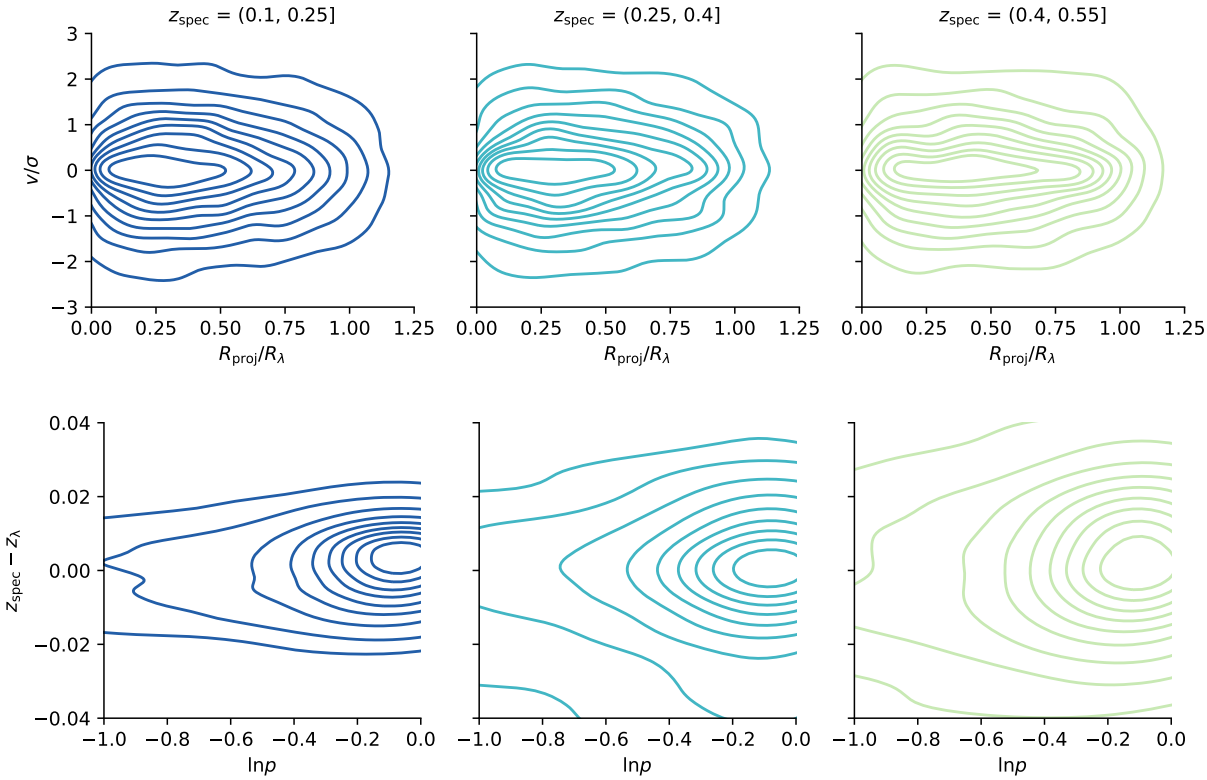

Abbildung 2.14: Adapted from Ider Chitham et al. in prep. Diagnostic plots of stacked member galaxies relative to their host properties as determined by the automated membership procedure (Sect. 2.3.1). 2,039 clusters with $\lambda > 5$ and $|v| < 3000$km/s from the SPIDERS DR16 cluster sample reanalysed with imaging from The Legacy Surveys (Ider Chitham et al., 2020, described fully in Sect. 3.4 of this thesis). **Upper panel:** velocity-distance plot. The velocity offset of each spectroscopic member galaxy relative to the velocity dispersion of its host cluster as a function of projected distance from the optical cluster centre in units of the redMaPPer scale-radius. **Lower panel:** redshift-probability plot. offset of all red-sequence members with spectroscopic redshifts relative to the redMaPPer photometric redshift (z_λ) as a function of ln(p) i.e. the natural logarithm of the redMaPPer membership probability. Both plots are based on those created on a per-cluster basis as part of the SPIDERS visual inspection procedure (Clerc et al., 2020; Kirkpatrick et al., 2021).

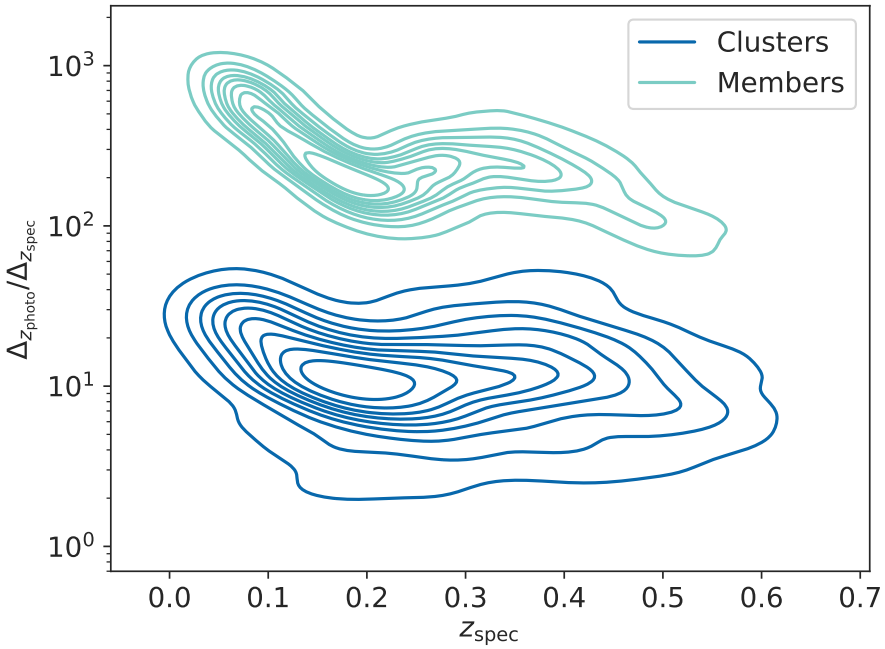

Abbildung 2.15: Adapted from Ider Chitham et al. (2020). The ratio of photometric and spectroscopic uncertainties for clusters and member galaxies as a function of spectroscopic redshift for the SPIDERS DR16 sample reanalysed with photometry from The Legacy Surveys (Sect. 3.4). Here, $z_{\rm photo}$ represents z_λ (Sect. 2.2.2) for clusters and $z_{\rm red}$ (Sect. 2.2.2) for members. $z_{\rm spec}$ represents $z_{\rm biwt}$ for clusters and $z_{\rm spec}$ for galaxies. The distribution of clusters and individual member galaxies are represented by independently normalised coloured contours. This indicates that on average the improvement provided by spectroscopic redshifts is a roughly a factor of ~ 10 for clusters and a ~ 200 for members.

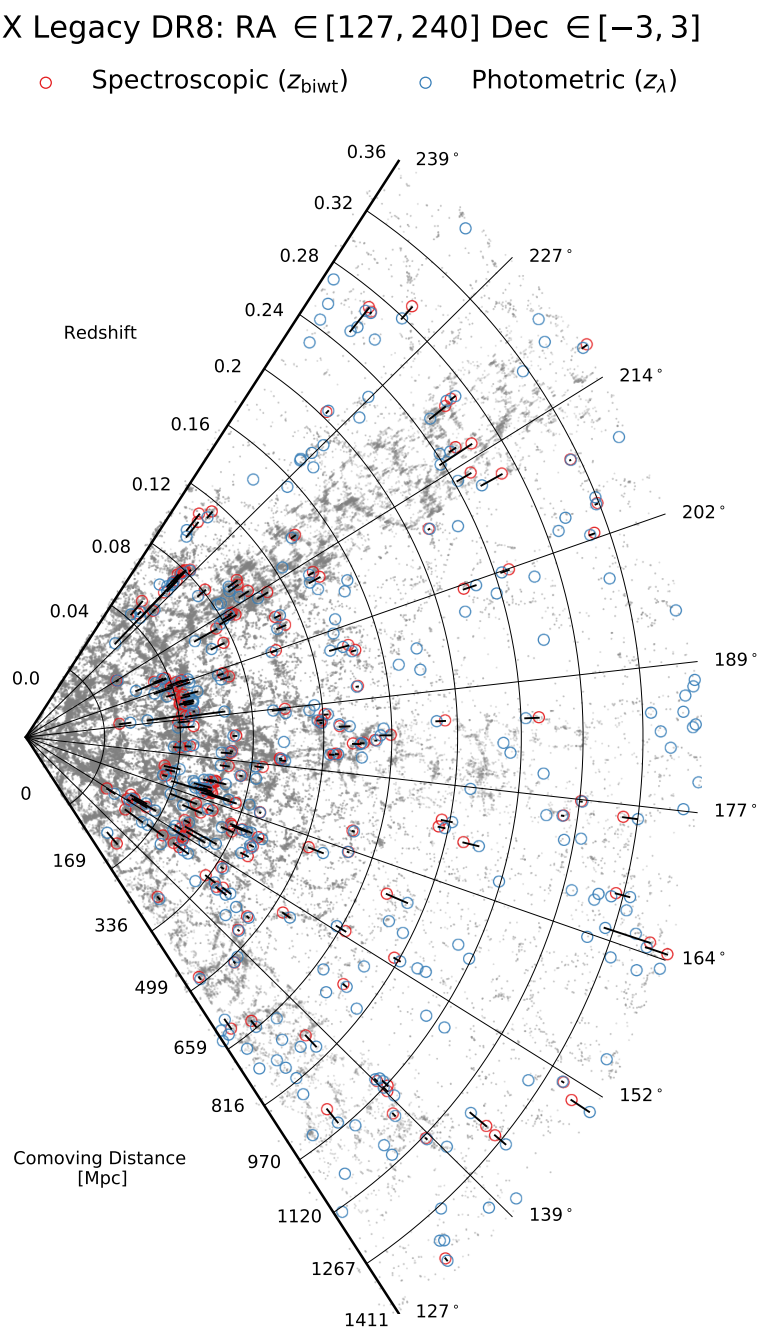

Abbildung 2.16: Adapted from Ider Chitham et al. in prep. Grey points show the position of spectroscopic galaxies over the GAMA (G09, G12, G15; Liske et al., 2015) regions using all publicly available archival spectroscopic redshifts (Sect. 2.3.2). Blue circles depict the position of CODEX cluster candidates with photometric redshifts derived from redMaPPer applied to DR8 of The Legacy Surveys (Sect. 3.4) with $\lambda > 20$. Red circles show a subset of clusters with spectroscopic redshifts (at least three spectroscopic member galaxies) as determined by the automated spectroscopic membership algorithm described in Sect. 2.3.1. Spectroscopic clusters (red) are more closely aligned with the large scale structure traced by the position of the underlying galaxies in comparison to photometric (blue) clusters due to the relative accuracy of the redshifts, which differ by 6.9Mpc on average in this region.

.

Kapitel 3

Method: creating spectroscopically confirmed cluster catalogues

As the uncertainty on cosmological parameters scales with the surveyed volume (therefore area and maximum observable redshift), the spectroscopic follow-up of *eROSITA* clusters with SDSS-V and 4MOST aims to utilise the largest possible area of sky in order to maximise the scientific output. For 4MOST and SDSS-V the goal is to follow up eRASS clusters spectroscopically over 13,500 deg^2 of extra-galactic sky. Spectroscopy will improve the cosmological constraining power of the *eROSITA*-DE cluster sample by ensuring robust cluster identifications and improving the accuracy in the luminosity and distance determination. It will also enable precise measurements of the 3-dimensional cluster power-spectrum and provide an estimate for cluster masses based on the bias-mass relation.

To facilitate the spectroscopic follow-up of eRASS clusters, Ider Chitham et al. in prep. uses all available public photometric and spectroscopic datasets to produce a cluster catalogue that covers almost the entire sky. This catalogue is redMaPPer-based and includes spectroscopic cluster redshifts measured using the techniques outlined in Kap. 2. The cluster redshift range, $z \in [0.05, 1.20)$, depends almost entirely[1] on the inhomogeneous photometric coverage from DESI Legacy Imaging DR9 (Dey et al., 2019), unWISE (Mainzer et al., 2014; Meisner et al., 2018), NSC DR2 (Nidever et al., 2020), SkyMapper DR2 (Onken et al., 2019), Pan-STARRS DR2 (Flewelling, 2018) and VST KIDS DR4 (Kuijken et al., 2019), DES Y1 (Drlica-Wagner et al., 2018) and HSC PDR2 (Aihara et al., 2019).

This thesis, however, focuses on creating a state of the art catalogue of spectroscopically confirmed X-ray selected clusters, *prior* to launch of *eROSITA*, within the SDSS-IV/SPIDERS cluster programme. This chapter is based on the optical follow-up of sources from the ROSAT All-Sky Survey (i.e. the predecessor of eRASS) and includes details of the original SPIDERS catalogue (Sect. 3.1, described in Clerc et al., 2016, based on SDSS-III/BOSS DR8). It also describes updates made by the author to improve the number of spectroscopic members per cluster (Sect. 3.2, contributions to Clerc et al., 2020, based

[1]The on-sky distribution of archival spectroscopic galaxies (Fig. 2.13) used for calibration purposes as well as the luminosity threshold value $L_{\rm thresh}$ adopted in Eq. 2.4 also have an impact on the redshift range that a redMaPPer cluster catalogue can span. Please refer to Sects. 2.2.2 & 3.5.1 and for further details.

on Pan-STARRS1 DR1), the generation of a synthetic cluster catalogue used to validate the visual inspection process (Sect. 3.3.1, contributions to Kirkpatrick et al., 2021, based on the Magneticum simulations) and the creation of a volume-limited version of the SPIDERS catalogue in preparation for cosmological analysis (Sect. 3.4.2, contributions to Ider Chitham et al., 2020; Kirkpatrick et al., 2021, based on DESI Legacy Imaging DR8). Additionally, a brief overview is provided for the next generation of spectroscopic cluster catalogues based on the eRASS. This includes the effective optical survey area for selection of the surveys used in Ider Chitham et al. in prep and the development of a 4MOST sky fibre allocation algorithm based on systematic maps created for each optical survey. The depth map creation (Sect. 3.4.2) and volume-limiting procedures (Sect. 3.4.2) are inspired by the pioneering work of Rykoff et al. (2015, 2016).

3.1 State of the art catalogues of X-ray clusters from SDSS-IV/SPIDERS DR16

The SPectroscopic IDentification of *eROSITA* Sources observational program (SPIDERS) within SDSS-IV/eBOSS (Dawson et al., 2016; Blanton et al., 2017) aims to follow up X-ray detected clusters (Clerc et al., 2016) and AGN (Dwelly et al., 2017) using the BOSS spectrographs (Smee et al., 2013) on the 2.5m SDSS telescope (Gunn et al., 2006). The SPIDERS cluster catalogue published as part of the sixteenth data release of the SDSS (DR16; Ahumada et al., 2020) consists of a set of spectroscopically validated galaxy clusters drawn from two X-ray galaxy cluster surveys:

- CODEX, the COnstraining Dark Energy with X-rays sample (Finoguenov et al., 2020), constructed by reanalysing the publicly available data from the *ROSAT* All-Sky Survey (RASS; Voges et al. 1999). It covers 10,800deg^2.
- XCLASS, the *XMM* CLuster Archive Super Survey catalogue constructed by reanalysing the publicly available *XMM* data (Clerc et al., 2012; Sadibekova et al., 2014). These serendipitous pointings span a total area of 3-4deg^2 in common with the SDSS imaging footprint.

The final DR16 SPIDERS cluster catalogues are thoroughly described in their respective companion papers Kirkpatrick et al. (2021) and Clerc et al. (2020).

Both catalogues are constructed by applying the red-sequence Matched-filter Probabilistic Percolation cluster-finding algorithm (redMaPPer; Rykoff et al., 2014) to sdss-III/BOSS DR8 imaging data (Fukugita et al., 1996; Aihara et al., 2011). This approach provides optical counterparts (over-densities of red galaxies) to the extended X-ray emission observed with *ROSAT* or *XMM*. Each optical counterpart is comprised of a set of potential member galaxies. The most probable members are observed spectroscopically (Clerc et al., 2016) and the sum of membership probabilities provides an estimate for the optical richness $\tilde{\lambda}_{\rm SDSS}$ (estimated using the original SDSS derived photometric redshift, $\tilde{z}_{\lambda_{\rm SDSS}}$; Rykoff et al., 2014). In this thesis, only the CODEX component of the SPIDERS

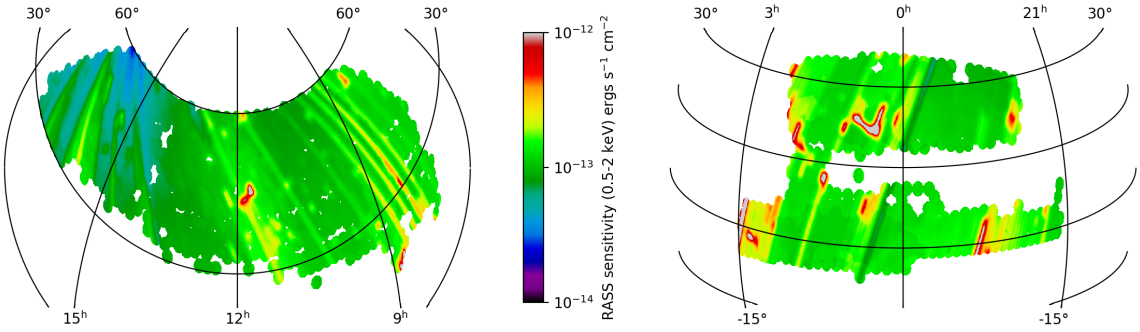

Abbildung 3.1: CODEX X-ray sensitivity from the RASS over the SPIDERS DR16 footprint (5,350 deg^2) in the 0.5-2.0 keV band (Clerc et al., 2020; Finoguenov et al., 2020). Coordinates are projected in the equatorial system. The left and right subplots illustrate the north and south Galactic caps respectively. The resolution is $N_{\rm side} = 512$ in the `HEALPIX` scheme, which equates to a pixel size of 7 arcmin. eBOSS spectroscopic plates used in SPIDERS DR16 sample are easily identifiable by their circular footprints, each with a 1.5 deg radius. The relevant optical systematics over the same region are illustrated in Fig. 3.2 for each photometric survey discussed in this chapter.

cluster catalogue is considered in order to simplify the modelling of the overall selection function during cosmological analyses (Sects. 4.1 and 4.2).

The SDSS-IV/SPIDERS spectroscopic data span 5,350deg^2 out of the total 10,800deg^2 sdss-III/BOSS footprint which is covered by CODEX. The 0.5-2.0 keV X-ray sensitivity of the RASS and the i-band optical galaxy magnitude limit of the SDSS over this region are illustrated in Fig. 3.1 and the top panel of Fig. 3.2 respectively. Within this area, the total number of spectroscopically validated and visually inspected clusters of galaxies with $\tilde{\lambda}_{\rm SDSS} > 10$ is 2,740 i.e. about 0.5 per square degree (Clerc et al., 2020; Kirkpatrick et al., 2021). For systems with $\tilde{\lambda}_{\rm SDSS} > 40$, a total of 920 out of 1047 are confirmed. The remaining 127 are either dubious candidates or high redshift systems ($\tilde{z} > 0.7$) lacking spectra. This emphasises the importance of spectroscopy for cluster confirmation.

3.2 Improving spectroscopic completeness at high-redshift

Over the course of the SPIDERS programme, several improvements to the target selection were made relative to the original redMaPPer implementation based on the SDSS (Clerc et al., 2020). These revised selection methods were specifically designed improve the spectroscopic confirmation rate of higher-redshift systems ($z > 0.5$) in chunks `eboss26` and `eboss27`. These additional targets take advantage of deeper optical datasets overlapping the CODEX sample such a CFHT (described in Kiiveri et al., 2021) and Pan-STARRS1. This section describes the data processing and analysis relevant to the spectroscopic follow-up of CODEX clusters based on Pan-STARRS1 optical photometry. This resulted in the

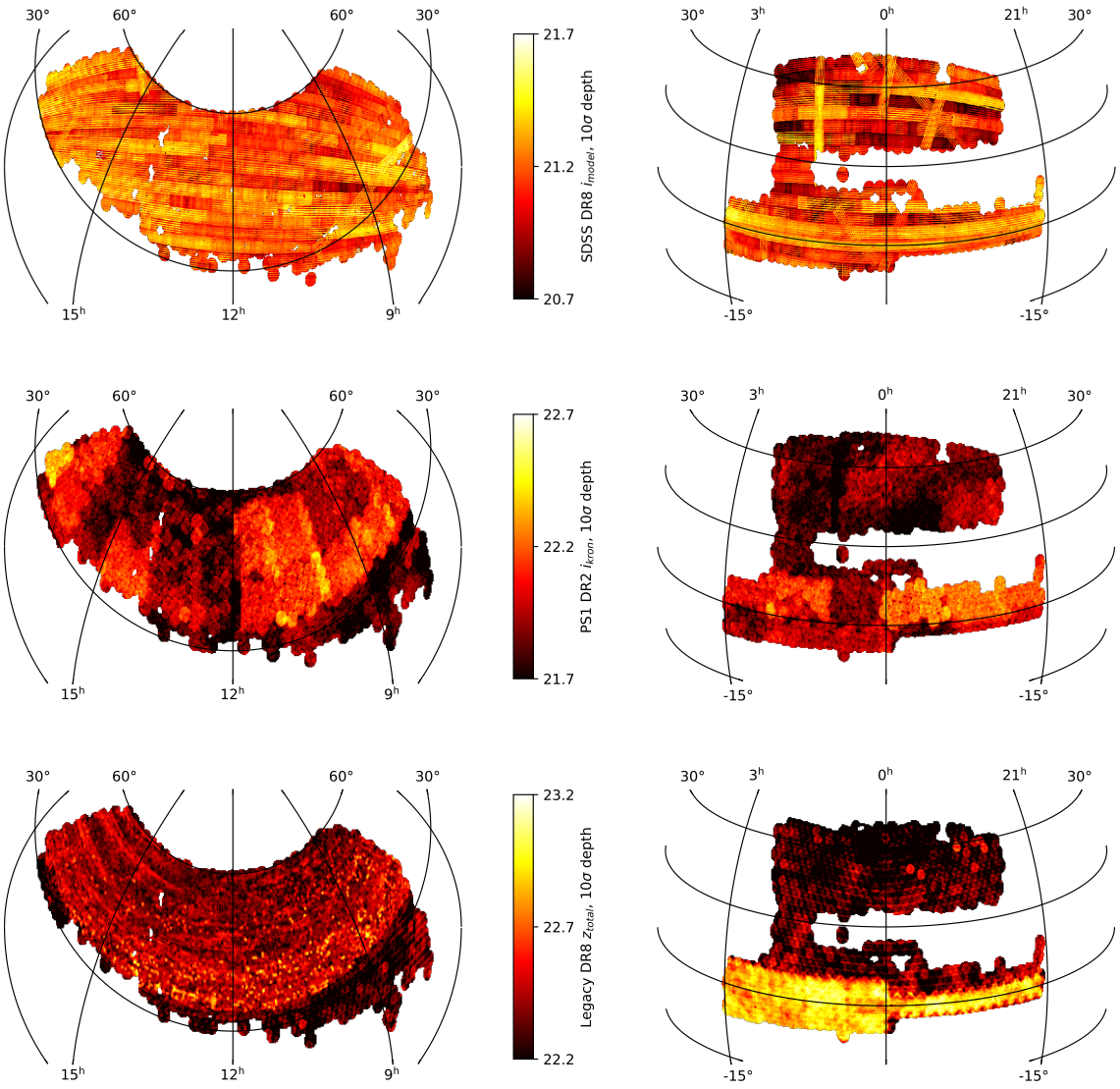

Abbildung 3.2: **Upper panel:** The limiting galaxy i band magnitude of the SDSS DR8 ($N_{\rm side} = 2048$). Although not directly implemented in the construction of the original redMaPPer (Rykoff et al., 2014) and CODEX (Finoguenov et al., 2020) catalogues, this map illustrates the optical systematics that affect the confirmation of X-ray selected clusters in the original SPIDERS DR16 catalogue (Clerc et al., 2020; Kirkpatrick et al., 2021) over the 5,350 deg^2 footprint. This panel is adapted from Rykoff et al. (2015) and Clerc et al. (2020). **Middle panel:** The same depth map ($N_{\rm side} = 1024$) for Pan-STARRS DR1 (Sect. 3.2, please note that the range of colour bar is shifted by one magnitude with respect to the upper panel). **Lower panel:** z-band depth map for The Legacy Surveys over the same region ($N_{\rm side} = 2048$). This shows the optical systematic relevant to the construction of the SPIDERS DR16 cosmological cluster sample (Sect. 3.4). The latter the only map of the three which is utilised by the relevant cluster confirmation algorithm when considering the optical properties of the X-ray selected clusters (Ider Chitham et al., 2020). It is also the only region that differs in the footprint size, with the northern border of the south galactic cap illustrating a slight loss of area relative to the original DR16 definition.

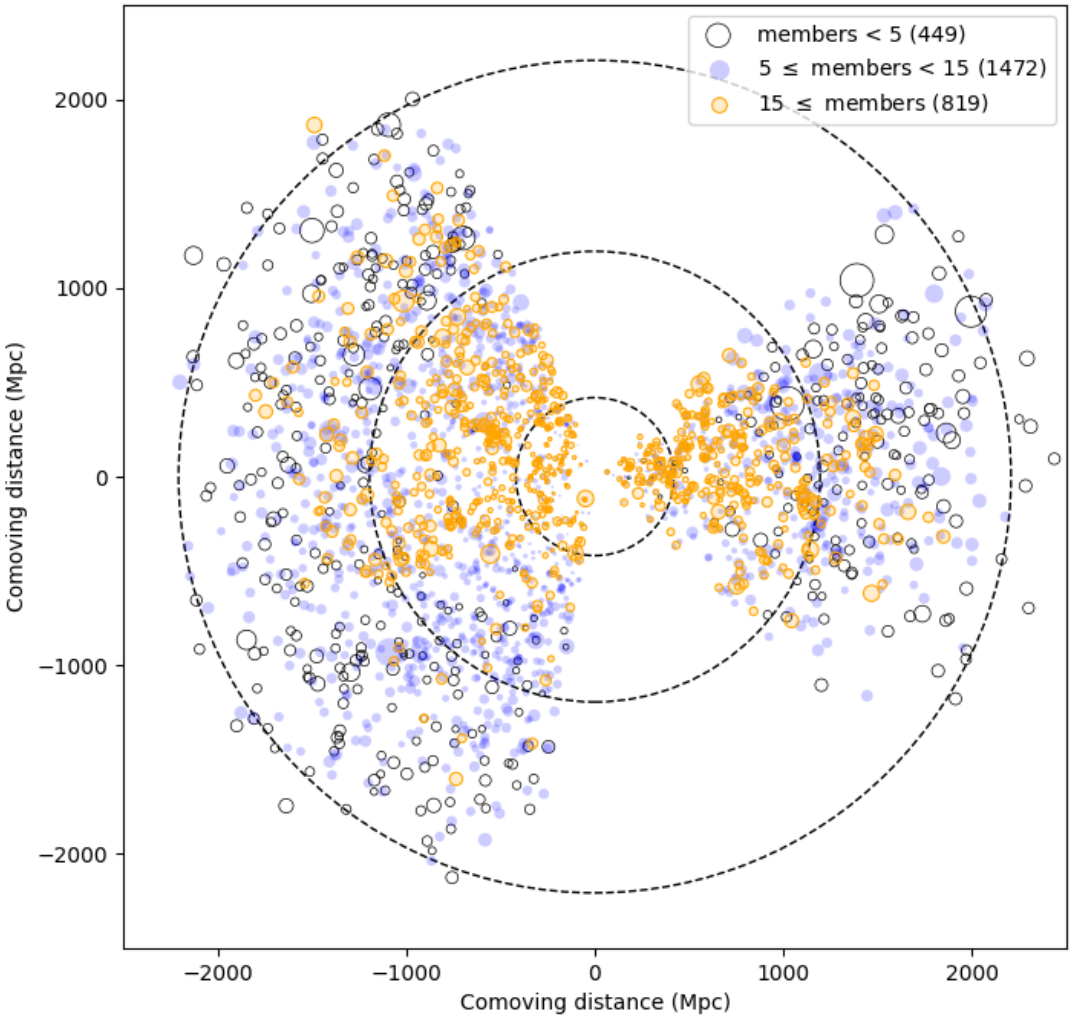

Abbildung 3.3: Adapted from Clerc et al. (2020). Representation of the *original* (SDSS based) SPIDERS DR16 catalogue of 2,740 spectroscopically validated CODEX clusters (see Kirkpatrick et al., 2021, for a full description). Circles represent a spectroscopically confirmed and visually inspected system. The colour represents the number of spectroscopic members and the size scales with the X-ray luminosity (and therefore mass). This sample is *not* volume-limited. The effect of Malmquist bias is shown by the increase in the average cluster luminosity with increasing redshift. Comoving distances in positive x direction tend towards right ascension of 0.0 deg and comoving distances in the positive y direction tend toward a right ascension of 90.0 deg.

submission of 8,897 spectroscopic targets for $\sim$ 250 SPIDERS clusters at high-redshift $z \in [0.50, 0.67)$. 150 cluster candidates are rich systems in the Northern galactic cap, while the other $\sim$ 100 are poorer systems (on average) and spread across the full BOSS footprint.

The Panoramic Survey Telescope and Rapid Response System (Pan-STARRS) (Kaiser, 2004) is a wide-field imaging facility on the summit of Haleakala, Hawaii with a single 1.8m telescope with a 7 deg^2 field of view, a gigapixel camera with on-chip guiding capabilities, and a g, r, i, z, y filter set (Tonry et al., 2012; Magnier et al., 2016). It is split into several separate surveys. The most relevant to extragalactic astronomy are the (relatively) shallow 3π survey covering the entire sky north of declination -30 deg and the Medium Deep Survey (MDS), which covers 10 smaller area fields to much greater depths. Cluster finding with PS1 has been previously explored several times. Notably, Murphy et al. (2012) developed an algorithm designed for use with a mock dataset of the PS1 MDS fields and the SDSS to find clusters by applying a photometric galaxy filter in colour-magnitude space. The algorithm works by identifying regions of high surface density with Voronoi diagram and then grouping them into clusters with a Friends-of-Friends technique. Ebeling et al. (2013) also optically confirmed clusters using observational Pan-STARRS datasets in a comprehensive search for distant, X-ray luminous galaxy clusters at $z > 0.5$. This search used the RASS and Pan-STARRS MDS as part of the pilot study for the extended Massive Cluster Survey (eMACS). Two cluster candidates were identified at $z \sim 0.6$. As the MDS has a limiting 5σ i-band magnitude of 26.3, it is significantly deeper than the 5σ 22.5 limit of the 3π survey, therefore instead of testing the efficiency or reliability of optical cluster confirmations obtained by eMACS based on PS1-3π images, their pilot study assessed the reliability of the faintest (low X-ray photon count) RASS sources as credible detections of massive high-redshift clusters.

The algorithm chosen for the high-redshift iteration on the SPIDERS target provision with PS1 is based on a custom version of the Multi-Component Matched Filter (MCMF) Cluster Confirmation tool (Sect. 2.2.3, Klein et al., 2018) due to the close similarity to scanning-mode functionality of redMaPPer (Sect. 2.2.2. Rykoff et al., 2014) as well as the luminosity, radial and colour filters. This algorithm is capable of ingesting the equatorial coordinates for each galaxy cluster candidate and determining the most probable photometric redshift by finding peaks in the richness-redshift distribution over the interval (0, 1) with steps of $\delta z = 0.001$. The following subsections describe the creation of custom photometric catalogues and the application of the cluster finding algorithm itself.

3.2.1 Source Extraction from Pan-STARRS1 DR1 images

In order to optimise the cluster confirmation process, photometric catalogues are necessarily extracted from the original DR1 images[2]. Compressed 3π-PS1 co-added science, weight and flag FITS image files are downloaded from the public PS1 cutout server endpoint[3] in g, r, i and z bands. Each science and weight image were then uncompressed with a non-standard

[2]Prior to the release of PS1-DR2, there were known problems in the PS1-DR1 source object catalogues, which meant they were not suitable for optical cluster confirmation. These issues have since been rectified.

[3]http://ps1images.stsci.edu/cgi-bin/ps1cutouts

image flux scaling determined by the BSOFTEN and BOFFSET keywords in each FITS header[4]. The flux is non-linearly scaled using an arcsinh transformation that converts them to a pseudo-magnitude scale image data related to the asinh magnitudes (aka luptitudes) that are used in the SDSS (York et al., 2000). FITS images therefore uncompressed using the transformation given by Eq. 3.1 (Waters et al., 2016).

$$v = \texttt{BZERO} + \texttt{BSCALE} * j \tag{3.1}$$
$$x = v * \ln(10)/2.5 \tag{3.2}$$
$$\phi = \texttt{BOFFSET} + \texttt{BSOFTEN} * 2 * \sinh(x) \tag{3.3}$$

Here BZERO and BSCALE are the standard FITS keywords for converting an integer image to a floating-point value. If j is the original integer pixel value, these equations convert to the float pixel value v and then to a standard linear flux, ϕ.

Photometric catalogues of galaxies are then created from the uncompressed science co-adds using Source Extractor (SExtractor, Bertin & Arnouts, 1996) 2.25.0 and Point Spread Function Extractor (Bertin, 2011, PSFex) 3.18.2. SExtractor automatically measures the background and its RMS noise, subtracts the background, filters (convolves with specified profile), finds objects (thresholds), de-blends detections, measures shapes and positions, cleans (reconsider detections, accounts for neighbours), performs photometry and provides information that can be used to separate star-like or galactic objects, then outputs a catalogue. The relevant magnitudes types are

- MAG_AUTO. Measured by summing the flux in an ellipse scaled to the Kron radius (Kron, 1980).
- MAG_MODEL. Measured by fitting the object with a given model and estimating the flux for this model.
- MAG_PSF. Determined during the PSF-fitting process.
- MAG_DETMODEL. Similar to MAG_MODEL but first carries out the model fitting on the detection image, and then fits the overall normalisation of this model to each single-band image separately. MAG_DETMODEL thus has a consistent galaxy model for the same galaxy across all filters, which is essential to provide reduce the scatter in the red-sequence colour measurements.

The custom pipeline configuration for the PS1 source extraction process is split into several steps:

- Extract i-band sources with SExtractor using the default PSF model.
- Create i-band PSF model with PSFex.

[4]Flag images from the server are already uncompressed due to the relatively small file size.

- Re-extract sources with `SExtractor` and the newly created PSF model in *dual* image mode. This allows one to get generate the `DETMODEL` magnitude.

- Repeat this process with g, r and z-bands, using i-band files as a reference for the dual image (forced photometry) mode as this provides the best combination of 3π stack 5σ depth (23.1) and median seeing (1.11 arcsec)[5] This finds the objects in one image and then applies the apertures and positions found on another image.

For each of the considered skycell regions, each source extracted single-band catalogue is merged into a single, multi-band catalogue. Publicly available data from the PS1 Catalogue Archive Server `ps1casjobs`[6] are utilised in addition to the co-added images to calibrate the zero-points and correct for galactic extinction after over the same sky-region for each cluster candidate within a search radius of 0.4 deg. This zero-point and dust correction[7] calibration step is carried out on a per skycell basis by matching all sources and then averaging over the difference in magnitude for each band, colour and magnitude type to determine the correction term.

In order to minimise the effect of artefacts from the source extraction process which tend to be prominent near the border of the multi-band catalogues a secondary merge is performed to combine each of the nine calibrated, multi-band skycell catalogues into a single catalogue that covers the vicinity of each cluster and surrounding area (required for background subtraction). While recursively concatenating each multi-band skycell catalogue, an additional data-cleansing procedure is carried out in the overlap region between neighbouring skycells[8] (a minimum of 240 pixels in each direction). The sources of each catalogue are sky-matched in this region and trimmed in such a way that clips sources in the area union that are closer to the border relative to the centre of their originating skycell from with respect but as close to the centre of the respective skycell as possible while still retaining unique sources that only appear in one of the catalogues. After safely trimming artefacts, remaining duplicates in the trimmed overlap region are eliminated within 0.250 arcsec. This is an appropriate distance since it is very close to the native image scale of the detector (0.258 arcsec). During this combination step, each colour and band is also converted to DECam magnitudes/colours such that it is possible to directly utilise the red-sequence models developed by Hennig (2017) and calibrated by Klein et al. (2018).

Creating photometric catalogues for PS1 in this way produces less "ghosts" and CCD related issues relative to the public catalogues, although artefacts around bright stars are more frequent. This can lead to apparent over-densities of spurious source detections which would otherwise bias the measurement of richness, photometric redshift and quality of spectroscopic targets. To safely remove these false positives in the source detection procedure, an additional mask is created for each cluster candidate by utilising the GAIA DR1(Gaia Collaboration et al., 2016) G-band stellar catalogue and an estimate for the saturation

[5] See http://panstarrs.stsci.edu/ for comprehensive band comparison.
[6] http://mastweb.stsci.edu/ps1casjobs
[7] This assumes that dust extinction does not vary at the resolution of a single skycell.
[8] `https://outerspace.stsci.edu/display/PANSTARRS/PS1+Sky+tessellation+patterns`.

radius (Eq. 3.4) that removes any sources with compromised photometry (Coupon et al., 2018).

$$r\,[\mathrm{arcsec}] = \begin{cases} 708.9 \times \exp(-G_{\mathrm{Gaia}}/8.41), & G_{\mathrm{Gaia}} < 9 \\ 694.7 \times \exp(-G_{\mathrm{Gaia}}/4.04), & G_{\mathrm{Gaia}} \geq 9 \end{cases} \tag{3.4}$$

3.2.2 Target selection of cluster members

Given these source catalogues, the `SPREAD_MODEL`, a measure produced by `SExtractor` of how star-like or galaxy-like is used to filter stars from galaxies. The cut of `SPREAD_MODEL` $<$ 0.0025 provides a clean separation between stars and galaxies down to a magnitude limit of $m_i = 22.0$ (Klein et al., 2018). Convolving the colour weights with a Gaussian smoothing kernel then provides a local measurement of the red-sequence galaxy density in the region of each source. Members are then selected by applying a threshold to product of the weight and density within a radius of 1.5Mpc from the optical centre down to $m_z = 21.0$. This ensures targets can be efficiently observed, using the BOSS spectrograph, down to the desired signal-to-noise ratio. These are interpolated with a cubic spline using the i-band magnitudes of the extracted sources to estimate the intrinsic colour scatter. The width of the (pseudo) Gaussian PDF is then taken as the root of the sum of squares of the intrinsic colour scatter and the detection model error (as indicated by `SExtractor`) for each source. The weight value for each source is then convolved using 2D Gaussian smoothing with a kernel width of 0.2Mpc. Density values are then normalised by the median[9] density across the area of merged sky (between 1-4 skycells) on a per cluster candidate basis. The red-sequence galaxy density map for the CODEX cluster `1_4117` is illustrated in Fig. 3.4 at a photometric redshift of $z = 0.604$. The target priority for observations in SDSS-IV is taken as the product of the median colour weight of each source and the median red-sequence galaxy density re-scaled over the range [0, 100) on a per cluster basis.

3.3 Creating a simulated catalogue of spectroscopically observed clusters

To test pipelines designed for the data processing and cosmological of eRASS clusters using simulations, one must create a synthetic dataset as indistinguishable to real X-ray and optical observations as possible. This must take into consideration variations in the effective *eROSITA* X-ray exposure time and multi-band optical photometry uncertainties across inhomogeneous wide-area optical surveys. Although such state of the art simulations (Comparat et al., 2020a) are processed using an inject-retrieve method in Ider Chitham et al. in prep., at the time of this thesis submission these results were preliminary. The following subsection therefore describes an earlier iteration of an analogous exercise used to validate the spectroscopic redshift assignment and visual inspection pipelines used in

[9]The median of the $g-r$, $r-i$ and $i-z$ colour weights and density maps produces the most robust results in the case of missing or bad PS1 photometric data.

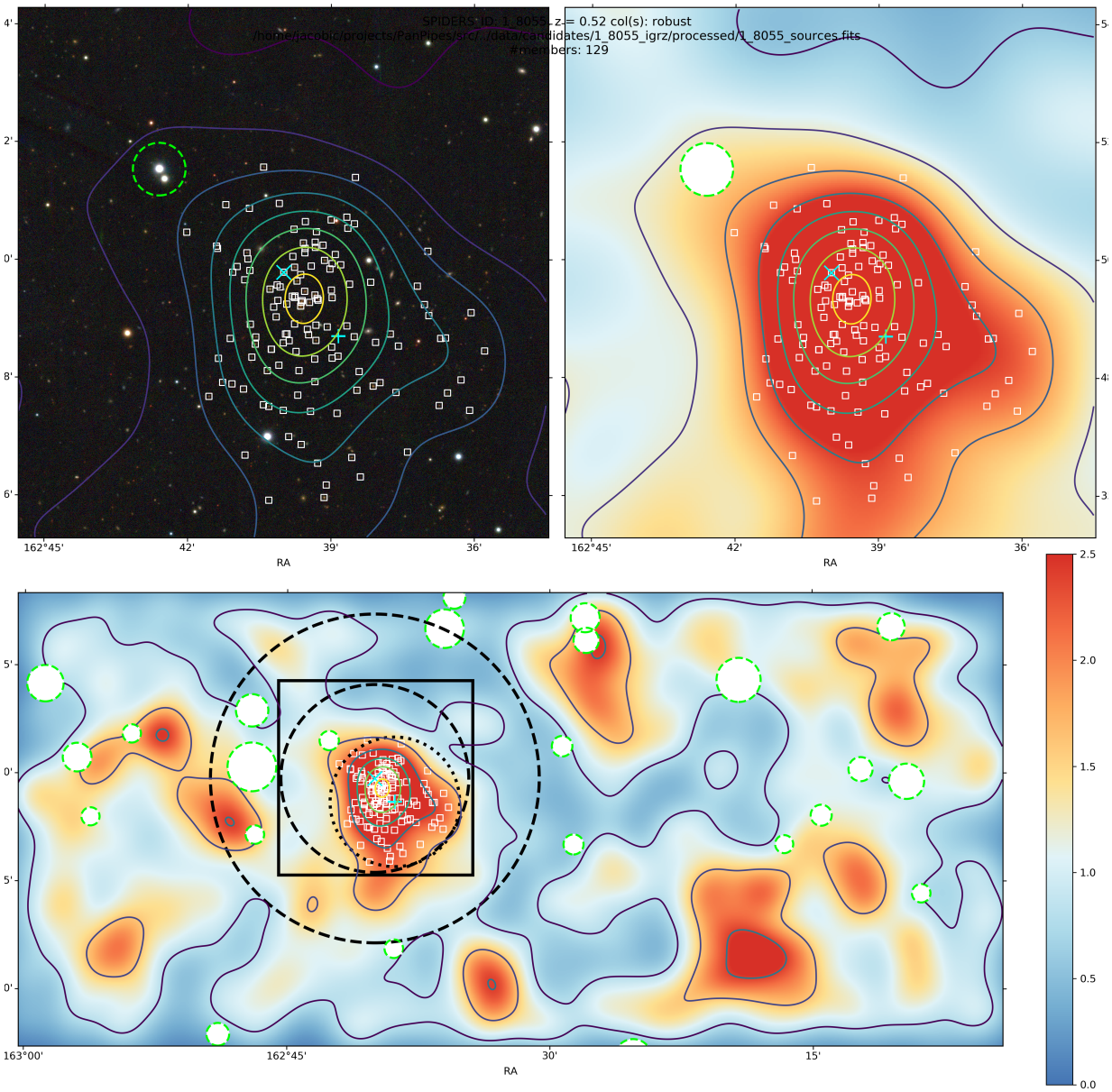

Abbildung 3.4: Adapted from Ider Chitham et al. in prep. **Lower panel:** MCMF density map for CODEX cluster `1_8055` at the photometric redshift of 0.52. Contour levels illustrate the local projected density of red-sequence galaxies at the photometric redshift. The red circle is centred on the X-ray centre of the cluster with a radius of 2.0Mpc. Red squares indicate the location of a subset of galaxies which have been previously observed with reliable spectroscopic redshifts. White circles are stars masked according to Eq. 3.4. **Upper panel:** a zoomed in region of the lower-panel. The left sub-figure shows the density map, while the right shows the optical *i*-band PS1 image used during the source extraction process Sect. 3.2.1.

the creation of the SPIDERS DR16 cluster sample (Clerc et al., 2020; Kirkpatrick et al., 2021).

3.3.1 Validating spectroscopic redshift assignment and visual inspection

In the context of spectroscopic target provision, to quantitatively understand the performance of red-sequence based cluster confirmation tools such as redMaPPer and MCMF as well as the subsequent selection effects such as fibre collisions and the final visual inspection, it is necessary apply an analogous pipeline to a simulated dataset that resembles the observational data. Three professional astronomers from the SPIDERS team took part in the experiment to visually inspect a synthetic catalogue of cluster targets. Each system was independently inspected by 2.4 astronomers on average and 97% of clusters have been successfully validated of which approximately 10% showed significant project effects with multiple spectroscopic structures along the line of sight.

The artificial SPIDERS dataset[10] is based on the Magneticum hydrodynamical simulations[11] (Hirschmann et al., 2014; Dolag et al., 2015) under the assumption of WMAP7 cosmological parameters (Komatsu et al., 2011), utilising all but the highest redshift slice. The resulting catalogue of simulated galaxy clusters contains 683 systems up to $z \sim 1.2$. This provides SDSS-like observer frame photometry based on Bruzual (2007) stellar population models with a underlying Chabrier (2003) initial mass function in the VEGA system, where attenuation has been accounted for.

The colours and uncertainties of galaxies in the Magneticum simulation do not show sufficient resemblance to observations for optical cluster confirmation algorithms to be directly applied to the synthetic galaxy catalogues without applying more sophisticated techniques to blend them into observational datasets (e.g. Ider Chitham in prep.). A more appropriate method to validate the spectroscopic redshift assignment and visual inspection process within SPIDERS DR16 is therefore to apply an artificial selection which resembles the filters used in redMaPPer and MCMF (described in Sects. 2.2.2 and 2.2.3). These selection processes are included in the list below, which contains all of the selection processes considered in the generation of the synthetic catalogue of spectroscopically confirmed systems in the order that they are applied:

- **Photometric depth cut**. In order to ensure that the limiting photometric depth of the simulated data is consistent with, the SDSS (the primary optical survey used to create the SPIDERS cluster catalogue), it is necessary to apply cuts to reflect the apparent magnitude limits of the SDSS i.e. g: 22.2, r: 22.2, i: 21.3, z: 20.5 with a completeness of 95%[12].

[10] https://github.com/jacobic/Magneticum.git
[11] http://www.Magneticum.org/data.html
[12] https://www.sdss.org/dr12/scope/\protect\let\futurelet\@let@token\let\let\relax

- **A galaxy luminosity (absolute magnitude) cut** of $M_* + 1$. This minimises the scatter in optical richness for SDSS-like data (Rykoff et al., 2012) and is applied in the i-band, due to its usage as the reference band for cluster detection in both SDSS and PS1.

- **A radial filter and projected contamination**. For each galaxy cluster with virial radii in the range of 1.8–2.2Mpc, every galaxy within a sphere of $5R_{vir}$ is selected as preliminary member galaxies. Every galaxy from all redshift snapshots (up to $z \sim 1.2$) that falls along the line of sight within a projected 2D area with a radius of 0.4 deg that is not within the corresponding 3D selection is selected as contamination i.e. field galaxies in the foreground and background or other projected structures.

- **A red-sequence galaxy selection**. The fraction of red-sequence galaxies changes as a function of the simulated photometric cluster redshift and 3D distance from the centre of cluster mass. This relation (Eq. 3.5) can be derived from a bi-modality in the population of galaxies in the stellar mass - star formation rate plane. This fraction (Eq. 3.5) is derived analytically from the Magneticum simulation data, although has since been super-seeded by a slightly modified form (Comparat & Ider Chitham in prep.) which is calibrated with respect to observations (Hennig et al., 2017; Nishizawa et al., 2018; Murata et al., 2020).

$$f = ((r + 0.01)^{-\frac{1}{4}} - 0.01r - 0.47)(1 + z)^{\frac{5}{8}} \tag{3.5}$$

  Here, f is the fraction of red-sequence galaxies and r is the distance from the cluster centre in units of the cluster virial radius.

  This relation is applied by iterating outwards from the cluster centre in concentric annuli each with a width of 0.5 arcmin, out to a maximum radius of 14 arcmin. This upper limit was chosen to match the stacked sky distribution of SPIDERS cluster targets derived from CODEX in chunk `eboss1` (656 systems) as shown in Fig. 12 of Clerc et al. (2016). Given the fraction of galaxies to sample in each annuls, priority is then given to those with the minimum absolute age difference relative to the median age of the true cluster members.

- **A 2 arcsec fibre i-band magnitude cut**. The same cut applied for all SPIDERS targets in order to maximise the redshift determination efficiency of the targets: $17.0 <$ `FIBER2MAG_I` < 21.2 (Clerc et al., 2016). As stellar mass is provided for each galaxy in the Magneticum simulations, fibre magnitudes can be determined by applying a correction to the apparent magnitudes. This correction factor is determined in in two steps. First it is necessary to calculate the size of each galaxy using a stellar mass - size relation and secondly one must predict the surface brightness profile of each galaxy. Eq. 3.6 provides a model for the size of galaxies as a function of their stellar mass for early type galaxies (Shen et al., 2003; Lange et al., 2015).

$$r_e = r_0 \left(\frac{\mathcal{M}_*}{\mathcal{M}_\odot}\right)^{\alpha} \left(1 + \frac{\mathcal{M}_*}{\mathcal{M}_0}\right)^{\beta-\alpha}, \tag{3.6}$$

where r_e is the effective half-light radius in kpc, $\mathcal{M}_*$ is the mass of the galaxy and $\mathcal{M}_0$ is transition mass between low/high-mass galaxies with $\alpha = 0.10$, $\beta = 0.78$, $r_0 = 7.74 \times 10^{-5}$ and $\mathcal{M}_0 = 2.43 \times 10^{10} M_\odot$ (Lange et al., 2015). To convert the total apparent i-band magnitude to an aperture magnitude, a Sérsic profile (Sérsic, 1963; Sersic, 1968; Graham & Driver, 2005) is used. This describes the intensity of a galaxy as a function of radius (Eq. 3.7).

$$I(r) = I_e \exp\left[-b_n\left(\left(\frac{r}{r_e}\right)^{1/n} - 1\right)\right], \tag{3.7}$$

Here the Sérsic index, $n = 4$, for early-type galaxies (a de Vaucouleurs $r^{1/4}$ light profile). I_e is the intensity at the half-light radius. b_n is a function defined such that $\Gamma(2n) = 2\gamma(2n, b_n)$, where Γ is the complete gamma function and γ is the incomplete gamma function (Ciotti, 1991).

- **Target density selection**. The target density is controlled by iterating through the targets starting with the highest priority first and finishing with the lowest priority, with each iteration all other targets that lie within the collision radius of the eBOSS spectrograph are removed from the sample. This approximation does not consider higher order terms such as multiple observations or overlapping tiles.

  For redMaPPer selected galaxies in the observed dataset, membership probabilities are used as the priority metric. In this case, the probability of membership is approximated as the min-max normalisation of the absolute i-band magnitude of the galaxies after pseudo-red-sequence selection in the interval [0.05, 1).

- **Automated membership selection**. The SPIDERS automated cluster redshift pipeline is applied as described in Sect. 2.3. In the case of convergence, this automatically removes most of the spectroscopic contamination, and in some cases also true cluster members that are not strongly bound from a dynamical perspective.

- **Visual inspection**. For the manual visual inspection procedure described in Clerc et al. (2016) and Kirkpatrick et al. (2021), the probability of membership must be visualised as a function of the offset between the spectroscopic galaxy redshift and the photometric cluster redshift. Photometric redshifts are determined for the simulated clusters by degrading their redshifts relative to the measured bias distribution of SPIDERS clusters. This is found to be an approximate normal distribution $\mathcal{N}(0, -0.00236, 0.01636)$. For simplicity this distribution is assumed to be independent of cluster redshift, when in reality, the distribution tends to becomes broader and more biased with increasingly high redshifts.

As a result of this experiment, an estimated 97% of clusters in the SPIDERS DR16 catalogue are expected to have been successfully validated, with $\sim$ 10% subject to significant project effects. This shows that although there is a large improvement relative

to the redshifts provided by photometric only methods, there is still a small (but non-negligible) cosmological implication with respect to cluster count analyses when including high-redshift and low-richness systems (discussed in Sect. 4.1.12).

3.4 Creating a volume limited cluster sample for cosmological analysis

The original visually inspected SPIDERS DR16 catalogue is the largest spectroscopically confirmed sample of X-ray selected galaxy clusters to date with the highest mean number of spectroscopic members per system. Despite this, it is still subject to the shallow photometric depth of the SDSS. When photometric data are incomplete, a redshift dependent correction (scale) factor, is applied to correct the richness (Rykoff et al., 2014). This factor effectively boosts the value and uncertainty of the richness according to Fig. 3.5. As a result, at redshifts above ~ 0.35 the optical mass proxy (richness) has high-scatter and the fraction of successfully spectroscopically confirmed systems drops significantly due to the quality of the optical data.

The impact of the scale-factor is challenging to model accurately. It is therefore beneficial to avoid the complex uncertainties that the scale-factor induces when performing a cosmological analysis with the SPIDERS DR16 cluster sample. To achieve this, the SPIDERS DR16 catalogue must be recreated using the highest quality (deepest) available photometric data. This increases the cosmological constraining power of the sample by reducing scatter for the optical mass proxy, increasing the fraction of systems at $z > 0.35$ that have redshift measurements and reducing the level of contamination due to chance associations with active galactic nuclei at high-redshift. This section describes this catalogue recreation process using The Legacy Surveys, which is carried out prior to the richness-based cosmological analysis described in Sect. 4.1.

3.4.1 DESI Legacy Imaging Surveys

The DESI Legacy Imaging Surveys (hereafter The Legacy Surveys; Dey et al., 2019, $\sim 14,300$ deg^2) consist of The DECam Legacy Survey (DECaLS Dey et al., 2019), The Beijing-Arizona Sky Survey (BASS Zou et al., 2017) and The Mayall z-band Legacy Survey (MzLS Silva et al., 2016). The eighth data release (DR8) of The Legacy Surveys also includes a variety of Dark Energy Camera (DECam) imaging from a variety of other sources such as the Dark Energy Survey (DES; The Dark Energy Survey Collaboration, 2005; Dark Energy Survey Collaboration et al., 2016, $\sim 5,000$ deg^2). These surveys span the entire BOSS (CODEX) footprint (see Fig. 3.6) and therefore provide ancillary photometry in g, r and z bands that is at least 1-2 magnitudes deeper than the SDSS.

The availability of superior photometric data from DR8 of The Legacy Surveys allows one to obtain a lower-scatter optical mass-proxy ($\tilde{\lambda}$) than the original CODEX measurement of richness ($\tilde{\lambda}_{\rm SDSS}$). This is achieved by remeasuring richness while also estimating a new optical centre within 400 kpc of the peak of the extended X-ray emission for each

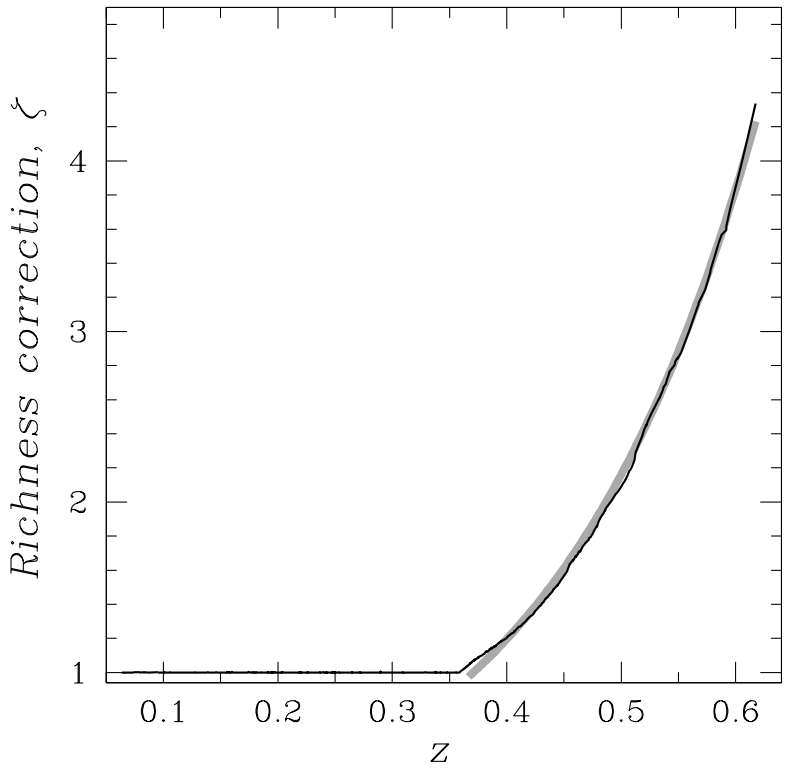

Abbildung 3.5: Adapted from Finoguenov et al. (2020). Multiplicative richness correction applied to the original CODEX cluster catalogue due to photometric depths of the SDSS. The black and grey lines illustrate the correction factor, $\zeta(z) \equiv S(z)$, and the Finoguenov et al. (2020) analytical approximation i.e. $\zeta(z > 0.37) \approx e^{5.5(z-0.35)} - 0.12$. One of the main benefits of reanalysing the CODEX X-ray sources catalogue using The Legacy Surveys to create a cosmological sample of SPIDERS clusters is that this correction factor is not required. This improves completeness of the catalogue, reduces the scatter in the richness-mass relation and simplifies the modelling.

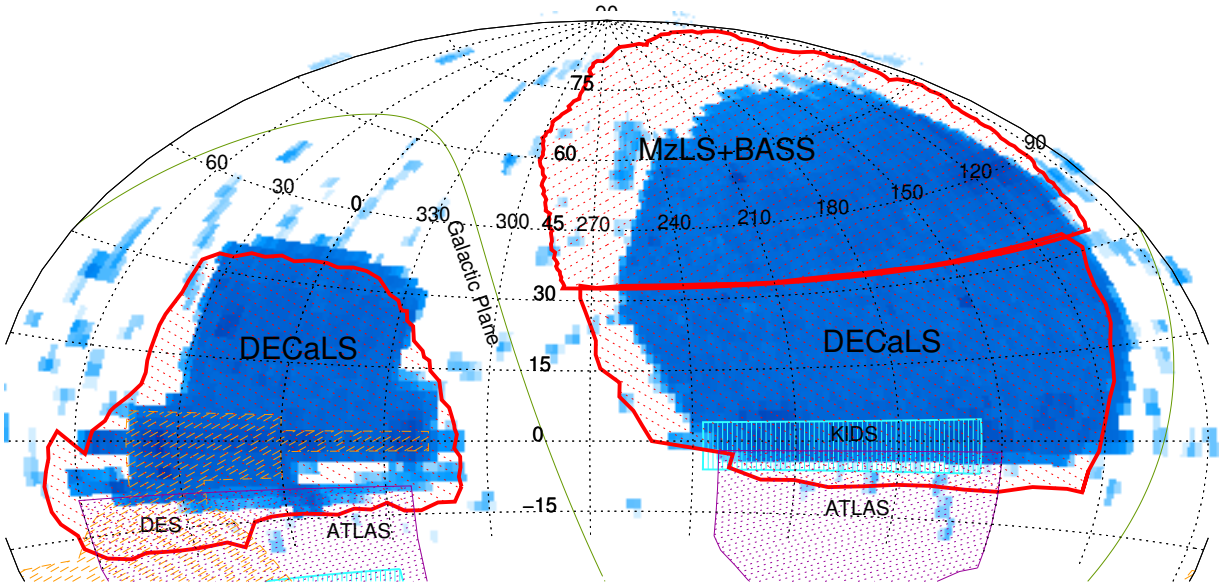

Abbildung 3.6: Adapted from Dey et al. (2019). Footprint of the eighth data release of the DESI Legacy Imaging Surveys. The red borders indicate the areas for each individual survey: DECaLS (Dey et al., 2019), DES (The Dark Energy Survey Collaboration, 2005; Dark Energy Survey Collaboration et al., 2016), MzLs (Silva et al., 2016) and BASS (Zou et al., 2017). Spectroscopic redshifts available in the SDSS (Ahumada et al., 2020), 2dF (Colless et al., 2001) and BOSS (Smee et al., 2013) are shown by blue pixels with darker areas indicating a higher density of spectra. As the SDSS is almost entirely encapsulated by The Legacy Surveys, it provides the best available photometric data to reanalyse the SPIDERS DR16 cluster catalogue for cosmological studies (Ider Chitham et al., 2020). The PS1-3π survey (Chambers et al., 2016) covers the entire figure (declination > -30).

cluster using redMaPPer v6.6 configured in *scanning-mode*. More specifically, it evaluates the optical properties (e.g. richness, redshift, membership and centring probabilities) of an input cluster catalogue at each point on a grid of redshift to determine an initial estimates before remeasuring them without a quantised redshift grid. This mode uses prior knowledge of the positions of cluster centres, producing a sample that is primarily defined by the original selection method used to construct the input catalogue (e.g. X-ray or SZ Finoguenov et al., 2020; Bleem et al., 2020), rather than searching for over-densities of red galaxies in the conventional *cluster-finding* mode which produces purely optically selected samples. The specific *scanning-mode* method used in this analysis determines the maximum likelihood richness (and photometric redshift) before the final percolation procedure, this differs from the *scanning-mode* used to create the original CODEX catalogue (Finoguenov et al., 2020) which determined the maximum richness and respective redshift. The maximum search radius for central galaxy candidates is 400kpc relative to the initial X-ray (RASS) position. This parameter is consistent with Finoguenov et al. (2020).

To ensure the SPIDERS DR16 sample is as complete as possible, spectroscopic redshifts are also re-evaluated by matching previously observed spectroscopic galaxies from the archival catalogue (Sect. 2.3.1) with the newly defined member catalogue, which is generated when remeasuring richness with The Legacy Surveys. The updated spectroscopic galaxies are then re-processed with the original SPIDERS automated cluster redshift pipeline (Sect. 2.3.2) by iteratively considering their distribution in phase space (Clerc et al., 2012, 2020; Kirkpatrick et al., 2021) and performing bootstrap re-sampling. A high-purity sub-sample of these clusters is used to constrain cosmology (Sect. 4.1) for which visual inspection is also carried out by a single inspector[13].

3.4.2 Photometric Configuration

The re-analysis of the SPIDERS cluster sample using The Legacy Surveys is based on a redMaPPer run from $\tilde{z} \in [0.05, 0.72)$ with a z-band (reference band) magnitude limit of 23.5. This makes use of $g-r$ and $r-z$ colours with a transition between the two at a redshift of 0.35. The spectroscopic training set used for the calibration of the default Bruzual & Charlot (2003) red-sequence models easily exceeds the minimum redMaPPer requirements over the redshift range of interest (Rykoff et al., 2014). This calibration procedure is carried out over the entire area of The Legacy Surveys, which uses all available spectroscopic galaxies in the literature, and is repeated for three iterations to ensure convergence.

Although the most significant improvements to the optical richness and contamination fraction are due to the deeper photometry and spectroscopic richness, there are also modest improvements due to updates within the redMaPPer algorithm itself (Rykoff et al., 2016). For instance, the estimation of the colour background is more accurate in redMaPPer v6.6 (relative to v5.2 for CODEX), due to the improved consideration of the local masking information and maps of limiting magnitude that capture the systematic variations in

[13] 7 bootstrap spectroscopic redshift estimates have been adjusted in the visual inspection process after the pipeline raised flags regarding the convergence of the velocity clipping procedure. 4 spectroscopic central galaxies were assigned/adjusted in the visual inspection process due to mis-centring.

the photometry as a function of sky position. The treatment of depth related systematic effects is vital because the observing strategies and depth of each of contributing imaging survey are largely contrasting. The creation of such depth maps and masks are described in Sects. 3.4.2.

Depth Map

The redMaPPer depth map is generated via a parametric depth model (Rykoff et al., 2015) at a maximum resolution of healpix `NSIDE`=4096 (Górski et al., 2005). If a healpixel has an insufficient number of galaxies for the model to converge, the depth is approximated by recursively expanding out to the next largest pixel in the nested scheme until it does.

A brief summary of the model is as follows: A galaxy with flux F (in nanomaggies) is related to its apparent magnitude m by

$$F = 10^{-0.4(m-22.5)} \tag{3.8}$$

This can also be expressed in terms of the expected number of signal photons, $\langle S \rangle$, reaching the detector

$$\langle S \rangle = k t_{\rm eff} F \tag{3.9}$$

Here, $t_{\rm eff}$ is the effective exposure time and k is a constant. The respective number of noise-photons N, is then given by

$$\langle N \rangle = k t_{\rm eff} F_{\rm noise} \tag{3.10}$$

where $F_{\rm noise}$ is the total effective flux due to noise originating from sky noise, read noise etc. With a total of N signal photons, an estimate for the unbiased galaxy flux is given by

$$\hat{F} = \frac{S+N}{k t_{\rm eff}} - F_{\rm noise} \tag{3.11}$$

N and S follow Poisson distributions, therefore the variance of $\hat{F}$ is

$$\sigma_F^2 = \frac{\langle S \rangle + \langle N \rangle}{k^2 t_{\rm eff}^2} = \frac{F + F_{\rm noise}}{k t_{\rm eff}} \tag{3.12}$$

which can also be expressed in terms of the corresponding magnitude error as

$$\begin{aligned} \sigma_m(F; F_{\rm noise}, t_{\rm eff}) &= \frac{2.5}{\ln 10} \frac{\sigma_F}{F} \\ &= \frac{2.5}{\ln 10} \left[\frac{1}{F k t_{\rm eff}} \left(1 + \frac{F_{\rm noise}}{F} \right) \right]^{1/2} \end{aligned} \tag{3.13}$$

In practice, $F_{\rm noise}$ is substituted for the 10σ limiting flux $F_{\rm lim}$ with $F_{\rm lim}/\sigma_F = 10$. It is then possible to solve for $F_{\rm noise}$ using Eq. 3.14, where the limiting magnitude, $m_{\rm lim}$ is the magnitude associated with a galaxy with a flux of $F_{\rm lim}$.

$$F_{\rm noise} = \frac{F_{\rm lim}^2 k t_{\rm eff}}{10^2} - F_{\rm lim} \tag{3.14}$$

Eqs. 3.8, 3.13 and 3.14 now define the model for the magnitude error $\sigma_m(m; m_{\rm lim}, t_{\rm eff})$ of a galaxy of magnitude m which can be calculated for each healpixel of an arbitrary optical/NIR survey. The cost function (Eq. 3.15) is then minimised to measure the 10σ galaxy magnitude limit and exposure times. To minimise the impact of gross outliers, only a subset of galaxies are used with a high signal-to-noise ratio ($\frac{m_{\rm gal}}{\sigma_m} > 5$). Bootstrap re-sampling of the galaxies is also carried out 50 times.

$$E(m_{\rm lim}, t_{\rm eff}) = \sum_{\alpha} \left| \sigma_m^{\rm obs} - \sigma_m(m_\alpha | m_{\rm lim}, t_{\rm eff}) \right| \tag{3.15}$$

Masking

A footprint mask is generated from the bitmasks set in The Legacy Surveys random point catalogues[14]. Points that reduce the fraction of good coverage from unity have the following bitmask[15] bits set:

- **Bright and medium bright stars**. These correspond to stellar sources from GAIA DR2 (Gaia Collaboration et al., 2018) and the Tycho (Høg et al., 2000) catalogues with $G_{\rm Gaia} < 13$ and $G_{\rm Gaia} < 16$ respectively. The corresponding bits are `BRIGHT`, `MEDIUM`. The radius at which stars are masked is determined by Eq. 3.4. This is analogous to Eq. 3.4 and the methods outlined in Sect. 3.2.

$$r\,[\text{arcsec}] = 0.262 \times \min[1800, 150 \times 2.5^{\frac{11-G_{\rm Gaia}}{3}}]. \tag{3.16}$$

- **Bright galaxies** from the Siena Galaxy Atlas (Moustakas et al., 2021) which is largely based on HyperLeda galaxy catalogues (Makarov et al., 2014). This corresponds to the `GALAXY` bit.

- **Globular Clusters & Planetary Nebulae** from the New General Catalogue and Index Catalogue [16]. This corresponds to the `CLUSTER` bit.

- **Bad sources**. If the (reference) z-band is photometric measurement is saturated or masked in anyway or if sources touch a pixel where source fitting failed. These bits correspond to `SATUR_Z`, `ALLMASK_Z` and `BAILOUT`.

- **Non primary sources**. When producing the galaxy source catalogue for redMaPPer to ingest, all sources with a stellar PSF morphological model type or the `NPRIMARY` maskbit is not set are also eliminated. This removes stellar contamination and duplicates sources in the overlap regions between bricks.

[14] http://legacysurvey.org/dr8/files/#random-catalogues
[15] http://legacysurvey.org/dr8/bitmasks
[16] https://github.com/mattiaverga/OpenNGC

To ensure that the cluster catalogue is also volume limited, it is important to extend this mask to the redshift dimension. This combination of the sky and redshift mask is the basis of the optical selection function outlined in Sect. 4.1.5. The variations of the local limiting magnitude of The Legacy Surveys translates to a maximum observable redshift as a function of sky position, $z_{\rm max}$. This directly impacts the fraction of X-ray selected galaxy clusters which can be optically confirmed by redMaPPer, with deeper regions making it possible extend to higher redshifts. This redshift mask is created from the (reference) z-band depth map and the calibrated models of the red-sequence (Sect. 2.2.1) by specifying a galaxy luminosity threshold, $L_{\rm thresh}$, at which members are detectable at least the 10σ level. Clusters which meet the volume limiting criteria of $z \leq z_{\rm max}$ are retained in the final cluster catalogue and therefore regions with extremely shallow data are excluded from the subsequent analysis. This prevents contributions from areas that are likely to induce complex systematic uncertainties. In order to be consistent with the optimal cut used in the redMaPPer luminosity filter, the threshold luminosity is set to $L_{\rm thresh} = 0.2L_{*}$. Although, this is relatively conservative, it is motivated by the fact that the primary use richness, with respect to this revised catalogue, is for it to be used as an optical mass proxy in a cosmological experiment Sect. 4.1 and therefore ensuring that its measurement is not extrapolated in any way simplifies the modelling the uncertainty of the observed richness (i.e. Eq. 4.3).

In order to determine the optical selection function in a computationally efficient way, it is convenient to sample the footprint and redshift masks using random points. As the observability of clusters depends on their projected optical extent on the sky as well as redshift, it necessary to generate and weighted a random cluster catalogue such that it has the same redshift and richness distribution as the volume limited catalogue of observed clusters. This process, first described by Rykoff et al. (2016) and later applied by Ider Chitham et al. (2020); Bleem et al. (2020), is achieved using the following steps:

- Firstly random samples of $\{\lambda, z_{\lambda}\}$ are drawn from the observed cluster catalogue and the redshift mask is used to ensure the randoms sample the survey volume in the same sample.
- The z-band depth map and footprint mask are then used calculated the local masking fraction, $f_{\rm mask}$, and richness extrapolation scale-factor, $S(z)$, so that only random points which have $f_{\rm mask} < 0.2$ and $\lambda/S > 20$ are considered. This must match the same cuts made on the observed cluster catalogue.
- To ensure that regions where clusters are discarded are sufficiently sampled (i.e. poor systems at high-redshift), the randoms are weighted by a factor $w = n_{\rm samp}/n_{\rm keep}$, i.e. the ratio of random points generated from a of $\{\lambda, z_{\lambda}\}$ pair, $n_{\rm samp}$, and the number of points that pass the masking and luminosity threshold criteria, $n_{\rm keep}$.
- This makes sure that the distribution of randoms reflects the data with respect to richness, redshift while considering the same boundaries and depth variations that impact the observations.

- It is then possible to determine the effective area for cluster confirmation as a function of redshift where $z < z_{\rm max}$ by multiplying the area of a single healpixel by number of pixels in the footprint, the fraction of good coverage across the footprint mask and the weights w, provided by the randoms and then smoothing via cubic spline interpolation.

3.5 Application to the next generation of cluster surveys with *eROSITA* & 4MOST

This section applies the methods described in Sect. 3.4.2 to the 4MOST *eROSITA* cluster redshift survey. This includes estimating the effective area as a function of redshift for a variety of optical surveys relevant to the photometric and spectroscopic follow-up of eRASS clusters (Sect. 3.5.1), as well as the development of an algorithm designed to allocate sky fibre positions for the 4MOST spectrograph (Sect. 3.5.2).

3.5.1 Effective optical survey area for *eROSITA*

Given the masking and depth estimation procedure described in the previous subsections, it is possible to estimate the effective optical survey area as a function of redshift. This is a crucial ingredient necessary to perform cosmological cluster count analyses with X-ray selected clusters which utilise optical photometry to estimate their optical properties (e.g. photometric redshift, optical richness etc.). The area curves for a selection of optical surveys are illustrated in Fig. 3.7 as a function of $L_{\rm thresh}$, which corresponds to luminosity cut (how much of the galaxy luminosity function is considered) when generating the position dependent redshift mask applied (Sect. 3.4.2). The upper limits of the shaded regions show the maximum available effective area over the entire sky and the lower limits show the fraction over the *eROSITA*-DE hemisphere. Such curves are the foundation of the volume-limited optical selection function (Sect. 4.1.5) after accounting for the re-weighting using a random cluster catalogue which resembles the real data after accounting for cuts in richness, redshift and the desired maximum masking fraction. Fig. 3.7 clearly demonstrates the effective survey area increases as the luminosity threshold is relaxed from $L_{\rm thresh} = 0.2L_*$ (optimal configuration to minimise the scatter in richness) to $L_{\rm thresh} = 0.4L_*$. This is intuitive as less conservative luminosity cuts result in a higher fraction of the optical survey which can be used to optically confirm clusters at a given redshift.

3.5.2 Sky fibre allocation for the 4MOST spectrograph

The 4MOST *eROSITA* Cluster Redshift Survey (Finoguenov et al., 2019a; de Jong, 2019) aims to obtain secure spectroscopic redshifts for the largest possible number of *eROSITA* clusters which have been selected as extended X-ray sources. As the probability of cluster detection with *eROSITA* is extremely low near the galactic plane (due to absorption of X-ray photons causing a decrement sensitivity), it is only feasible to observe areas at

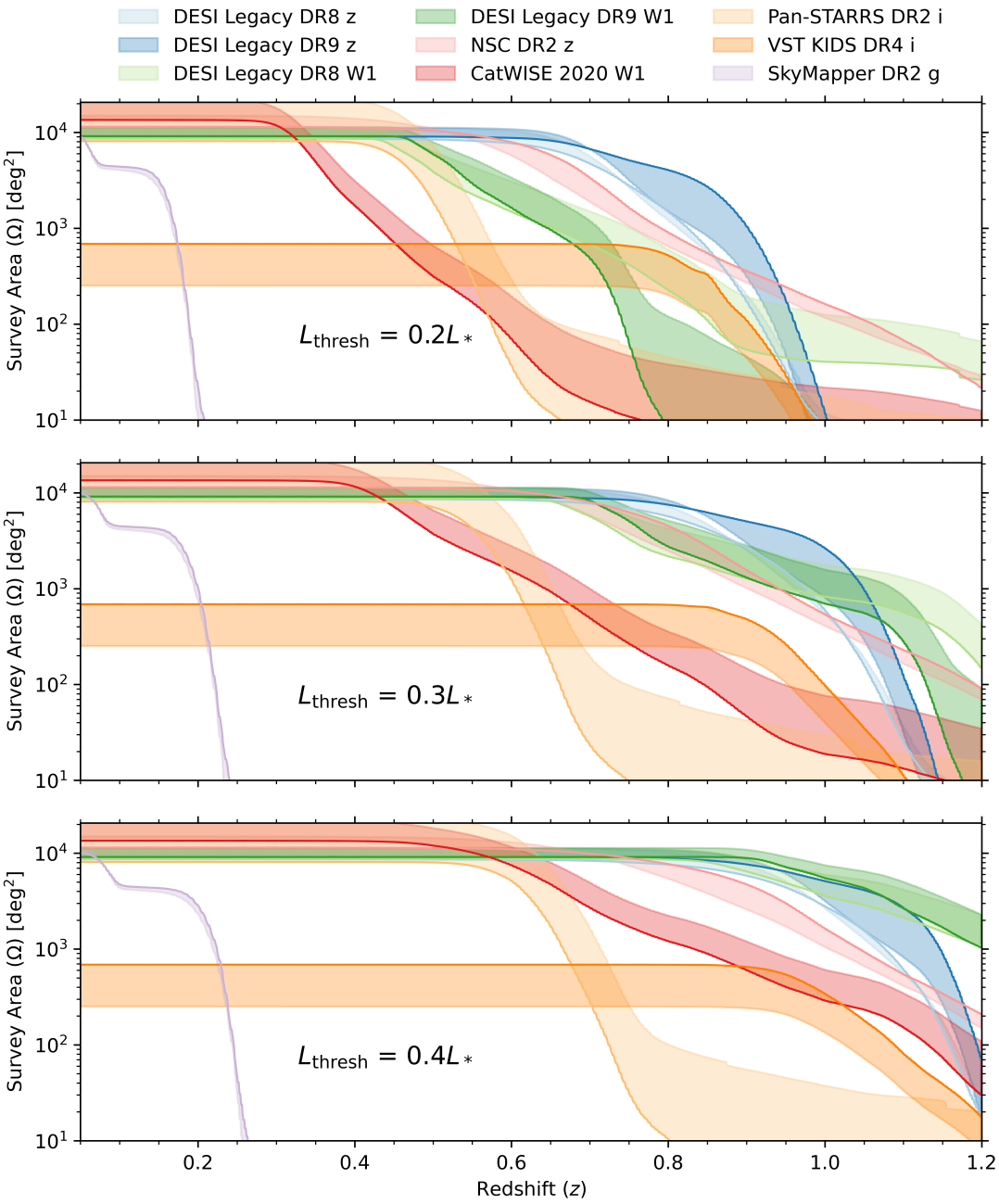

Abbildung 3.7: Adapted from Ider Chitham et al. in prep. Effective area for a selection of optical surveys as a function of redshift and L_{thresh} (the cut applied to the cluster galaxy luminosity function when applying a position dependent redshift mask based on the survey depth). The upper boundary for a given survey represents the maximum area and the lower limit shown by the solid line represents the fraction over the *eROSITA*-DE hemisphere. DESI Legacy Imaging DR8 and DR9 includes the use of the z-band as well as the near infrared $W1$ band provided by unWISE (Mainzer et al., 2014; Meisner et al., 2018). Area curves for CatWISE 2020 (Eisenhardt et al., 2020; Marocco et al., 2021), NSC DR2 (Nidever et al., 2020), SkyMapper DR2 (Onken et al., 2019), Pan-STARRS DR2 (Flewelling, 2018) and VST KIDS DR4 (Kuijken et al., 2019) are also shown according to the colour in the legend.

high Galactic latitude (with $b > 10$). In order to facilitate the measurement of robust spectroscopic cluster redshifts, each cluster needs to be sampled by at least three member galaxies with secure redshifts (e.g. Clerc et al., 2016). For low-redshift clusters detected by *eROSITA* ($z < 0.4$), the projected distance between cluster galaxies exceeds the fibre collision radius due to the fact that they are closer to the observer. They therefore appear more extended, making it possible to obtain a much higher number of galaxy redshifts per cluster especially in the central regions. At the high redshift end of the cluster population ($z > 0.9$), galaxies are faint and therefore challenging to target due to the uncertainty on the photometric information used in the cluster confirmation procedure as well as the long exposure times required to obtain high-signal-to-noise spectra.

For the 4MOST *eROSITA* Cluster Redshift Survey, the target selection is limited to a fibre magnitude of $K_{AB} \lesssim 17.25$ (VISTA) or $r_{AB} < 21.5$ (DECam) resulting in density of approximately two hundred targets per square degree. Consequently, it is only possible to spectroscopically confirm the highest redshift systems via observations of their brightest cluster galaxies (BCGs). Instead of applying a strict magnitude limit for BCGs, observations will continue until a signal-to-noise ratio of five is obtained (required for redshift determination). During spectroscopic observations, a fraction of spectroscopic fibres must be allocated to sky positions to obtain sample spectra in regions where there are no astronomical sources. The outputs of these fibres are then aggregated to form a single, high signal-to-noise sky spectrum, which is then necessarily scaled and subtracted from the spectrum of each object.

The algorithm described below (Ider Chitham et al. in prep) is designed to automatically allocate optimal sky fibre positions for the 4MOST spectrograph. This algorithm has been successfully validated and tested with real observations of clusters using the 2dF spectrograph mounted at Anglo Australian Telescope (Lewis et al., 2002; Sharp et al., 2006) over the South Ecliptic Pole as well as a field south of The Legacy Surveys footprint border[17] (Ider Chitham et al. in prep.):

1. Generate randoms over the `healsparse` spectroscopic survey mask at a minimum resolution of `NSIDE`=2^{16}. This mask can consist of an arbitrary number of circular plates. The density of random points should be at least one thousand times the density of sky fibres as down sampling will occur later in the chain (so the maximum initial density can also be arbitrarily large).

2. Remove areas with compromised photometry according to the astronomical source components of the mask described in Sect. 3.4.2 at a resolution of `NSIDE`=2^{16}. This eliminates regions covered by bright and medium bright stars, bright galaxies, globular cluster, planetary nebulae and the Magellanic clouds.

3. Eliminate randoms within 50 arcsec of any spectroscopic target and within 2 arcsec of any photometric source in any band using the most complete combination of available

[17]These regions are covered by DECam imaging data processed in the second data release of the NOIRLab Source Catalogue (NSC DR2, Nidever et al., 2020, analogous to the data described in Sect. 3.4.1)

optical and near infrared surveys. These exclusion radii are an example used when targeting cluster members (to avoid fibre collisions between sky fibres and science targets). The optimal exclusion radii are dependent on the observing strategy and density of science targets.

4. Remove areas with a low fraction of good photometry coverage across each plate (e.g. assert that it is at least 90%). In the case that the photometric data are not complete it is import to disregard patchy regions which have not been observed to sufficient completeness to ensure that previous step is robust enough. The minimum recommend resolution for the coverage map is `NSIDE`=4096.

5. Down sample randoms to the desired density using the local depth of the primary optical photometry (i.e. the deepest available photometry used in the previous step) as a probability weight. For example, one could select 100 randoms per plate (depending on the needs of the observing strategy) to create a final sample of sky fibre positions in regions where optical photometry is deepest, thereby optimising the previous step(s). The weighting based on the liming depth can also be scaled (in a non-linear fashion) to improve performance further.

A graphical representation is also provided in Fig. 3.8 for a single spectroscopic plate with a two degree field of for spectroscopic observations of cluster galaxies and neighbouring galaxies which reside between the cluster candidates at similar photometric redshifts to considered systems[18]. This test field was chosen for observations due to having the highest density of photometrically selected galaxies in clusters by redMaPPer and two systems at similar redshifts.

[18]These coordinates are artificially centred at the equatorial origin.

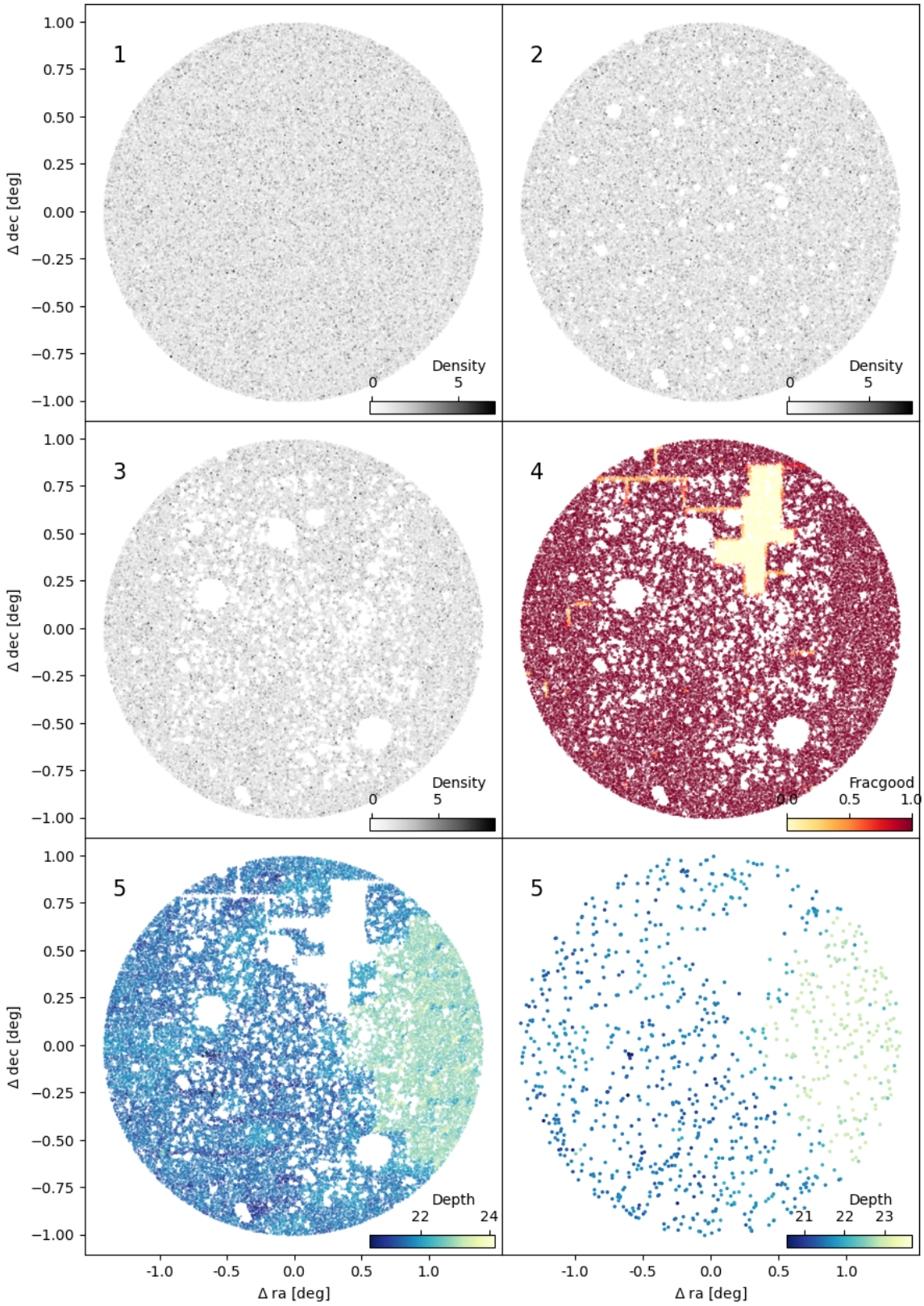

Abbildung 3.8: Adapted from Ider Chitham et al. in prep. Prototype sky fibre algorithm for the 4MOST *eROSITA* cluster redshift and AGN surveys. The number in the top-left corner of each sub-figure corresponds to the numbered list of steps outlined in Sect. 3.5.2.

Kapitel 4

Results: cosmological interpretation of cluster abundances

This chapter presents a cosmological analysis based on the properties of X-ray selected clusters of galaxies from the CODEX survey which have been spectroscopically followed up within the SPIDERS programme as part of the sixteenth data release (DR16) of SDSS-IV. Sect. 4.1 describes a cosmological analysis using optical richness as a mass proxy and Sect. 4.2 provides a summary of an analogous analysis which has been modified to use velocity dispersion instead of optical richness. These sections are derived entirely from the author's major contributions to Ider Chitham et al. (2020) and Kirkpatrick et al. (2021) respectively. The description of the X-ray selection function is adapted from Ider Chitham et al. (2020); Finoguenov et al. (2020).

When referring to observables throughout this chapter, the following notation is consistently used: a tilde above an observable $\tilde{\mathcal{O}}$ symbolises an *observed* quantity; one without a tilde, $\mathcal{O}$, refers to the *true*, unobserved quantity (e.g, of the underlying halo). Conditional probabilities such as the probability of A given B_μ are denoted as $P(A \,|\, B)$.

4.1 SPIDERS Cosmology: Optical Richness

The cosmological sub-sample of SPIDERS clusters used in this section contains a total of 691 clusters over an area of 5,350deg^2 with newly measured optical properties provided by a reanalysis of the CODEX source catalogue using redMaPPer and the DESI Legacy Imaging Surveys (DR8, as described in Sect. 3.4). Optical richness is used as a proxy for the cluster mass, and the combination of X-ray, optical and spectroscopic information ensures that only confirmed virialised systems are considered. Clusters are binned in observed redshift, $\tilde{z} \in [0.1, 0.6)$ and optical richness, $\ln \tilde{\lambda} \in [25, 148)$ and the number of clusters in each bin is modelled as a function of cosmological and richness-mass scaling relation parameters. A high-purity sub-sample of 691 clusters is used in the analysis and best fit cosmological parameters are found to be $\Omega_{m_0} = 0.34^{+0.09}_{-0.05}$ and $\sigma_8 = 0.73^{+0.03}_{-0.03}$. The redshift evolution of the self-calibrated richness-mass relation is poorly constrained due to

the systematic uncertainties associated with the X-ray component of the selection function (which assumes a fixed X-ray luminosity-mass relation with $h = 0.7$ and $\Omega_{m_0} = 0.30$). Repeating the analysis with the assumption of no redshift evolution is found to improve the consistency between both cosmological and scaling relation parameters with respect to recent galaxy cluster analyses in the literature. All other cosmological parameters are fixed to their fiducial values assuming a flat ln λCDM cosmological model with a constant dark energy equation of state $w = -1$, a present day CMB temperature of 2.7255K (Fixsen, 2009) and 3.046 effective neutrino species (Planck Collaboration et al., 2018) with $H_0 = 70 \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $\Omega_{b_0} = 0.048$ and $n_s = 0.96$.

The definition of richness is sensitive to the calibration procedure within redMaPPer. Influential factors include the initial spectroscopic galaxy training set, the number of iterations used to tune the models of red-sequence galaxies, the measurement of the background as well as the quality of the photometry (i.e. the level of optimisation with respect to colour and total flux). As almost all these things differ from the original CODEX redMaPPer processing with the SDSS, the remeasured richness based on The Legacy Surveys is found to be systematically lower than the original $\tilde{\lambda}_{\mathrm{SDSS}}$ definition. The median ratio between the two definitions of richness is found to evolve with redshift as

$$\tilde{\lambda} = \frac{\tilde{\lambda}_{\mathrm{SDSS}}}{1.04 + 0.17e^{5.4(\tilde{z}-0.36)}} \,. \tag{4.1}$$

For SDSS redMaPPer selected clusters, richness extrapolation occurs when $\tilde{z} \gtrsim 0.36$ (Rykoff et al., 2014; Finoguenov et al., 2020). Such a boosting effect can result in richness being systematically up-scattered, hence the exponent in the denominator of Eq. 4.1.

4.1.1 SPIDERS spectroscopy

The SDSS-IV/SPIDERS spectroscopic data span 5,350deg^2 out of the total 10,800deg^2 SDSS-III/BOSS footprint which is covered by CODEX. Within this area, the total number of spectroscopically validated and visually inspected clusters of galaxies with $\tilde{\lambda}_{\mathrm{SDSS}} > 10$ is 2,740 i.e. about 0.5 per square degree (Clerc et al., 2020; Kirkpatrick et al., 2021). For systems with $\tilde{\lambda}_{\mathrm{SDSS}} > 40$, a total of 920 out of 1047 were confirmed. The remaining 127 are either dubious candidates or high redshift systems ($\tilde{z} > 0.7$) lacking spectra.

The range over which observables (redshift and richness) are modelled directly impacts the accuracy and precision of the posterior distributions of the cosmological and richness-mass relation parameters derived in this work. Due to the trade-off between accuracy and precision, one must optimise the size of the cluster sample to ensure that it is statistically significant enough to be sensitive to the parameters of the model while simultaneously excluding regions of the observable parameter space which are challenging to characterise. The latter is necessary to reduce the influence of unquantifiable systematic uncertainties. This section justifies the cuts applied to the observable parameter-space which define the cosmological cluster sub-sample used in this analysis. Fig. 4.1 illustrates the redshift and richness distribution of 4,448 CODEX clusters from The Legacy Surveys volume-limited[1]

[1]For details of the volume limiting procedure, please refer to Sect. 4.1.5.

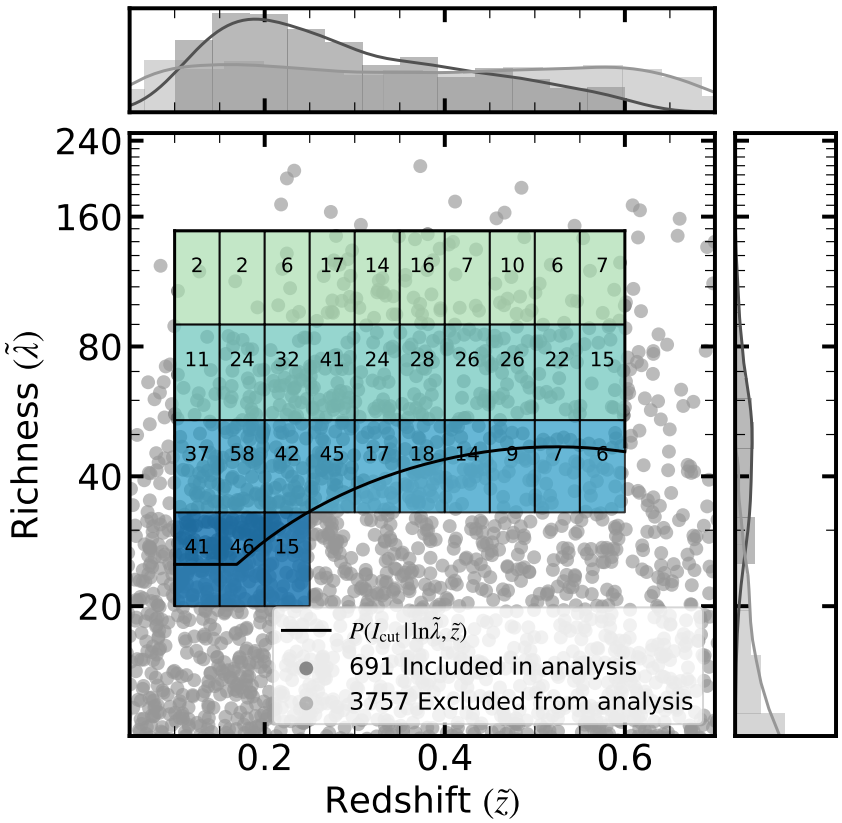

Abbildung 4.1: Taken from Ider Chitham et al. (2020). The volume limited cluster sample used in the cosmological analysis produced by analysing the original CODEX source catalogue with redMaPPer and The Legacy Surveys over the SPIDERS DR16 footprint. The solid black line represents the redshift dependent richness cut described by $P(I_{\rm cut} \mid \ln\tilde{\lambda}, \tilde{z})$ (Eq. 4.19). Boxes represent the bins used in this analysis, annotated with the respective clusters counts (post richness cut) and colours represent the richness bins highlighted in Fig. 4.2. The dark grey distributions illustrate clusters that are included in this analysis i.e. the 691 out of 4,448 clusters that fall within the coloured boxes *and* above the black line. The light grey distributions illustrate clusters excluded from the analysis.

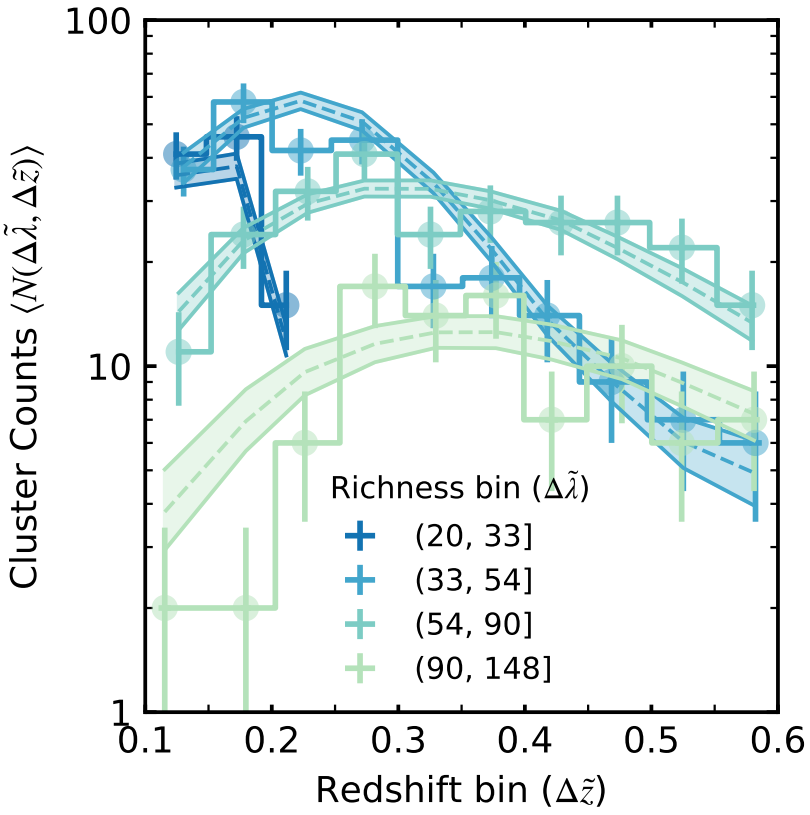

Abbildung 4.2: Taken from Ider Chitham et al. (2020). The abundance of SPIDERS clusters as a function in bins of observed redshift ($\Delta\tilde{z}_j$) and richness ($\Delta\ln\tilde{\lambda}_i$) where $\tilde{z} \in [0.1, 0.6)$ and $\tilde{\lambda} \in [25, 148)$. Steps represent the observed data, the width and height correspond to the size of the bin and the magnitude of the diagonal (Poisson) elements of the covariance matrix, respectively. The position of the vertical bars indicates the mean redshift in each bin. The shaded regions trace the expectation value provided by the model (with a normal prior on the intrinsic scatter Bleem et al., 2020), centred on the median, which corresponds to the best-fit cosmology. The lower and upper limit are similarly set by the 15% and 85% confidence intervals. These distributions are calculated directly from the stored expectation values of cluster counts for the MCMC chains used to create the contours shown in Fig. 4.10.

.

redMaPPer catalogue constructed over the SPIDERS DR16 footprint. The cosmological sub-sample of 691 clusters is shown by the points in the coloured boxes that are above the redshift dependent richness cut (Eq. 4.19) where the X-ray selection function is most sensitive (Sect. 4.1.5). The abundance modelling for these clusters is described in the subsequent section (Sect. 4.1.2) and presented in Fig. 4.2.

659 clusters in the cosmological sub-sample have more than three spectroscopic member galaxies and an additional 26 systems have a single visually inspected central spectroscopic galaxy. This results in a mean of 15.7 spectroscopic members per cluster redshift, as shown top panel of Fig. 4.3, with statistical uncertainties of the order $\Delta_z/(1+\tilde{z}) \sim 7 \times 10^{-4}$. The best available redshifts for the remaining 6 clusters are provided by photometric redshift estimates obtained during the richness re-measurement process (detailed in Sect. 3.4.2).

The lower-limit of the observed-redshift range is set to 0.1 to avoid sub-optimal redMaPPer performance. Although this is less conservative than the limit of 0.2 used by Abbott et al. (2020), it is justified by the fact that photometric redshifts of clusters in the cosmological sub-sample for $\tilde{z} \in [0.1, 0.2)$ have extremely low-bias (~ 0.002) as well as the largest number of spectroscopic members per cluster ($N_{\rm mem} \sim 20$) as shown in Fig. 4.3. The upper redshift limit is set to 0.6 to exclude clusters in the regime of spectroscopic incompleteness (Clerc et al., 2020).

Clusters with $\tilde{\lambda} < 25$ are also excluded from this analysis. This hard limit is motivated by the fact that Abbott et al. (2020) showed that excluding $\tilde{\lambda}_{\rm DES} < 30$ systems, systematically shifts the posterior distribution of Ω_{m_0} closer to that constrained by other cosmological probes (Beutler et al., 2011; Ross et al., 2015; Alam et al., 2017; Scolnic et al., 2018), indicating the presence of unknown effects related to low-richness systems. The numerical difference relative to the hard-limit imposed here is because the definition of richness for The Legacy Surveys is systematically lower ($\tilde{\lambda} \approx 0.87\tilde{\lambda}_{\rm DES}$, Sect. 4.1.9). On top of this hard-limit at low-richness, an additional conservative redshift-dependent richness cut (Sect. 4.1.5) is applied. This ensures an even smaller proportion of low-richness systems are considered above a redshift of 0.2 (Sect. 4.1.5 and Fig. 4.1). Clusters with $\tilde{\lambda} > 148$ are also excluded from this analysis as they make up only 8/691 when $\tilde{z} \in [0.1, 0.6)$ and due to their rarity, a Gaussian likelihood model is not justified (Rozo et al., 2010).

Photometric clusters make up less than one percent of the cosmological cluster sub-sample and have a mean redshift and richness of 0.40 and 73.6 respectively. Of the 685 spectroscopic clusters, the fraction of systems with spectroscopic and photometric redshifts that differ by more than three times the estimated photometric uncertainty is also less than one percent (red outliers are shown in the middle panel of Fig. 4.3). The dashed purple and dotted red lines in the lower panel of Fig. 4.3 indicate that the photometric redshifts used in the cosmological cluster sub-sample are almost unbiased ($\tilde{z} - \tilde{z_\lambda} \sim 0.005$), with low scatter ($\sigma_z/(1+\tilde{z}) \sim 0.01$).

Conducting a cosmological analysis for the 2,740 spectroscopically validated clusters in the original SPIDERS DR16 sample with a revised mass-proxy from The Legacy Surveys would require the completeness of optical and spectroscopic observations as a function of cluster properties for the SDSS (Clerc et al., 2020) to be modelled. Given the sparsity and high quality of the redMaPPer photometric redshifts, the systematic effect impact

of supplementing the spectroscopic sub-sample with 6 photometric clusters is negligible relative to the benefits of simplifying the overall selection function.

4.1.2 Cosmological model

This section presents the modelling used to interpret the cluster number counts in the observed data. The adopted form for the HMF is that of Tinker et al. 2008, with an accuracy of at the $10-20\%$ level, which is sufficient for this study. This HMF is mapped to the richness observable through a scaling relation (Sect. 4.1.3). The selection function (Sect. 4.1.5) is then convolved to obtain a prediction of the number counts in the same binning scheme as the observations (Sect. 4.1.4). Finally, cosmological and scaling relations parameters are jointly fit by comparing the model to the data assuming a Poisson likelihood (Sect. 4.1.6). To ensure the likelihood function and code works as expected, a series of validation tests have also been carried out as detailed in Sect. 4.1.8.

Cluster masses and radii are consistently defined by a spherical over-density 200 times that of the critical density of the Universe (unless stated otherwise).

4.1.3 Scaling Relation

The scaling relation relates the true halo mass ($\mu = \ln M$) to the true richness ($\ln \lambda$) as a function of true redshift (z). The natural logarithm of the expected richness is given by

$$\langle \ln \lambda(\mu, z) \rangle = A + B\,(\mu - \mu_{\rm pivot}) + B_z \ln\left(\frac{1+z}{1+z_{\rm pivot}}\right) . \tag{4.2}$$

Here, $z_{\rm pivot}$ and $\mu_{\rm pivot}$ are pivot values of redshift and the mass which equate to 0.263 and the natural logarithm of $3.08 \times 10^{14} {\rm M}_\odot h^{-1}$. For convenience, these values are chosen to be the median redshift and mass for the cosmological cluster sub-sample (Sect. 4.1.1) when assuming masses inferred from X-ray luminosity (Finoguenov et al., 2020). The remaining quantities are dimensionless free parameters: A is the expectation value at the pivot point $\langle \ln \lambda(\mu_{\rm pivot}, z_{\rm pivot}) \rangle$, B_μ is the coefficient of halo mass dependence, B_z is the linear coefficient of redshift evolution.

The overall variance on richness is modelled as

$$\sigma^2(\mu, z) = \sigma^2_{\rm int}(\mu, z) + \sigma^2_{\rm noise}(\ln \lambda, z) = \sigma^2_{\rm int} + \lambda^{-1} . \tag{4.3}$$

where $\sigma_{\rm int}$ is the intrinsic scatter about the expected richness, which is assumed to be independent of halo mass and true redshift (e.g. Capasso et al., 2019c), and $1/\sqrt{\lambda}$ is the Poisson noise on the true richness from summing the membership probabilities of galaxies.

The priors for the richness-mass scaling relation parameters are summarised in Table. 4.3 and initialised to the best-fit values of Capasso et al. (2019c). The prior for the normalisation of the richness-mass relation is set to the natural logarithm of the observed-richness range used in this analysis i.e. $A \in [3.0, 5.0]$. As richness scales with mass, the prior of mass dependency is set to enforce a positive trend with a wide range of possible

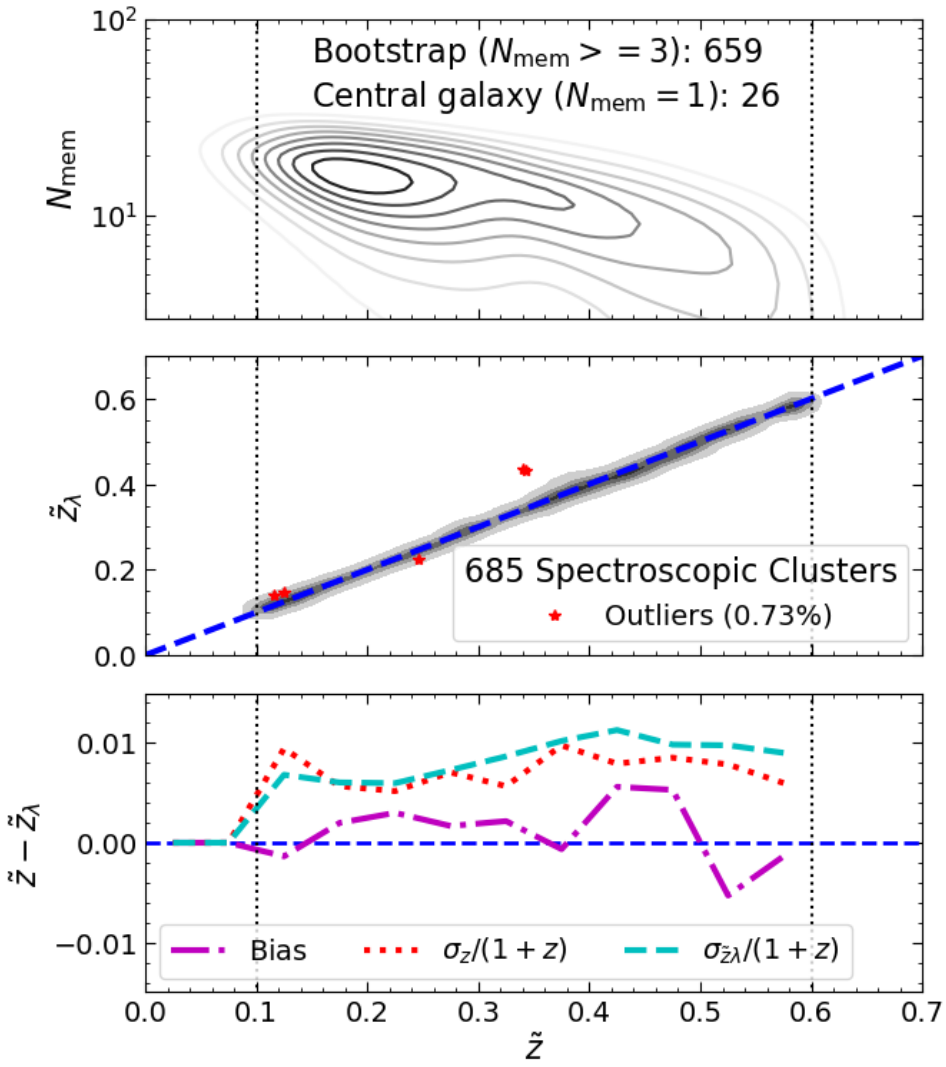

Abbildung 4.3: Taken from Ider Chitham et al. (2020). Cosmological cluster sub-sample after remeasuring spectroscopic redshifts for CODEX clusters using member galaxies selected as part of the reanalysis with The Legacy Surveys (Sect. 4.1.1). Upper panel: distribution of number of spectroscopic members, $N_{\rm mem}$ as a function of redshift. Middle panel: spectroscopic redshift $\tilde{z}$ versus photometric redshift $\tilde{z}_\lambda$. Outliers are defined when the discrepancy between spectroscopic and photometric redshifts is more than three times the uncertainty on the photometric redshift ($\tilde{z} - \tilde{z_\lambda} > 3\sigma_{\tilde{z}_\lambda}$). Lower panel, median quantities over each redshift bin used in this analysis (bin width of 0.05) bias: $(\tilde{z} - \tilde{z_\lambda})$, scatter: $\sigma_z/(1+\tilde{z}) = 1.4862 \times (|\tilde{z} - \tilde{z_\lambda}| - (\tilde{z} - \tilde{z_\lambda}))/(1+\tilde{z})$ and photometric redshift uncertainties: $\sigma_{\tilde{z}_\lambda}/(1+\tilde{z})$.

values $B_\mu \in [0.0, 10.0)$ (e.g. Kiiveri et al. subm.). The prior for the redshift evolution is set to $B_z \in [-5.0, 5.0]$ (Abbott et al., 2020) and several different priors are considered for the intrinsic scatter in Sect. 4.1.11. These intrinsic scatter priors include that of the recent DES analysis, $\sigma_{\rm int} \in [0.1, 0.5]$ (justified in Appendix B of Abbott et al., 2020), a more restrictive variant of this, $\sigma_{\rm int} \in [0.1, 0.3]$, as well a normal prior centred on the best fit value from Bleem et al. (2020). In addition to these basic priors, restrictions are also made to ensure only physically plausible parameter combinations are considered i.e. $\langle \ln \lambda(\mu, z) \rangle > 1$ for all values of true mass and true redshift.

4.1.4 Expected number counts

The expectation value of the number of galaxy clusters per bin $N(\Delta \ln \tilde{\lambda}_i, \Delta \tilde{z}_j)$ is given by Eq. 4.4, where the integration limits are consistently reduced to the subset of the parameter-space where the X-ray selection function is sensitive (Sect. 4.1.5) i.e. $z \in [0.05, 0.75]$, $\nu \in [-4, 4]$ and $\mu \in [30.7, 35.3]$ (or equivalently $\log_{10}(M_{200c}) \in [13.5, 15.5] \mathrm{M}_\odot$).

$$\langle N(\Delta \ln \tilde{\lambda}_i, \Delta \tilde{z}_j) \rangle = \int_{0.05}^{0.75} \mathrm{d}z \; \frac{\mathrm{d}V(z)}{\mathrm{d}z} \cdot \langle n(\Delta \ln \tilde{\lambda}_i, z) \rangle \int_{\Delta \tilde{z}_j} \mathrm{d}\tilde{z} \; P(\tilde{z} \,|\, z, \Delta \ln \tilde{\lambda}_i) \,, \tag{4.4}$$

Here $\mathrm{d}V(z)/\mathrm{d}z$ is differential comoving volume element for a flat universe (Hogg, 1999) and $P(\tilde{z} \,|\, z, \Delta \ln \tilde{\lambda}_i) = \mathcal{N}(z, \langle \tilde{z}_{ij} \rangle, \sigma_{\tilde{z}_{ij}})$, i.e. a normal distribution accounting for uncertainties on the mean observed redshift in each bin[2]. The quantity $\langle n(\Delta \ln \tilde{\lambda}_i, z) \rangle$ is the expected comoving number density of halos in the richness bin.

$$\langle n(\Delta \ln \tilde{\lambda}_i, z) \rangle = \int_{30.7}^{35.3} \mathrm{d}\mu \frac{\mathrm{d}n(\mu, z)}{\mathrm{d}\mu} \int_{\Delta \ln \tilde{\lambda}_i} \mathrm{d}\ln \tilde{\lambda} \; P(\ln \tilde{\lambda} \,|\, \mu, z) \,. \tag{4.5}$$

This calculation depends on the differential HMF defined in terms of μ; the natural logarithm of the halo mass within a radius, R, as well as $\sigma(\mu)$; the RMS fluctuation of the smoothed matter density field. Here $f(\sigma)$ is the Tinker et al. (2008) multiplicity function[3].

$$\frac{\mathrm{d}n(\mu, z)}{\mathrm{d}\mu} = \frac{3}{4\pi R^3(\mu, z)} \frac{\mathrm{d} \ln \sigma(\mu)^{-1}}{\mathrm{d}\mu} f(\sigma(\mu), z) \,. \tag{4.6}$$

A hard prior on Ω_{m_0} is given by the *mean* spherical over-density relative to the range used to calibrate the Tinker et al. 2008 HMF. When using a critical over-density definition of mass this prior translates as $\Delta_m(z) \equiv \Delta_c \Omega_m(z) \in (200, 3200)$ (Bocquet et al., 2019).

4.1.5 Selection function

The cosmological sub-sample of SPIDERS DR16 cluster is obtained via the aggregation of multiple selection processes: primary X-ray selection, followed by secondary optical selection. These components of the total selection are described in the following subsections.

[2]The PDF of observed redshift given the true redshift is modelled as a Dirac delta as justified in Sect. 4.1.12

[3]Eq. 4.6 is calculated using the `colossus` package (Diemer, 2018).

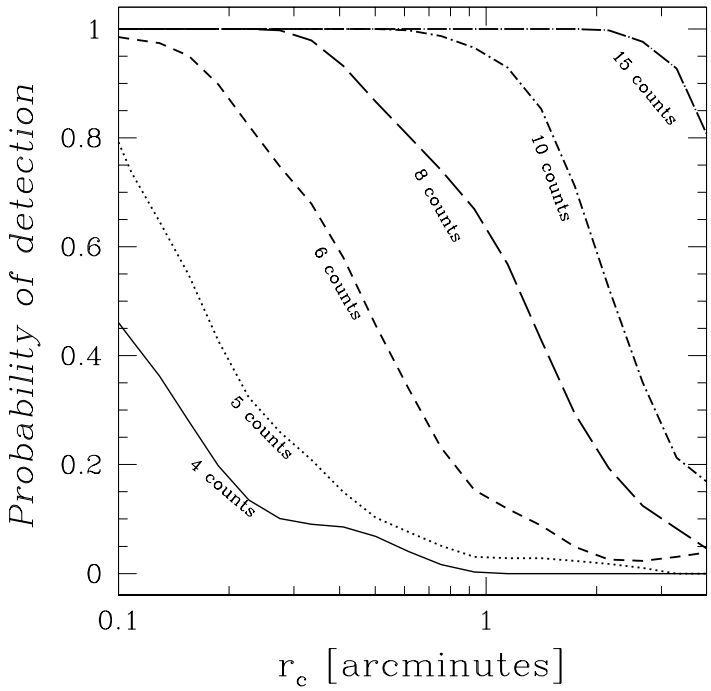

Abbildung 4.4: Adapted from Finoguenov et al. (2020). CODEX probability of X-ray cluster detection as a function of core radius $P(I|r_c, \eta^{\rm ob}, \beta(\mu))$. Each line corresponds to evaluating the selection function at different a number of X-ray photon counts $\eta^{\rm ob} \in \{4, 5, 6, 8, 10, 15\}$. r_c represents the core radii of the CODEX cluster cluster sample at $z = 0.25$.

The variable I is used to denote *selection* and each component of the total selection $I_{\rm tot}$ has a representative subscript e.g. $I_{\rm X}$ indicates X-ray selection, $I_{\rm Legacy}$ represents the selection due to applying redMaPPer to The Legacy Surveys and $I_{\rm cut}$ describes the redshift dependent richness cut applied to the X-ray selected clusters.

X-ray selection function[4]

The probability of detecting CODEX X-ray sources as a function as a function of X-ray photon counts ($\eta^{\rm ob}$) and surface brightness (S_B) shape parameters is denoted as $P(I|\eta^{\rm ob}, S_B)$. Shape parameters sample the parameters of a β-profile according to

$$S_B(r) = \left(1 + \left(\frac{r}{R_c}\right)^2\right)^{-3\beta+0.5}, \tag{4.7}$$

[4]The description of the CODEX X-ray selection function is adapted from Finoguenov et al. (2020) and Ider Chitham et al. (2020).

where r represents radius and R_c represents the core radius. Eq. 4.7 is normalised to the total number of true counts (η) and core radii are sampled as a function of temperature, T, using a fixed $\beta - T$ relation (Käfer et al., 2019).

In addition to the covariance between X-ray luminosity and shape, the covariance between X-ray luminosity and optical richness (Farahi et al., 2019) is also considered. This results in a distribution which describes the probability at each point in the observable parameter space given mass and redshift

$$P(\ln L_X, \ln R_c, \ln \lambda | \mu, z) = \frac{1}{(2\pi)^{3/2}|\Sigma|^{1/2}} \exp\left[-\frac{1}{2}\boldsymbol{X}^T \Sigma^{-1} \boldsymbol{X}\right]. \tag{4.8}$$

Here, the vector $\boldsymbol{X}$, is defined by three redshift dependent observable-mass relations (Käfer et al., 2019; Mulroy et al., 2019; Capasso et al., 2019c)

$$\boldsymbol{X} = \begin{pmatrix} \ln L_X - \langle \ln L_X | \mu, z \rangle \\ \ln R_c - \langle \ln R_c | \mu, z \rangle \\ \ln \lambda - \langle \ln \lambda | \mu, z \rangle \end{pmatrix}. \tag{4.9}$$

with the covariance matrix defined by

$$\Sigma = \begin{pmatrix} \sigma^2_{\ln L_X|\mu} & \rho_{\ln L_X \ln R_c|\mu}\sigma_{\ln L_X|\mu}\sigma_{\ln R_c|\mu} & \rho_{\ln L_X \ln \lambda|\mu}\sigma_{\ln L_X|\mu}\sigma_{\ln \lambda|\mu} \\ \rho_{\ln L_X \ln R_c|\mu}\sigma_{\ln L_X|\mu}\sigma_{\ln R_c|\mu} & \sigma^2_{\ln R_c|\mu} & \rho_{\ln R_c \ln \lambda|\mu}\sigma_{\ln R_c|\mu}\sigma_{\ln \lambda|\mu} \\ \rho_{\ln L_X \ln \lambda|\mu}\sigma_{\ln L_X|\mu}\sigma_{\ln \lambda|\mu} & \rho_{\ln R_c \ln \lambda|\mu}\sigma_{\ln R_c|\mu}\sigma_{\ln \lambda|\mu} & \sigma^2_{\ln \lambda|\mu} \end{pmatrix}. \tag{4.10}$$

Here, $\rho_{\ln R_c l|\mu} = -0.3$, $\sigma_{\ln R_c|\mu} = 0.36$ (Käfer et al., 2019), $\sigma_{\ln \lambda|\mu} = 0.2$ (Capasso et al., 2019c; Mulroy et al., 2019), $\rho_{\ln L_X \ln \lambda|\mu} = -0.3$ (Farahi et al., 2019), $\sigma_{\ln L_X|\mu} = 0.46(1 - 0.61z)$ (Mantz et al., 2016) and $\rho_{\ln R_c \ln \lambda|\mu} = 0.0$[5].

The dependency on $\ln \lambda$ only influences the selection matrix by offsetting the distributions of $\ln L_X$ and $\ln R_c$. It is therefore convenient to re-parameterise Eq. 4.8 in terms of ν (Eq. 4.11), which is defined as the fraction of the deviation of the true richness from its expectation value (Eq. 4.2) and the intrinsic scatter ($\sigma_{\rm int}$):

$$\nu \equiv \frac{\ln \lambda - \langle \ln \lambda(\mu, z) \rangle}{\sigma_{\rm int}}, \qquad \frac{{\rm d}\nu}{{\rm dln}\,\lambda} = \frac{1}{\sigma_{\rm int}}. \tag{4.11}$$

This makes it possible to independently vary $P(\ln \lambda | \mu, z)$ as a function richness-mass relation parameters (Eqs. 4.2 and 4.3) by splitting Eq. 4.8 into its separate components. This is useful as it avoids unnecessarily re-calculating the (computationally expensive) selection term.

$$P(\ln R_c, \ln L_X, \ln \lambda | \mu, z) = P(\ln R_c, \ln L_X | \mu, \nu, z) \cdot P(\ln \lambda | \mu, z). \tag{4.12}$$

While the remaining term can be decomposed further and analytically inverted into the following multivariate log-normal distribution

$$P(\ln R_c, \ln L_X | \mu, \nu, z) = P(\ln R_c | \mu, \ln L_X, z) \cdot P(\ln L_X | \mu, \nu, z) \tag{4.13}$$

[5]No covariance is assumed due to the the lack of a published reference at the time prior to the submission of Finoguenov et al. (2020)

where

$$P(\ln R_c|\mu, \ln L_X, z) = \mathcal{N}(\ln L_X, \langle \ln L_X|\mu, z\rangle + \rho_{\ln L_X \nu|\mu}\nu\sigma_{\ln L_X|\mu}, \sigma_{\ln L_X|\mu}\sqrt{1-\rho^2_{\ln L_X \ln\lambda|\mu}}) \tag{4.14}$$

and

$$P(\ln L_X|\mu,\nu,z) = \mathcal{N}(\ln R_c, \langle \ln R_c|\mu, z\rangle + \frac{\rho_{\ln L_X \ln R_c|\mu}(\ln L_X - \langle \ln L_X|\mu, z\rangle)\sigma_{\ln R_c|\mu}}{\sigma_{\ln L_X|\mu}\sqrt{1-\rho^2_{\ln L_X \ln\lambda|\mu}}}, \sigma_{\ln R_c|\mu}\sqrt{1-\rho^2_{\ln L_X \ln R_c|\mu}}) \tag{4.15}$$

These formulae demonstrate that covariance results in a systematic shift of the X-ray selection function distribution. This as illustrated as a modified probability distribution for non-zero values of ν as illustrated in Fig. 4.5.

Denoting the survey area Ω (deg^2) and sensitivity S (ergs s^{-1} cm^{-2}), which includes the effects of exposure and n_H, the CODEX survey selection function is defined as

$$\begin{aligned} P(I_{\rm X}\,|\,\mu,z,\nu)\Omega_{tot} = \int {\rm d}S \frac{{\rm d}\Omega}{{\rm d}S} \iiint {\rm d}\ln L_X\,{\rm d}\ln R_c\,{\rm d}\tilde{\eta}\cdot P(\ln R_c, \ln L_X|\mu,\nu,z) \\ \cdot P(\tilde{\eta}|\eta(\ln L_X, S, z)) \\ \cdot P(I|\tilde{\eta}, \beta(\mu), \ln R_c)\,, \end{aligned} \tag{4.16}$$

where

$$P(\tilde{\eta}|\eta) = \frac{\eta^{\tilde{\eta}}(\ln L_X, S, z)e^{-\eta(\ln L_X, S, z)}}{\tilde{\eta}!}\,. \tag{4.17}$$

Converting luminosity to counts uses the luminosity distance to the object $d_L(z)$, sensitivity S (counts per flux in ergs cm^{-2} s^{-1}) and K-correction $K(\langle T|L_X\rangle, z)$:

$$\eta(\ln L_X, S, z) = \frac{L_X S}{4\pi d_L(z)^2 K(\langle T|L_X\rangle, z)}\,. \tag{4.18}$$

The definition of $P(I_{\rm X}\,|\,\mu,z,\nu)$ includes a volume correction due to the survey mask described in Sect. 3.4.1. $P(I_{\rm X}\,|\,\mu,z,\nu)$ also has a small cosmological dependence via the luminosity distance and its impact increases with redshift. In this analysis, $P(I_{\rm X}\,|\,\mu,z,\nu)$ is evaluated at the fiducial cosmology $\Omega_{m_0} = 0.30$ and $h = 0.7$. This probability distribution is illustrated in Fig. 4.5 and systematic uncertainties due to this approximation are estimated in Sect. 4.1.12.

redMaPPer selection function

The first optical selection term, $P(I_{\rm Legacy}\,|\,\ln\tilde{\lambda},\tilde{z})$, accounts for any differences in the area coverage and masking between in the original SDSS based SPIDERS footprint and the one based on The Legacy Surveys used in this analysis. It also considers the fact that the

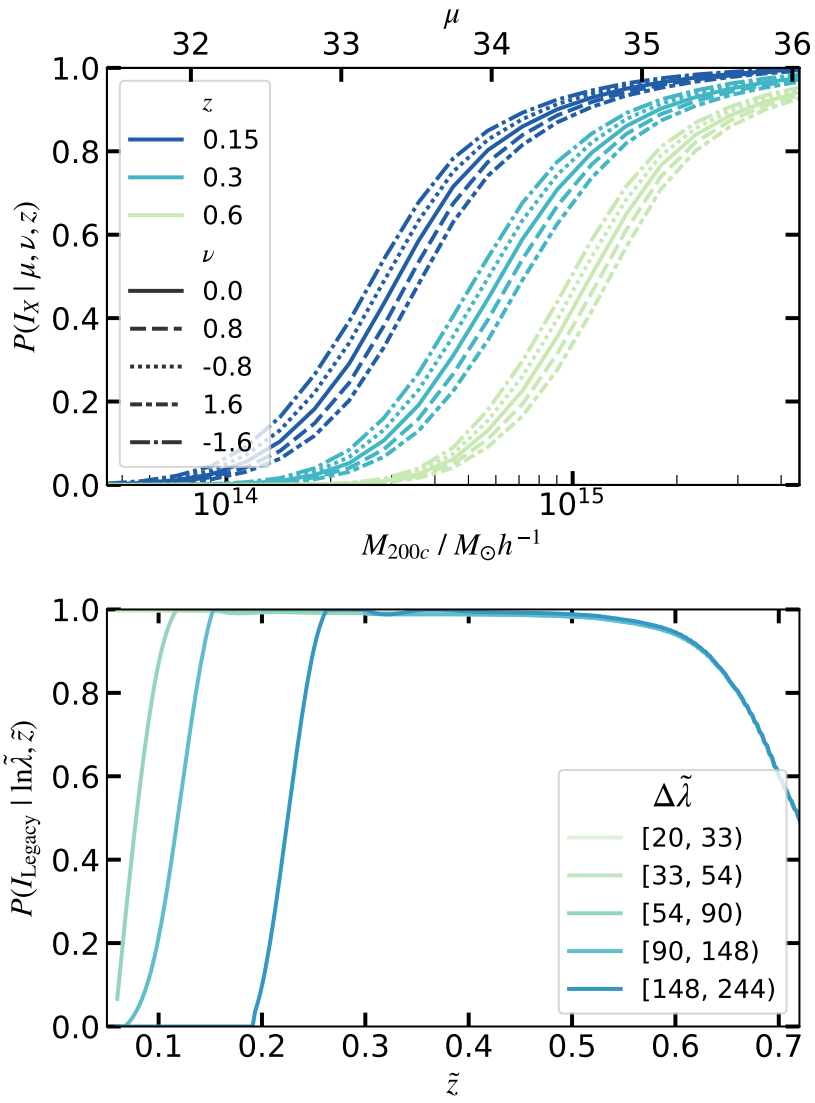

Abbildung 4.5: Taken from Ider Chitham et al. (2020). **Upper panel:** CODEX X-ray selection function, $P(I_{\rm X} \,|\, \mu, z, \nu)$ (defined by Eq. 9 of Finoguenov et al., 2020, for the SPIDERS DR16 footprint over The Legacy Surveys). Solid lines show the probability of detection as a function of mass ($\mu = \ln M_{\rm 200c}$) and true redshift (colour coded) when the expected richness of the richness-mass relation is equal to the true richness (i.e. $\nu = 0$, see Eq. 4.11). The dotted and dashed variants of the solid lines illustrate how the X-ray selection function changes when the difference between the expected richness and true richness is equal to -1.6, -0.8, 0.8 or 1.6 times the intrinsic scatter of the richness-mass relation. **Lower panel:** $P(I_{\rm Legacy} \,|\, \ln\tilde{\lambda}, \tilde{z})$, the optical selection function describing the completeness of the volume-limited redMaPPer catalogue of CODEX clusters with The Legacy Surveys over the SPIDERS DR16 footprint (Sect. 4.1.5). This considers the magnitude/luminosity cut ($L* < 0.2$), minimum richness ($\tilde{\lambda} > 5$) and maximum masking fraction (0.2).

sample is locally volume limited (Rykoff et al., 2016; Abbott et al., 2020; Bleem et al., 2020, discussed in Appendix 3.4.2) in order to reduce the Eddington (1913) bias. This simply means that the depth of the survey is used to estimate the maximum redshift at which galaxies at a z-band luminosity cut of $L* < 0.2$ can be observed in The Legacy Surveys (at the 10σ confidence level) as a function of position on the sky as described in Sects. 3.4.2 , 3.4.2 & 3.5.1. Clusters which exceed this maximum redshift and the maximum masking fraction (20% in this case) are excluded from the volume-limited catalogue. The $P(I_{\rm Legacy} \mid \ln\tilde{\lambda}, \tilde{z})$ distribution, illustrated in the lower panel of Fig. 4.5, is generated by redMaPPer for each redshift and richness bin using weighted randoms points for more detail, please refer to Sect. 3.4.2 of this thesis.

Redshift dependent richness cut

The second optical selection, $P(I_{\rm cut} \mid \ln\tilde{\lambda}, \tilde{z})$, describes a redshift dependent cut in observed-richness (Eq. 4.19) introduced to minimise contamination from spurious X-ray sources while retaining as many true systems as possible (also known as optical cleaning, e.g. Klein et al., 2018, 2019; Grandis et al., 2020). This is shown by the solid black boundary which slices through the coloured bins in Fig. 4.1.

$$P(I_{\rm cut} \mid \ln\tilde{\lambda}, \tilde{z}) = \theta(25\tilde{\lambda}_{\rm SDSS}(\tilde{z}/0.15)^{0.8}) \tag{4.19}$$

Here, θ is the Heaviside step function. The argument[6] corresponds to the threshold above which the X-ray detection of CODEX cluster exceeds a probability of $> 50\%$ ($P(I_{\rm X} \mid \mu, z, \nu) > 0.5$) under the assumption of the fiducial richness-mass relation (Capasso et al., 2019c). Such a conservative cut is motivated by the impact of the systematic uncertainties of the fixed X-ray luminosity-mass relation on the cosmological and richness-mass scaling relation parameters constrained in this work (discussed in Sect. 4.1.12). Although applying such a cut significantly reduces the total number of systems, and therefore the precision of the constrained parameters, the remaining sub-sample is extremely high-purity with a contamination level of $< 5\%$[7].

Total selection function

The probability of observing a cluster with $\tilde{\lambda}$ (given mass and redshift) is obtained by integrating the product of the overall selection, $P(I_{\rm tot} \mid \tilde{z}, \ln\tilde{\lambda}, \mu, z, \nu)$, with a log-normal distribution of observed richness with a mean with given by the true richness, $P(\ln\tilde{\lambda} \mid \ln\lambda, \mu, z)$.

[6]Please note that Eq. 4.19 is in terms of the original SDSS-based CODEX richness ($\tilde{\lambda}_{\rm SDSS}$) and is necessarily converted to The Legacy Surveys definition of richness using Eq. 4.1.

[7]Using a redshift dependent richness cut which corresponds to the 10% sensitivity level of the CODEX survey is estimated to result in approximately 5% of the selected X-ray sources being non-clusters (Finoguenov et al., 2020), therefore an equivalent cut at the 50% sensitivity level is expected to reduce this significantly.

The resulting term (Eq. 4.20) essentially provides the link between the richness-mass relation and the halo mass function (see definition in Sect. 4.1.4):

$$P(\ln\tilde{\lambda}\,|\,\mu,z) = \int_{\langle\ln\lambda(\mu,z)\rangle-4\sigma_{\rm int}}^{\langle\ln\lambda(\mu,z)\rangle+4\sigma_{\rm int}} \mathrm{d}\ln\lambda\; P(\ln\tilde{\lambda}\,|\,\ln\lambda,\mu,z)\cdot\; P(I_{\rm tot}\,|\,\tilde{z},\ln\tilde{\lambda},\mu,z,\nu)\,. \tag{4.20}$$

Here, $P(\ln\tilde{\lambda}\,|\,\ln\lambda,\mu,z) = \mathcal{N}(\ln\tilde{\lambda},\ln\lambda,\sqrt{\lambda})$, i.e. a log-normal distribution of observed richness, centred on the true value of richness with scatter described by the noise term in Eq. 4.3. The limits of the integral are determined by the minimum and maximum deviation from the fiducial richness-relation Eqs. 4.2 and 4.3 considered during the construction of the X-ray selection function. This can be simplified as an expression in terms of ν (introduced in Eq. 4.11) by changing limits and introducing an additional term $P(\nu) = \mathcal{N}(\nu,0,1)$. Expanding the total selection, $P(I_{\rm tot}\,|\,\tilde{z},\ln\tilde{\lambda},\mu,z,\nu) = P(I_X, I_{\rm Legacy}, I_{\rm cut}\,|\,\tilde{z},\ln\tilde{\lambda},\mu,z,\nu)$, into its separate components then gives

$$\begin{aligned} P(\ln\tilde{\lambda}\,|\,\mu,z) = \int_{-4}^{+4} \mathrm{d}\nu\; P(\nu)\cdot\; & P(\ln\tilde{\lambda}\,|\,\ln\lambda(\mu,z,\nu),z) \\ & \cdot\; P(I_{\rm X}\,|\,\mu,z,\nu) \\ & \cdot\; P(I_{\rm Legacy}\,|\,\ln\tilde{\lambda},\tilde{z}) \\ & \cdot\; P(I_{\rm cut}\,|\,\ln\tilde{\lambda},\tilde{z})\,. \end{aligned} \tag{4.21}$$

4.1.6 Parameter fitting & likelihood function

Cosmological and scaling relation parameters, contained in the vector θ, span a relatively high number of dimensions, therefore in order to heuristically fit a model to the binned data it is necessary to utilise a Monte Carlo Markov Chain (MCMC) sampler. The chosen sampler used in this work is `emcee` (Foreman-Mackey et al., 2013). The natural logarithm of the generic likelihood function has the following form

$$\ln\mathcal{L}(\boldsymbol{N}\,|\,\theta) \propto -\frac{1}{2}\left[(\boldsymbol{N}-\langle\boldsymbol{N}(\theta)\rangle)^T\,\boldsymbol{C}^{-1}\,(\boldsymbol{N}-\langle\boldsymbol{N}(\theta)\rangle)\right], \tag{4.22}$$

where $\boldsymbol{N}$ and $\langle\boldsymbol{N}(\theta)\rangle$ represent matrices of cluster counts in the *data* and expectation values provided by the *model*, respectively. Each element of the matrix represents a bin of $\Delta\ln\tilde{\lambda}_i$ and $\Delta\tilde{z}_j$, and $\boldsymbol{C}$ is the covariance matrix. The Poisson contribution to the covariance matrix is simply given by a diagonal matrix of the expectation value of counts in each bin $\boldsymbol{C}^{\rm Poisson} = \delta_{\rm ij}\langle\boldsymbol{N}\rangle_i$, where $\delta_{\rm ij}$ represents the Kronecker delta function. The contribution due to sample variance (Hu & Kravtsov, 2003) is ignored as it is subdominant in the regime of high mass (richness), therefore a Poisson likelihood is a good approximation given that the majority of low-richness systems are removed by the conservative redshift dependent richness cut Eq. 4.19.

4.1.7 Binning strategy

In order to determine the bin-size for the fiducial SPIDERS DR16 cosmological analysis, the experiment has been repeated over a grid of linear redshift bins $|\Delta\tilde{z}_j| \in \{0.1, 0.05, 0.025, 0.0125\}$ and logarithmic richness bins $|\Delta\ln\tilde{\lambda}_i| \in \{0.625, 0.5, 0.3125\}$ for the cosmological cluster sub-sample (Sect. 4.1.1, as well as the validation dataset detailed in Appendix 4.1.8). The aim being to find the binning scheme that minimises uncertainties from Poisson noise that otherwise limit the minimum width of the posterior distributions of both cosmological and scaling relation parameters (via Eq. 4.22). The precision for each run is given by the volume of the parameter space spanned by the Markov chains at the point of convergence. Using the sum of the squares of the fractional uncertainties on all the constrained parameters as an approximate performance indicator, the optimal (fiducial) configuration was to be $|\Delta\tilde{z}_j| = 0.05$, $|\Delta\ln\tilde{\lambda}_i| = 0.5$, corresponding to ten bins of observed redshift and 4 bins of observed richness.

4.1.8 Validation with Simulations

To conduct a robust test of the cosmological pipeline, a validation procedure is performed using a mock catalogue of clusters (Comparat et al., 2020a). The basis of the mock is a suite of MultiDark simulations (Klypin et al., 2016) generated under a *Planck* cosmology ($H_0 = 67.77\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $\Omega_{m_0} = 0.307115$, $\Omega_{b_0} = 0.048206$, $\sigma_8 = 0.8228$, $n_s = 0.96$, $w_0 = -1$, $N_{\mathrm{eff}} = 3.046$). The simulation (box volume of $1.0\,\mathrm{Gpc}^3$) is replicated to cover the full sky down to halo masses of $M_{500c} = 5 \times 10^{13} M_\odot$ for $\tilde{z} \in [0.1, 0.6)$.

Values of observed richness are generated by assuming a log-normal distribution of true richness with a mean equal to the expectation value from the SPTPol Extended Cluster Survey richness-mass relation (Bleem et al., 2020, derived at a *Planck*-like fiducial cosmology) with a standard deviation determined by Eq. 4.3. For simplicity, the relation is left in its original form of, i.e. as function of M_{500c}, with a redshift evolution which scales with the Hubble parameter, $E(z)/E(z_{\mathrm{pivot}})$, rather than the $(1+z)/(1+z_{\mathrm{pivot}})$ scaling assumed in Eq. 4.2. The SPIDERS DR16 polygon mask (Swanson et al., 2008) is then applied to ensure the coverage of the simulated catalogue matches the survey footprint. The number of clusters in the catalogue is reduced further using weighted down-sampling. Weights are determined for each halo using the probability provided by the CODEX X-ray selection function $P(I_{\mathrm{X}} \,|\, \mu, z, \nu)$ (Finoguenov et al., 2020, assuming $\Omega_{m_0} = 0.31$), which depends only on the true halo properties of the simulated clusters, and the richness-mass relation to transform the generated richness into ν space via Eq. 4.11).

The redshift-mask (derived from a redMaPPer depth map of the limiting galaxy magnitude and models of m_*) is then applied to produce a volume limited catalogue with $L_* < 0.2$ and $\tilde{\lambda} > 5$. The selection term to account for this, $P(I_{\mathrm{Legacy}} \,|\, \ln\tilde{\lambda}, \tilde{z})$, is estimated using random weighted points a for each bin of observed redshift and richness in the simulated catalogue as detailed in Sect. 4.1.5. Finally, a redshift dependent cut in richnes, $P(I_{\mathrm{cut}} \,|\, \ln\tilde{\lambda}, \tilde{z})$[8] is applied to the simulated cluster sample to create a synthetic catalogue

[8]Please note this is the same redshift dependent redshift cut as for the analysis using the observed

Tabelle 4.1: Summary of measured parameters and their initial priors for the simulated SPIDERS DR16 dataset as shown in Fig. 4.9. The format of this table is identical to that of the observed SPIDERS DR16 data Table. 4.3. Fiducial values of the input Bleem et al. (2020) richness-mass relation are omitted as they are only defined for M_{500c} rather than M_{200c}. For a comparison of the measured richness-mass relation relative to the input one, please refer to Fig. 4.8.

Parameter	Fiducial	Prior	Posterior
Ω_{m_0}	0.31	$\in [0.1, 0.8]$	$0.32^{+0.11}_{-0.06}$
σ_8	0.82	$\in [0.4, 1.2]$	$0.84^{+0.08}_{-0.06}$
A	-	$\in [3.0, 5.0]$	$3.93^{+0.19}_{-0.35}$
B_μ	-	$\in [0.0, 10.0]$	$1.10^{+0.25}_{-0.21}$
B_z	-	$\in [-5.0, 5.0]$	$0.08^{+0.87}_{-0.81}$
$\sigma_{\rm int}$	0.22	$\sim \mathcal{N}(0.23, 0.16)$	$0.30^{+0.16}_{-0.15}$

Ω_{m_0} represents the mean matter density of the universe at redshift zero.
σ_8 is the amplitude of the matter power spectrum.
A is $\langle \ln \lambda(\mu, z) \rangle$ at pivot mass scale and pivot redshift.
B_μ is the coefficient of halo mass dependence in $\langle \ln \lambda(\mu, z) \rangle$.
B_z is the coefficient of linear redshift dependence in $\langle \ln \lambda(\mu, z) \rangle$.
$\sigma_{\rm int}$ is the intrinsic scatter of the richness-mass relation.

that resembles the high-purity SPIDERS DR16 sub-sample (Sect. 4.1.1 & Eq. 4.19).

The SPIDERS DR16 likelihood function (Eq. 4.22) is applied to the synthetic sub-sample to attempt recover the input cosmological and scaling relation parameters. The cosmological parameters are recoverable within one standard deviation of their input values using a Tinker et al. 2008 HMF, see Figs. 4.7 and 4.9. In order to establish how accurately the input relation is recovered, the M_{500c} definition of the input relation is converted to a function of M_{200c} using a Child et al. (2018) halo concentration model and NFW profile (Navarro et al., 1997). The comparison between the measured richness-mass relation and the input relation at the pivot (median) mass and redshift of the simulated dataset is shown in Fig. 4.8 using consistent definitions of mass. The recovered normalisation, mass and redshift slopes are in good agreement with the input richness-mass relation.

4.1.9 Results

The best fit model (Sect. 4.1.4) to the abundance data (Sect. 4.1.1) is summarised visually in Fig. 4.2 and numerically in Table. 4.2. The corresponding posterior distributions of measured parameters are shown in Fig. 4.10 along with contours indicating the 68 and 95% confidence levels. The fiducial, prior and posterior values of these parameters are also

CODEX cluster sample (Eq. 4.19), however, it is kept in terms of the DES definition of richness.

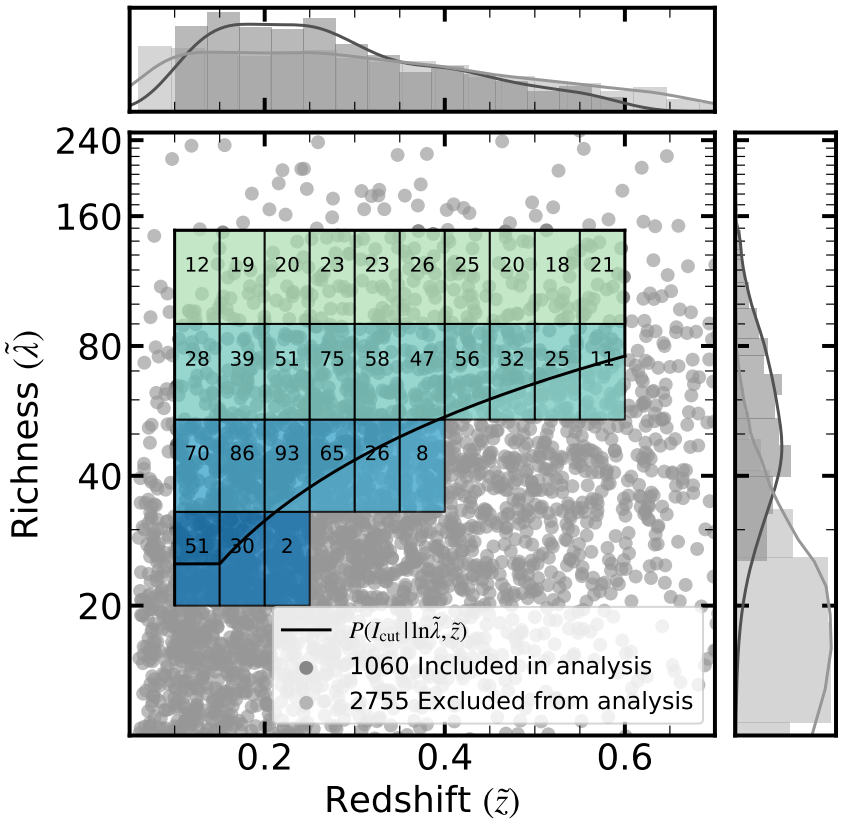

Abbildung 4.6: Taken from Ider Chitham et al. (2020). The simulated validation sample of SPIDERS DR16 clusters (3815 in total) in the richness-redshift plane. This is *after* applying the SPIDERS DR16 survey mask, volume-limiting procedure ($L_* < 0.2$, maximum mask fraction > 0.2 and $\lambda > 5$) and down-sampling using the CODEX X-ray selection function, $P(I_{\rm X} \,|\, \mu, z, \nu)$. The solid black line represents a redshift-dependent richness cut described by $P(I_{\rm cut} \,|\, \ln\tilde{\lambda}, \tilde{z})$ in Eq. 4.19. Boxes represent the bins used in the validation experiment annotated with the respective clusters counts (after applying $P(I_{\rm cut} \,|\, \ln\tilde{\lambda}, \tilde{z})$) and colours represent the richness bins highlighted in Fig. 4.7. The dark grey distributions illustrate clusters that are included in the analysis i.e. clusters that fall within the coloured boxes *and* that are above the black line. The light grey distributions illustrate clusters excluded from the analysis.

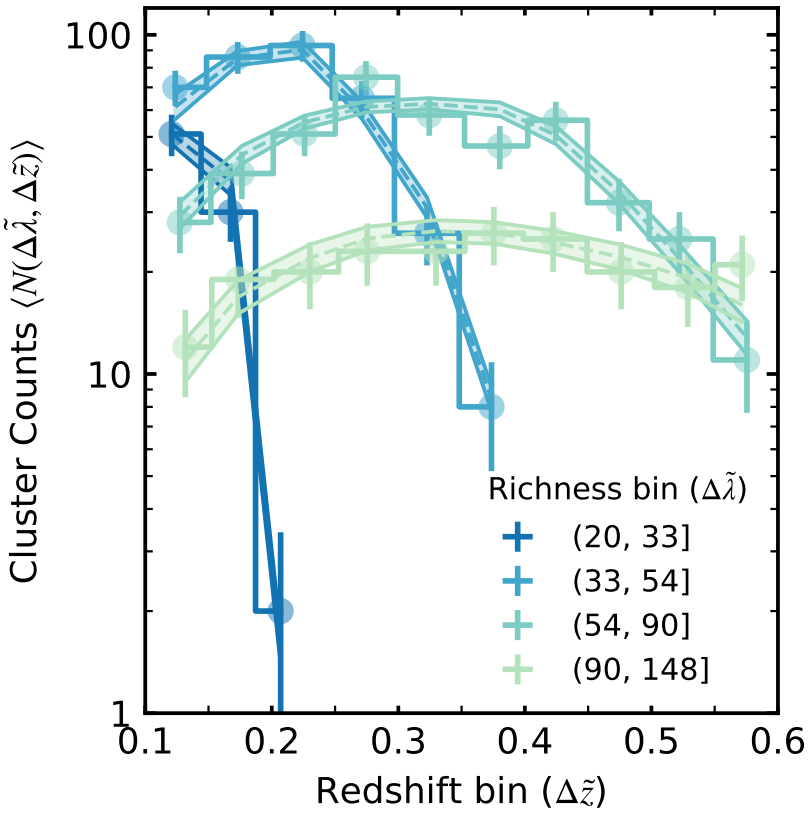

Abbildung 4.7: Taken from Ider Chitham et al. (2020). The abundance of simulated SPIDERS clusters as a function in bins of observed redshift ($\Delta\tilde{z}_j$) and richness ($\Delta\ln\tilde{\lambda}_i$) where $\tilde{z} \in [0.1, 0.6)$ and $\ln\tilde{\lambda} \in [25, 148)$. Steps represent the simulated data (Fig. 4.6), the width and height correspond to the size of the bin and the magnitude of the diagonal (Poisson) elements of the covariance matrix, respectively. The position of the vertical bars indicates the mean redshift in each bin. The shaded regions trace the expectation value provided by the model (with a normal prior on the intrinsic scatter Bleem et al., 2020), centred on the median, which corresponds to the best-fit cosmology. The lower and upper limit are similarly set by the 15% and 85% confidence intervals. These distributions are calculated directly from the stored expectation values of cluster counts for the MCMC chains used to create the contours shown in Fig. 4.9

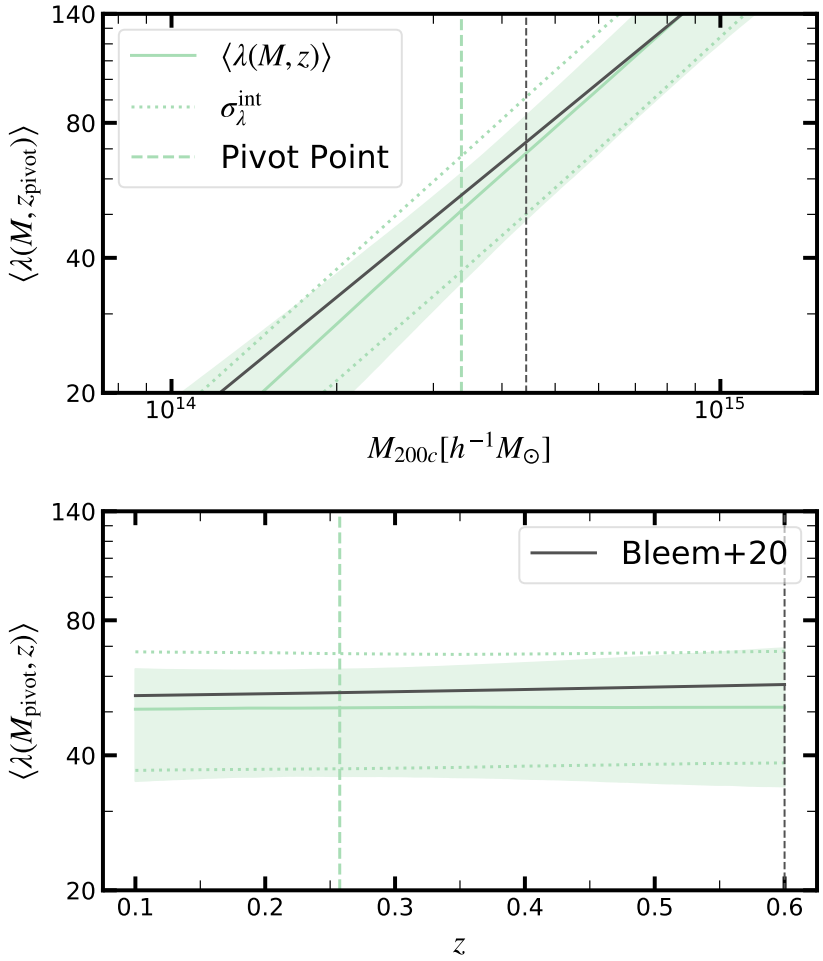

Abbildung 4.8: Taken from Ider Chitham et al. (2020). Constraints for the richness-mass relation evaluated at the pivot mass $M_{200c} = 3.37 \times 10^{14} M_\odot h^{-1}$ (lower panel) and pivot redshift 0.257 (upper panel) used in this analysis. The dotted line about the solid line indicates the intrinsic scatter (with normal prior from Bleem et al., 2020) about the mean (Eq. 4.2), while the coloured contours show the 16 and 85 percentiles of the draws from the MCMC chains used to constrain the cosmological and scaling relation parameters shown in Fig. 4.9

.

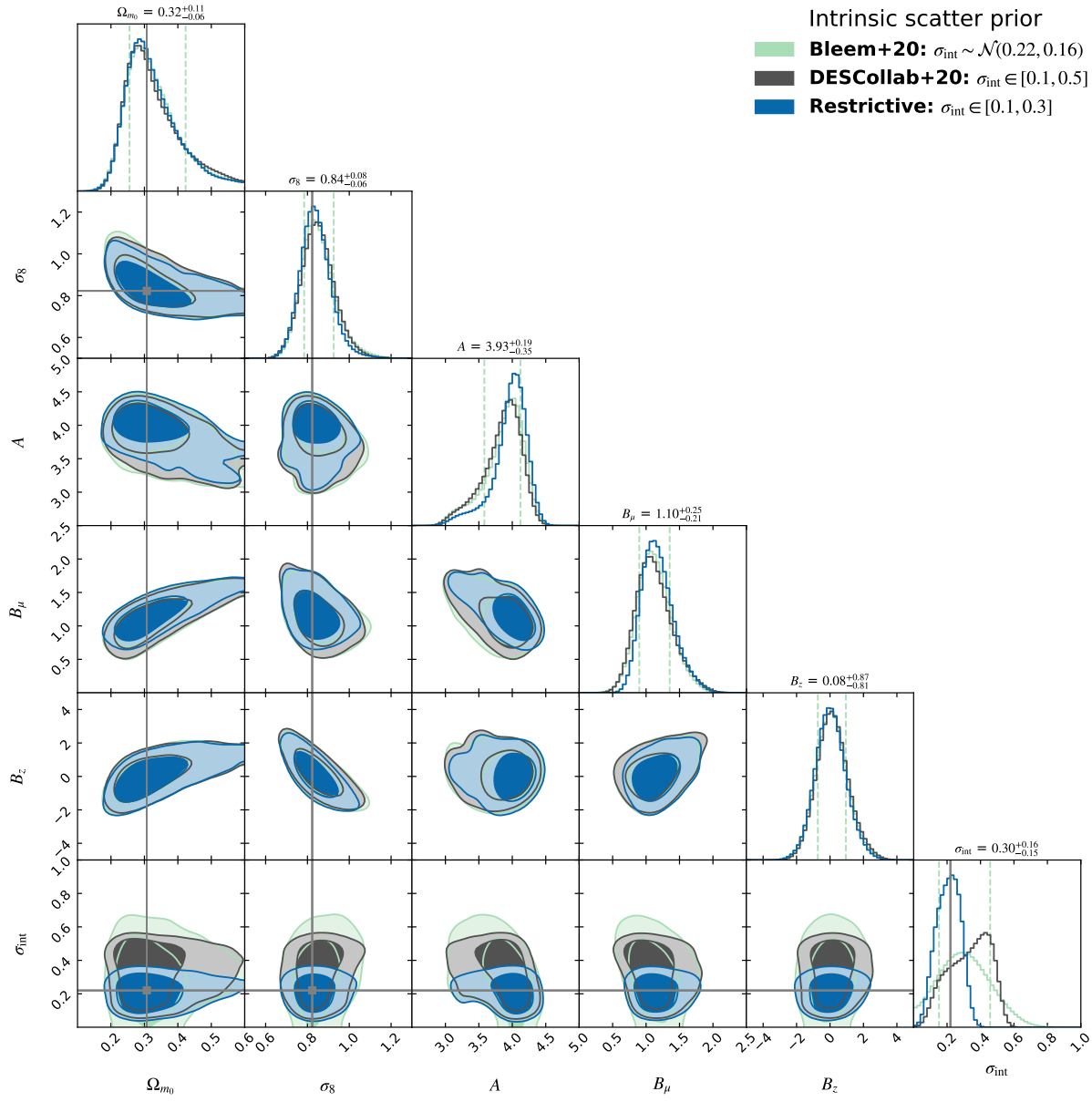

Abbildung 4.9: Taken from Ider Chitham et al. (2020). Cosmological parameters and scaling relation parameter constraints for the simulated SPIDERS DR16 dataset as summarised in Table. 4.1. Contours depict the one 68% and 95% confidence levels where posterior distributions are obtained using the full SPIDERS DR16 likelihood function outlined in Sect. 4.1.6. Input values of cosmological and intrinsic scatter parameters are marked by the grey lines, for a comparison of the input and recovered scaling relation, please refer to Fig. 4.8. The impact of the intrinsic scatter prior is shown by the different coloured posteriors (see legend). Although $\sigma_{\rm int}$ is only well constrained when using the Bleem et al. (2020) prior, the determination of every other parameter is found to be relatively insensitive to the alternate priors.

Tabelle 4.2: A tabular representation of Fig. 4.2, summarising cluster counts in bins of redshift and optical richness for the 691 SPIDERS DR16 galaxy clusters that comprise the cosmological cluster sub-sample. Parentheses contain values predicted by the best fit model (summarised in Table. 4.3) and statistical uncertainties taken from the diagonal (Poisson) elements of the covariance matrix.

$\Delta\tilde{z}/\Delta\tilde{\lambda}$	(25, 33]	(33, 54]	(54, 90]	(90, 148]
(0.1, 0.15]	$41(35 \pm 5)$	$37(37 \pm 6)$	$11(13 \pm 3)$	$2(3 \pm 1)$
(0.15, 0.2]	$46(37 \pm 6)$	$58(52 \pm 7)$	$24(22 \pm 4)$	$2(6 \pm 2)$
(0.2, 0.25]	$15(11 \pm 3)$	$42(58 \pm 7)$	$32(28 \pm 5)$	$6(9 \pm 3)$
(0.25, 0.3]	–	$45(52 \pm 7)$	$41(32 \pm 5)$	$17(11 \pm 3)$
(0.3, 0.35]	–	$17(35 \pm 5)$	$24(32 \pm 5)$	$14(12 \pm 3)$
(0.35, 0.4]	–	$18(22 \pm 4)$	$28(30 \pm 5)$	$16(12 \pm 3)$
(0.4, 0.45]	–	$14(14 \pm 3)$	$26(26 \pm 5)$	$7(11 \pm 3)$
(0.45, 0.5]	–	$9(9 \pm 3)$	$26(22 \pm 4)$	$10(10 \pm 3)$
(0.5, 0.55]	–	$7(6 \pm 2)$	$22(18 \pm 4)$	$6(9 \pm 3)$
(0.55, 0.6]	–	$6(5 \pm 2)$	$15(13 \pm 3)$	$7(7 \pm 2)$

Tabelle 4.3: Summary of measured parameters and their initial priors. The fiducial column describes the starting value of the Markov chains in the parameter space. In this column, cosmological parameters are those used to pre-compute the X-ray selection function described in Sect. 4.1.5 and scaling relation parameters are initialised to their fiducial values (Capasso et al., 2019c). Posterior distributions correspond to the twenty-fourth percentile about the median as summarised by the diagonal elements of Fig. 4.10.

Parameter	Fiducial	Prior	Posterior
Ω_{m_0}	0.30	$\in 0.1, 0.8]$	$0.34^{+0.09}_{-0.05}$
σ_8	0.780	$\in [0.4, 1.2]$	$0.73^{+0.03}_{-0.03}$
A	~ 3.68	$\in [3.0, 5.0]$	$3.96^{+0.15}_{-0.30}$
B_μ	0.98	$\in [0.0, 10.0]$	$1.07^{+0.22}_{-0.20}$
B_z	-1.08	$\in [-5.0, 5.0]$	$1.42^{+0.72}_{-0.69}$
$\sigma_{\rm int}$	0.22	$\sim \mathcal{N}(0.23, 0.16)$	$0.29^{+0.18}_{-0.19}$

Ω_{m_0} represents the mean matter density of the universe at redshift zero.
σ_8 is the amplitude of the matter power spectrum.
A is $\langle \ln \lambda(\mu, z) \rangle$ at pivot mass scale and pivot redshift.
B_μ is the coefficient of halo mass dependence in $\langle \ln \lambda(\mu, z) \rangle$.
B_z is the coefficient of linear redshift dependence in $\langle \ln \lambda(\mu, z) \rangle$.
$\sigma_{\rm int}$ is the intrinsic scatter of the richness-mass relation.

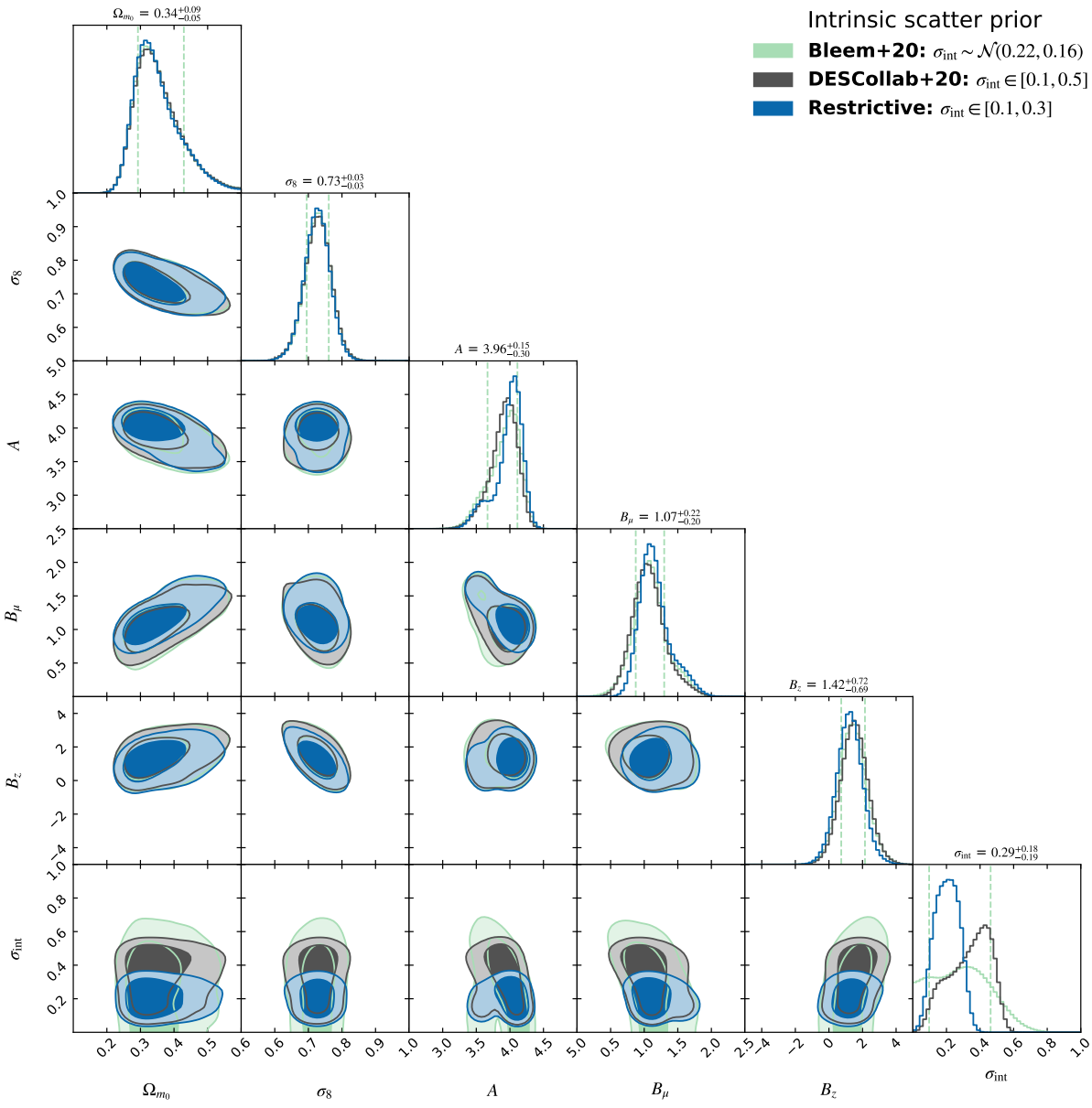

Abbildung 4.10: Taken from Ider Chitham et al. (2020). Cosmological parameters and scaling relation parameter constraints for SPIDERS DR16 as summarised in Table. 4.3. Contours depict the 68% and 95% confidence levels where posterior distributions are obtained using the likelihood function outlined in Sect. 4.1.6. The impact of the intrinsic scatter prior is shown by the different coloured posteriors (see legend).

summarised in Table. 4.3, with best fit values found to be $\Omega_{m_0} = 0.34^{+0.09}_{-0.05}$, $\sigma_8 = 0.73^{+0.03}_{-0.03}$, $A = 3.96^{+0.15}_{-0.30}$, $B = 1.07^{+0.22}_{-0.20}$, $B_z = 1.42^{+0.72}_{-0.69}$ and $\sigma_{\rm int} = 0.29^{+0.18}_{-0.19}$.

The pipeline used to constrain these parameters is validated in Appendix 4.1.8 using a synthetic volume-limited catalogue of CODEX clusters over the SPIDERS DR16 area (Comparat et al., 2020a) after generating richness via the Bleem et al. (2020) relation over the same richness and redshift range as the observed cosmological cluster sub-sample. The same likelihood model described in Sect. 4.1.2 is implemented on the simulated data and all input cosmological and richness-mass relation parameters are comfortably recovered within one standard deviation as shown by Figs. 4.8 and 4.9 and displayed numerically in Table. 4.1.

The following subsections attempt to interpret the parameters constrained from the observed cluster catalogue. This includes a comparison of the richness-mass relation to the most recent calibration experiments and the measured cosmological parameters to constraints obtained by other recent cluster cosmology experiments that cover a similar volume.

Richness-mass relation constraints

The richness-mass relation corresponding to the best fit parameters at the pivot mass, $3.08 \times 10^{14} {\rm M}_\odot h^{-1}$ and redshift 0.263 is shown in Fig. 4.11 alongside several recently calibrated richness-relations from the literature:

- Capasso et al. (2019c): derived from 428 SPIDERS clusters using dynamical mass calibration and the original SDSS based measurements of richness for CODEX clusters. The redshift range, $\tilde{z} \in [0.1, 0.7)$, is similar to the one presented in this paper although the richness range is much larger with $\tilde{\lambda}_{\rm SDSS} > 20$. Clusters are selected via the minimum number of spectroscopic members per cluster $N_{\rm mem} > 10$. The pivot mass and redshift are $M_{\rm 200c} = 3 \times 10^{14} {\rm M}_\odot$ and 0.18 respectively. The version of relation shown in Fig. 4.11 also accounts for correlated scatter between richness and X-ray luminosity using the CODEX selection function (Finoguenov et al., 2020). For comparative purposes, observed richness is converted to The Legacy Surveys definition using Eq. 4.1.

- Bleem et al. (2020): calibrated using the Joint SPT-redMaPPer Cluster Sample from the 2,770 deg^2 SPTPol Extended Cluster Survey and DES redMaPPer catalogue. This relation has a pivot redshift of 0.6 and pivot mass of $M_{\rm 500c} = 3 \times 10^{14} {\rm M}_\odot h^{-1}$. Masses are converted from $M_{\rm 500c}$ to $M_{\rm 200c}$ using a Child et al. (2018) halo concentration model and a NFW profile (Navarro et al., 1997). The conversion between DES and SDSS richness is carried out using $\tilde{\lambda}_{\rm SDSS} \approx 0.93 \tilde{\lambda}_{\rm DES}$ (McClintock et al., 2019a) in combination with Eq. 4.1. As DES photometric data are complete, no redshift scaling is applied to the richness value when applying the conversion between SDSS and The Legacy Surveys definitions. This leads to $\tilde{\lambda} \approx 0.87 \tilde{\lambda}_{\rm DES}$.

At the pivot redshift (0.263), the normalisation and mass evolution of the best fit relation is remarkably similar to that of Capasso et al. (2019c) and Bleem et al. (2020)

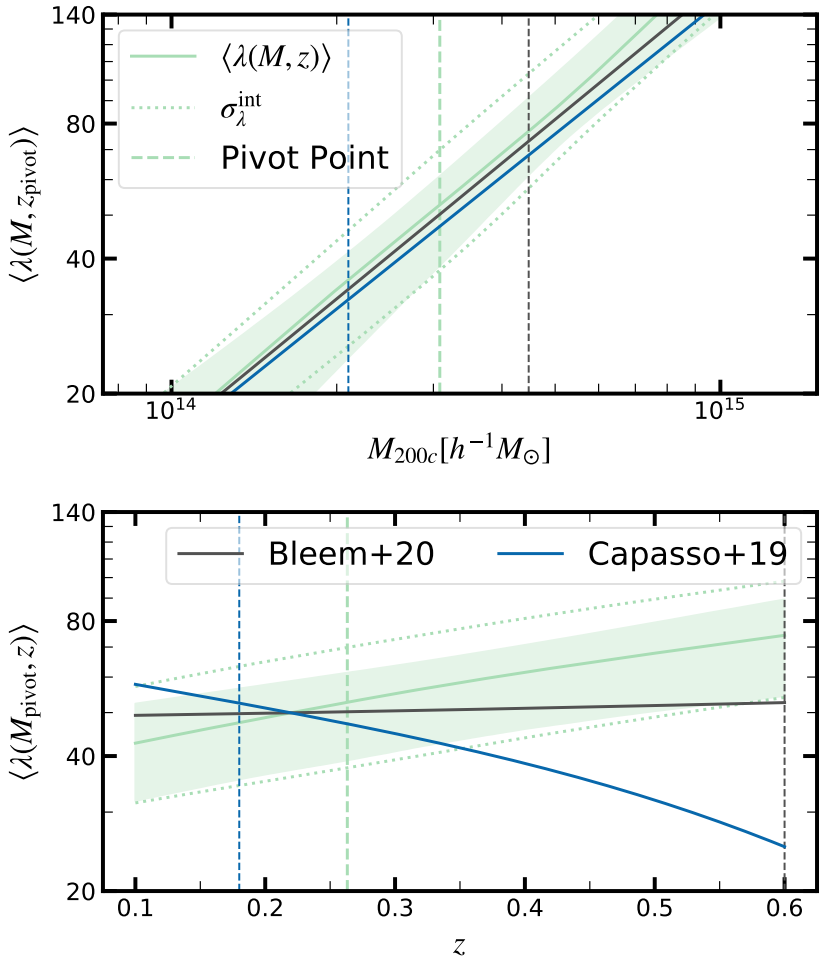

Abbildung 4.11: Taken from Ider Chitham et al. (2020). Constraints for the richness-mass relation evaluated at the pivot redshift and mass used in this analysis (0.263 and $3.08 \times 10^{14} \mathrm{M}_\odot h^{-1}$). The dotted line about the solid line indicates the intrinsic scatter ($\sigma_{\rm int}$ with normal prior from Bleem et al., 2020) about the mean (Eq. 4.2), while the coloured contours show the 16 and 85 percentiles of the draws from the MCMC chains used to constrain the cosmological and scaling relation parameters shown in Fig. 4.10.

although at the pivot mass ($3.08 \times 10^{14} \mathrm{M}_{\odot} h^{-1}$) the redshift trend is vastly different. This work indicates a moderate, positive redshift evolution rather than predicting richness which decreases with redshift (Capasso et al., 2019c), or little to no redshift evolution (Bleem et al., 2020). It is likely that the difference in the selection methods used to construct each sample contributes to the discrepancy in the redshift trends, although it is currently not possible to draw any strong conclusions due to the lack of statistical significance on the B_z parameter constraints. The wide posterior distribution of B_z in Fig. 4.10 is indicative of $\sim 50\%$ measurement uncertainties with systematic uncertainties also estimated to be $\gtrsim 50\%$ (Sect. 4.1.12). The degeneracy between B_z and cosmological parameters suggests that its value is in fact consistent with the derived cosmology (positive redshift evolution corresponds to high-Ω_{m_0} and low-σ_8 relative to a *Planck*-like cosmology). Repeating the experiment with a fixed value of $B_z = 0$, confirms this by resulting in cosmological parameters that are more closely aligned with the median value of of recent cluster analyses shown in Fig. 4.12 (i.e. it increases the degree of similarity with respect to the cosmological constraints of Mantz et al., 2015; Zubeldia & Challinor, 2019; Bocquet et al., 2019). The corresponding constraints on the richness-mass relation when $B_z = 0$ are also consistent with that of Bleem et al. (2020) shown in Fig. 4.11.

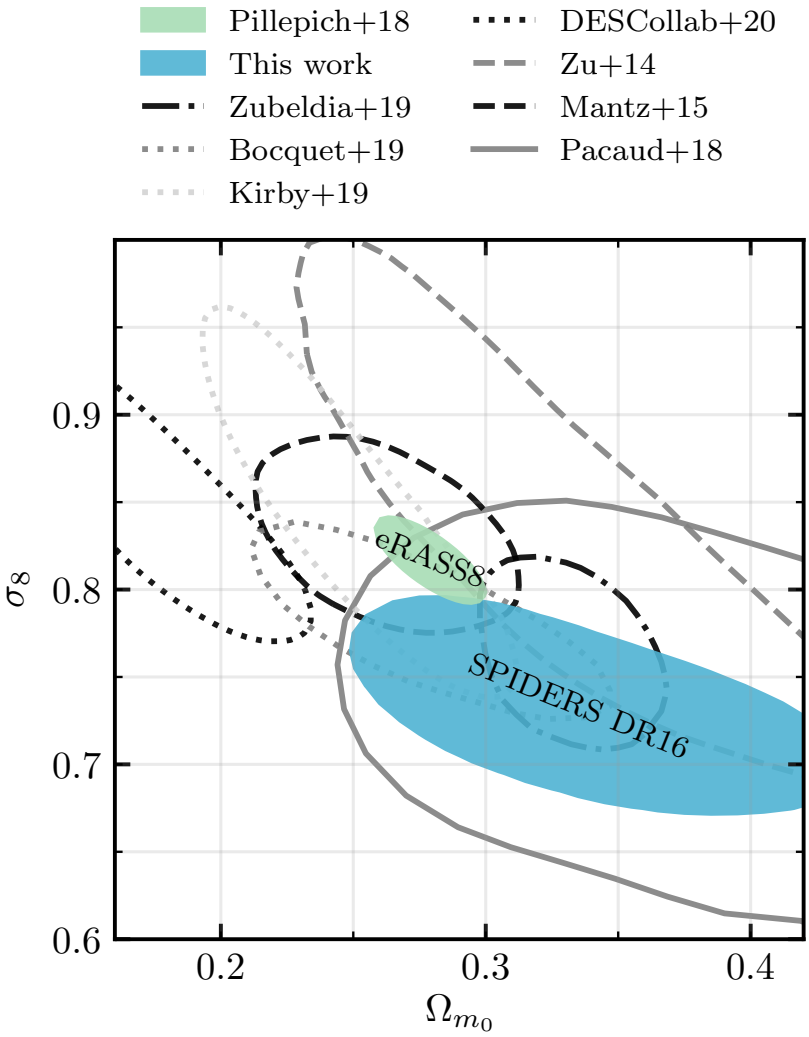

Abbildung 4.12: Taken from Ider Chitham et al. (2020). Cosmological parameter constraints obtained with the SPIDERS DR16 compared to previous work. Contours represent the 68% confidence level. Table. 4.4 contains a comparison of each reference shown in the figure above, summarising each selection method, sky area, redshift range, sample size, and data origin. The blue ellipse shows the constraints from this work and is identical to the inner contour shown in top-left panel of Fig. 4.10 when a normal prior is used for the intrinsic scatter of the richness-mass relation (Bleem et al., 2020). The green ellipse represents a conservative *forecast* of the constraining power of *eROSITA* after its four year all sky survey (Pillepich et al., 2018). Grey and black lines represent constraints from recent competitive cluster analyses, for more detailed descriptions of each experiment please refer to Sect. 4.1.10. Erratum: DESCollab+20 refers to the archival version of Abbott et al. (2020) and eRASS8 refers the final cumulative eRASS (eRASS:8).

Tabelle 4.4: A comparison of cosmological analyses as summarised in Fig. 4.12. Columns include references, the origin of the data source, the total number of clusters used in the fiducial analysis, the area covered in square degrees. The final column corresponds to the redshift range, i.e. the minimum and maximum cluster redshift used in the sample (analyses are not necessarily complete over this redshift range). The description in parentheses also specifies the type of redshifts used; spectroscopic (spec), photometric (photo) or a mixture (both).

Reference	Origin	Selection	# Clusters	Area [deg^2]	Redshift (Type)
This work	SPIDERS DR16	X-ray	691	5,350	0.10-0.60 (spec[9])
Pillepich et al. (2018)	eRASS:8 forecast (pessimistic)	X-ray	88,900	27,145	0.01-2.00 (photo)
Zubeldia & Challinor (2019)	*Planck* MMF3 + CMB lensing	SZ	433	26,814	0.00-1.00 (both)
Bocquet et al. (2019)	SPT + *Chandra* + lensing	SZ	377	2,500	0.29-1.13 (photo)
Abbott et al. (2020)	DES-redMaPPer + lensing	Optical	6,504	1,500	0.20-0.65 (photo)
Kirby et al. (2019)	SDSS-redMaPPer + lensing + *Chandra*	Optical	6,964	9,000	0.10-0.30 (photo)
Zu et al. (2014)	MaxBCG + lensing	Optical	10,815	7,398	0.10-0.30 (photo)
Mantz et al. (2015)	*Chandra* archive	X-ray	224	-	0.08-1.06 (spec)
Pacaud et al. (2018)	*XMM*-XXL	X-ray	178	50	0.05-1.00 (spec)

4.1.10 Cosmological constraints

Table. 4.4 contains a comparison of cluster count experiments from the literature, with Fig. 4.12 depicting the respective posterior distributions in the $\Omega_{m_0} - \sigma_8$ plane. Each experiment is unique in the sense they differ with the primary method of cluster selection and mass calibration, cluster sample size, survey volume and the type of redshifts used in the analysis (photometric, spectroscopic or a mixture of the two):

- Mantz et al. (2015): 224 bright clusters originating from three X-ray flux-limited samples of clusters (Ebeling et al., 1998; Böhringer et al., 2004; Ebeling et al., 2010) based on the RASS (Truemper, 1993). Mass proxies are provided from X-ray follow-up (94 clusters have high-quality *Chandra* data), weak gravitational lensing data (50 massive clusters) and measurements of the X-ray gas fraction.

- Pacaud et al. (2018): a sub-sample of 178 bright and spectroscopically confirmed serendipitous clusters from the *XMM*-XXL survey (Adami et al., 2018) with an X-ray selection function depending on the extent and extent-likelihood of the emission to ensure contamination levels are $\lesssim 50\%$ (Pacaud et al., 2006; Clerc et al., 2012; Pacaud et al., 2016). The redshift range is large, $\tilde{z} \in (0.05, 1.00)$, and masses are estimated using calibrated relations between mass-temperature (Lieu et al., 2016), luminosity-temperature (Adami et al., 2018) and the link between extent and core-radius (Pacaud et al., 2018).

- Pillepich et al. (2018): a forecast of cosmological constraints predicted from eight scans of *eROSITA* all sky survey (eRASS8). This version of the forecast is a conservative (pessimistic) estimate of the precision in the sense that only photometric redshifts are used.

- Zubeldia & Challinor (2019): 439 clusters from the MMF3 cosmology sample (Planck Collaboration et al., 2016d) with a thermal SZ signal-to-noise ratio greater than 6. CMB lensing mass estimates are provided for of all clusters with redshifts (Planck Collaboration et al., 2016b) which originate from a mixture a variety of ancillary data (Planck Collaboration et al., 2014, 2015a,b, 2016c,d). This includes a mixture of photometric and spectroscopic information from X-ray (Piffaretti et al., 2011), optical (Rykoff et al., 2014; Liu et al., 2015) and SZ (Hasselfield et al., 2013; Bleem et al., 2015) counterparts.

- Bocquet et al. (2019): 377 clusters from the South Pole Telescope 2500deg^2 survey (de Haan et al., 2016) with high detection significance ($\zeta > 5$) a purity of 95% (Bleem et al., 2015). The analysis uses a mixture of photometric (Bleem et al., 2015; Strazzullo et al., 2019) and spectroscopic (Bayliss et al., 2016; Khullar et al., 2019) redshifts over $\tilde{z} \in (0.29, 1.13)$ with complementary *Chandra* X-ray data (McDonald et al., 2013, 2017) for 89 clusters and weak-lensing data from the *Hubble Space Telescope* and Magellan (Schrabback et al., 2018; Dietrich et al., 2019) for 32 clusters.

The following optical selected cluster analyses all simultaneously model the weak lensing measurement, richness-mass relation and abundance of clusters using photometric redshifts:

- Zu et al. (2014): a sub-sample of 10,815 SDSS MaxBCG clusters (Koester et al., 2007a) which is almost volume-limited over the redshift range $\tilde{z} \in [0.1, 0.3)$ with $\sim 90\%$ completeness and purity for $M_{200m} >= 10^{14} h^{-1} \mathrm{M}_{\odot}$.
- Kirby et al. (2019): improve on the work of Costanzi et al. (2019b), which is based on 6,964 SDSS redMaPPer clusters, by deriving X-ray gas masses for the 30 richest systems using supplementary *Chandra* data. The cluster sample is volume limited with $\tilde{z} \in [0.1, 0.3]$.
- Abbott et al. (2020): a volume-limited sample of 6,504 DES redMaPPer clusters over $\tilde{z} \in [0.2, 0.65]$. Changes to the original Costanzi et al. (2019b) model were implemented after *un-blinding*, including an updated (tighter) prior on the intrinsic scatter and corrections to account for systematic biases in the weak lensing mass estimates. The contours shown in Fig. 4.10 correspond to the results of the *un-blinded* analysis.

The work presented here is differs from all of the above in that it is the only volume-limited X-ray selected cluster sample that uses a self calibrated richness-mass relation as well spectroscopic redshifts over a large area and redshift range without supplementary mass measurements. The X-ray selection function does, however, assume a pre-calibrated X-ray luminosity-mass relation, with the magnitude of the shift between the Mantz et al. (2015) constraints and the SPIDERS DR16 constraints expected to represent the order of systematic uncertainty caused by fixing this relation. As the accuracy of the mass-calibration process is usually the largest source of systematic uncertainty impacting cosmological analyses with clusters, the difference between constraints presented in this work and that of SZ selected samples (Zubeldia & Challinor, 2019; Bocquet et al., 2020) is expected to be driven primarily by the ability to accurately constrain observable-mass relation parameters, especially at high-mass and high-redshift. The deviation from the constraints of Kirby et al. (2019) and Abbott et al. (2020) is partially explained by a more sophisticated physically motivated model of richness related biases compared to this work (Sect. 4.1.12), but mostly due to the systematic effects that impact the weak-lensing mass measurements of optically selected redMaPPer clusters. Although in general, the constraints derived from the SPIDERS DR16 cosmological cluster sub-sample are in good agreement with previously published measurements, the systematic uncertainties caused the X-ray selection function alone are estimated to be least $\gtrsim 20\%$ (Sect. 4.1.12), with an additional unknown contribution from richness-related biases (Sect. 4.1.12) which have yet to be quantified for CODEX clusters.

4.1.11 Impact of intrinsic scatter

When the intrinsic scatter (σ_{int}) of the richness-mass relation is assumed to be independent of mass and redshift (Eq. 4.3), it has shown to be degenerate with the normalisation (A)

of the relation (e.g Bleem et al., 2020) as well as with cosmological parameters (Costanzi et al., 2019b). In the case of richness-mass relations derived from a sub-sample of CODEX clusters; $\sigma_{\rm int}$ and A are consistently found to be almost completely degenerate (e.g. Capasso et al., 2019c, Kiiveri et al. subm.). This strong degeneracy between $\sigma_{\rm int}$ and A is also apparent in this work (Fig. 4.10), with $\sigma_{\rm int}$ being a prior dominated quantity. When a restrictive prior is used $\sigma_{\rm int} \in [0.1, 0.3)$, the degeneracy is less obvious, but using the wider uniform prior used by Abbott et al. (2020) $\sigma_{\rm int} \in [0.1, 0.5)$, the degeneracy is much more significant. This is also the case when using normal priors centred on the best fit values from recently calibrated richness-mass relations (Saro et al., 2015; Capasso et al., 2019c; Mulroy et al., 2019; Bleem et al., 2020, and Kiiveri et al. subm.) e.g. $\sigma_{\rm int} \approx \mathcal{N}(0.2, 0.15)$. This choice of prior also effects a mild bi-modality in the posterior distribution of the intrinsic scatter. This is likely related to the wide prior used for the normalisation, the poorly constrained redshift evolution parameter and large residuals between the data and model at the most extreme values of richness and redshift shown in Fig. 4.2. As the intrinsic scatter is poorly constrained, this in turn, directly impacts the posterior distribution of the normalisation and mass evolution, although cosmological parameters and the redshift trend of the richness-mass relation are relatively unaffected. This effect, although less prominent, is also observed in the posterior distributions of simulated SPIDERS DR16 catalogue (Fig. 4.9). Although informative, using a Gaussian prior for the intrinsic scatter centred on the best fit value of Bleem et al. (2020) allows for the widest possible solutions cosmological analysis of the observational dataset (Table. 4.3) and simulated catalogue (Table. 4.3). For this reason, best fit values of all parameters are consistently stated for runs which adopt the Bleem et al. (2020) prior.

4.1.12 Discussion

The input cosmological and richness-mass relation parameters are successfully retrieved using the likelihood model on the simulated catalogue in the validation procedure (Appendix 4.1.8). This confirms that there are no major numerical issues with the cosmological pipeline. However, as the observed data (Fig. 4.2) appears much noisier than the simulated data (Fig. 4.7) it is apparent that it is considerably more complex than the model employed in this work. Failure to account for such complexities can cause shifts in the measured parameters relative to their true value. This section discusses several possible sources of systematic uncertainties and their relevance in the context of this work.

Mass function systematics

The systematic uncertainty imposed by the chosen HMF model, is subdominant relative to the measured precision for all constrained parameters. This is confirmed by recovering almost identical posterior distributions after repeating the abundance modelling with and a Tinker et al. (2008), Bocquet et al. (2016) and Despali et al. (2016) multiplicity function in Eq. 4.6. In theory, this source of uncertainty could be reduced further by using a high accuracy HMF emulator (e.g. McClintock et al., 2019b; Nishimichi et al., 2019; Bocquet

et al., 2020). However, as this analysis assumes $h = 0.7$, the cosmological parameter space covered by the majority of the (publicly available) emulators is not sufficiently wide to cover the broad posterior distributions of Ω_{m_0} and σ_8.

Redshift bias

The benefit of using spectroscopic redshifts means that the redshifts are extremely accurate (unbiased) relative to photometric redshifts (Clerc et al., 2020; Kirkpatrick et al., 2021), therefore there is no contribution to the systematic uncertainty due to photometric redshift estimation. The median statistical uncertainties of spectroscopic redshifts (Sect. 4.1.1) within the SPIDERS DR16 sample over the redshift range considered in this analysis, $z \in [0.1, 0.6)$, is of the order $\Delta_z/(1+\tilde{z}) \sim 7 \times 10^{-4}$. When compared to photometric redshifts derived from the SDSS (the DESI Legacy Imaging Surveys; Sect. 3.4.1), equivalent statistical uncertainties are found to be ~ 0.0097 (0.0075), meaning the impact of redshift measurement uncertainties is reduced by a factor of 13 (10). SPIDERS DR16 spectroscopic redshifts are therefore an excellent approximation to the true cluster redshift, allowing one to safely avoid marginalising over the redshift uncertainties when modelling of the abundance of SPIDERS clusters (i.e. the PDF which considers the uncertainties on observed redshifts in Eq. 4.4 can be approximated as a Dirac Delta function).

Richness bias

During the visual inspection procedure within the SPIDERS cluster pipeline (Sect. 4.1.1), it is possible to disentangle structure along the line of site. This reduces the impact that projection effects have on the number of redshift outliers, with $\sim 10\%$ of SPIDERS systems being multi-component (Clerc et al., 2020). The value of richness, however, can still be severely affected, usually resulting in up-scattered richness values for structure which cannot be de-blended. Ignoring this effect has been shown to systematically shift σ_8 high and Ω_{m_0} low, relative to their true values and can increase the instrinsic scatter (Costanzi et al., 2019a). Recent cosmological analyses using optically selected clusters (Costanzi et al., 2019b; Abbott et al., 2020) include specially designed terms within their abundance model to account for biases caused by projection effects. For SPIDERS, however, an equivalent term has yet to be derived. To study this effect in detail requires accurate methods to jointly mock X-ray observables and optical richness, which will be addressed in future work.

Percolation effects describe the process of assigning member galaxies which appear in multiple cluster candidates to the cluster with the maximum redMaPPer likelihood (Rykoff et al., 2014). In some cases, this can cause low-richness systems to be blended into richer systems when aligned along the line of sight. This results in a deficiency in the number of low-richness systems and higher values of richness for systems which are assigned the additional member galaxies. Modifying the percolation radii parameters to extreme values when constructing the DES redMaPPer cluster sample was shown to have a cosmologically negligible impact on the total number of systems (Abbott et al., 2020).

Although, the impact is also expected to be small for this study, percolation effects have yet to be quantified for the CODEX cluster sample when measuring richness with redMaPPer configured in *scanning-mode*.

The optical centre chosen by redMaPPer also has a systematic effect on the value of observed richness as it defines the radial aperture over which the membership probabilities of galaxies are summed. The performance of centring algorithm(s) in redMaPPer has been studied at length (e.g. Hoshino et al., 2015; Hikage et al., 2018) with a focus on comparisons to fiducial X-ray selected centres from high-quality archival data (Rozo & Rykoff, 2014; Zhang et al., 2019). As X-rays trace the dominant component of baryonic mass in the ICM, X-ray centres are in general more closely linked to the true centre for relaxed clusters. However, as CODEX clusters are detected using *ROSAT* data, the positional accuracy of the X-ray centres is $\sim$ 3 arcmin (Clerc et al., 2012). This means that on average, optical centres found by redMaPPer, which use prior information of CODEX X-ray source positions, are expected to be much closer to the fiducial cluster centres than the priors themselves.

For SPIDERS clusters analysed with The Legacy Surveys with $\tilde{\lambda} > 20$, the mean distance between optical centre determined by redMaPPer and the original RASS X-ray source is found to be 0.19Mpc with a standard deviation of 0.10Mpc, however, whether or not the optical centres of CODEX clusters analysed with The Legacy Surveys are more accurate than those found in purely optical selected clusters has yet to be studied in detail. Given that the difference between the value of richness at the optical centre and the mean value of richness averaged over the five most likely centres (within of 0.40Mpc of the X-ray centre) is less than the measured uncertainty on the richness for 99.4% of systems, it is unlikely that centring issues significantly contribute to the systematic uncertainties of this analysis.

Selection bias

Selection bias is the largest contribution to the total error budget. This is because the parameters of the X-ray luminosity-mass relation used in the construction of the CODEX selection function are fixed (Lieu et al., 2016; Giles et al., 2016) along with $h = 0.7$ and $\Omega_{m_0} = 0.30$ (Finoguenov et al., 2020). This causes any inaccuracies in these parameters to be absorbed into the posterior distributions of the richness-mass relation and cosmological parameters. In an attempt to minimise the overall systematic impact of these fixed parameters, only systems which are above the 50% sensitivity limit of the survey are taken into consideration when defining the cosmological cluster sub-sample (Sect. 4.1.5). Assuming these parameters of the X-ray luminosity-mass are known within 10% of their true values, the resulting contribution to the systematic uncertainty on the abundance prediction would be a maximum of $\sim$ 20%. Conducting the cosmological analysis without marginalising over Ω_{m_0} in the X-ray selection function, however, is estimated[10] to cause systematic

[10]Systematic uncertainties are estimated by repeating the cosmological analysis on the data with $\Omega_{m_0} = 0.24$ and $\Omega_{m_0} = 0.36$ fixed in the X-ray selection function, $P(I_{\rm X} \,|\, \mu, z, \nu)$, with a restrictive prior on the intrinsic scatter $\sigma_{\rm int} \in [0.1, 0.3]$.

uncertainties on the richness-mass relation of at least 5% on the normalisation, 8% on the mass evolution and 50% on the redshift evolution. The respective impact on cosmological parameters is $\sim 20\%$ for Ω_{m_0} and $\sim 15\%$ for σ_8.

4.1.13 Summary

Ω_m and σ_8 are constrained along with the parameters of the richness-mass scaling relation using a high-purity sub-sample of 691 CODEX clusters over the SPIDERS DR16 footprint (5,350deg^2) after a reanalysis of the original CODEX source catalogue using redMaPPer and The Legacy Surveys. The cluster abundance is modelled by considering the X-ray selection, a redshift dependent cut in observed-richness and the selection effect due to applying redMaPPer to The Legacy Surveys in bins of observed richness and redshift. The value of having spectroscopic redshifts is demonstrated via a fine-grain binning strategy that enables a relatively high cosmological constraining power. Despite this, the extrapolation of a complex selection model in the high redshift regime leads to underestimated statistical uncertainties due to its fixed cosmology and luminosity-mass relation. This also contributes to the systematics that affect the redshift evolution of the richness-mass relation which indicates an increasing over-abundance of rich clusters with redshift given the best-fit cosmology. In order to improve the results presented here, it is necessary to marginalise over all cosmologically dependent parameters including the Hubble constant, baryon fraction, neutrinos etc. as well as the luminosity-mass and richness-mass scaling relations in all aspects of the model (e.g. selection processes and HMF emulation). To improve the accuracy of the model, the likelihood function should also be modified to include projection/percolation effects and consider sample variance in the covariance matrix.

Modelling the low-photon statistics of RASS combined with optical richness from The Legacy Surveys is challenging, and the quality of the X-ray and photometric data can be improved to facilitate the usage of lower-scatter mass proxies which also simplify the selection modelling. The self-calibration procedure of the richness-mass relation would also largely benefit from accurate mass measurements from gravitational lensing, dynamics and/or high quality X-ray follow-up (e.g. Mantz et al., 2015; Kirby et al., 2019; Bocquet et al., 2020). The combination of *eROSITA* (Merloni et al., 2012), wide/deep photometric surveys such as LSST (Ivezić et al., 2019) and multi-object spectrograph like 4MOST (de Jong et al., 2012) and DESI (DESI Collaboration et al., 2016) will provide the next generation of large spectroscopic cluster samples. These will benefit from the investments into the SPIDERS programme and supersede the DR16 dataset to provide high-precision competitive cosmological constraints with well understood systematics.

4.2 SPIDERS Cosmology: Velocity Dispersion[11]

Although the primary cosmological analysis for the CODEX component of the SPIDERS DR16 cluster sample is carried out using optical richness as a mass proxy (see Sect. 4.1 Ider Chitham et al., 2020). Constraints can be derived in an analogous way using observed velocity dispersion, $\tilde{\sigma}$, as a dynamical mass proxy. This derivation of this quantity is described fully in Sect. 2.3.

The most accurate estimates of velocity dispersion come from the systems at low redshift with the highest number of spectroscopically confirmed member galaxies. Therefore, in this version of the analysis, the original SPIDERS catalogue (Clerc et al., 2020; Kirkpatrick et al., 2021) is utilised as opposed to the revised catalogue with the improved optical information from The Legacy Surveys (Ider Chitham et al., 2020, summarised in Sect. 4.1 of this thesis). This decision is justified by the higher accuracy of the velocity dispersion estimates due to a more rigorous visual inspection process with a larger number of inspectors (at least 3 per cluster) as well as the fact that the improved richness-estimator is not required and that the SDSS is volume limited over a redshift range of $\tilde{z} \in [0.1, 0.3)$.

4.2.1 Adapting the cosmological model

To conduct a cosmological analysis using velocity dispersion and the original SPIDERS D16 catalogue, a modification to the PDF which describes the relation between the observed and true quantities is necessary. This means $P(\ln\tilde{\lambda} \mid \ln\lambda(\mu, z))$ must be replaced with $P(\ln\tilde{\sigma} \mid \ln\sigma(\mu, z))$ in the original likelihood function (Eq. 4.21), where $P(\ln\tilde{\sigma} \mid \ln\sigma(\mu, z))$ is a normal distribution centred on the value of the natural logarithm of the true velocity dispersion ($\ln\sigma$) added to a mean bias value, with a scatter of both statistical and intrinsic origin.

Bias

The mean bias can be modelled as a function of the radial aperture through a multiplicative factor $B_{\rm ap}$ (see upper panel of Fig. 4.13), the number of galaxies entering the measurement ($N_{\rm gal}$) and the interloper fraction[12] as a function of cluster redshift and radial aperture (Saro et al., 2013). The interloper fraction (upper panel of Fig. 4.14) sets the value of the multiplicative bias factor $B_{\rm inter}$ on the velocity dispersion measurement (Ferragamo et al., 2020, see lower panel of Fig. 4.13). Given that only the most massive galaxies contribute to the measurement, it is necessary to also include an additional bias factor of 0.98 (Ferragamo et al., 2020). When combined, these sources of bias provide a model for the expected velocity dispersion (Eq. 4.23), which is consistent with that derived from SPIDERS-like resampling of bright nearby clusters (Zhang et al., 2017).

[11] This description of the SPIDERS velocity dispersion cosmological analysis is adapted from the author's major contribution to Kirkpatrick et al. (2021).

[12] The interloper fraction refers to the fraction of galaxies that enter the computation of velocity dispersion despite being located outside of the virially bound cluster region.

$$\langle\sigma\rangle = \frac{0.98\sigma}{0.9775 + \left(\frac{0.72}{N_{\rm gal}-1}\right)^{1.28}} \cdot B_{\rm ap}(\Gamma_{\rm ap}) \cdot B_{\rm inter}(f_{\rm inter}(\Gamma_{\rm ap}, z), N_{\rm gal}). \tag{4.23}$$

Here, the ratio of the between the aperture radius and R_{200} is

$$\Gamma_{\rm ap} = \frac{R_{\rm ap}}{R_{\rm 200c}(\mu, z)} \tag{4.24}$$

Scatter

The variance on the natural logarithm of observed velocity dispersion (Σ^2_σ, Eq. 4.25) is approximated as the sum of the squares of two separate relations. The first accounts for the statistical scatter on $\ln(\tilde{\sigma}/\sigma)$ as a function of $N_{\rm gal}$ to the scatter ($\Sigma_{\rm statistical} = -0.037 + 1.04/\sqrt{N_{\rm gal}}$, Saro et al., 2013) and the second considers the additional contribution due to the interlopers and aperture selection ($\Sigma_{\rm inter}(\Gamma_{\rm ap}(\mu, z), M_{\rm vir}(\mu, z))$, shown in the lower panel of Fig. 4.14, Saro et al., 2013)

$$\Sigma^2_\sigma = \Sigma^2_{\rm statistical}(N_{\rm gal}) + \Sigma^2_{\rm inter}(\Gamma_{\rm ap}(\mu, z), M_{\rm vir}(\mu, z)) \tag{4.25}$$

Assuming that $\ln\tilde{\sigma}$ is normally distributed around $\ln\sigma$ with a bias term, the additive bias b on $\ln\tilde{\sigma}$ can be written as

$$b(R_{\rm ap}, M_{\rm vir}(\mu, z), N_{\rm gal}) = \langle\ln\sigma\rangle - \ln\sigma = \ln\langle\sigma\rangle - \frac{\Sigma^2_\sigma}{2}(\mu, z, N_{\rm gal}). \tag{4.26}$$

The total scatter on $\ln\tilde{\sigma}$ is then be modelled the scatter due to measurement uncertainties (Σ_σ) added to to an intrinsic scatter term (Σ_0).

$$\Sigma = \sqrt{\Sigma^2_\sigma(\mu, z, N_{\rm gal}) + \Sigma^2_0}. \tag{4.27}$$

The probability to observe a velocity dispersion given its true value is then given by

$$P(\ln\tilde{\sigma}|\ln\sigma(\mu, z), z) = \mathcal{N}(\ln\tilde{\sigma}, \ln\sigma + b(R_{\rm ap}, M_{\rm vir}(\mu, z), N_{\rm gal}), \Sigma(R_{\rm ap}, M_{\rm vir}(\mu, z), N_{\rm gal})). \tag{4.28}$$

As this distribution is dependent on $N_{\rm gal}$, the observed catalogue is tabulated and inverted as a callable function $N_{\rm gal}(\sigma(\mu, z), z)$, which subsequently defines the bias b and the scatter Σ_σ of the observed values. Here, the true velocity dispersion, $\sigma(\mu, z), z)$, is simply a function of mass and true redshift (Carlberg et al., 1997)

$$\sigma(\mu, z) = 10H(z)\frac{R_{\rm 200c}(\mu, z)}{\sqrt{3}}\,. \tag{4.29}$$

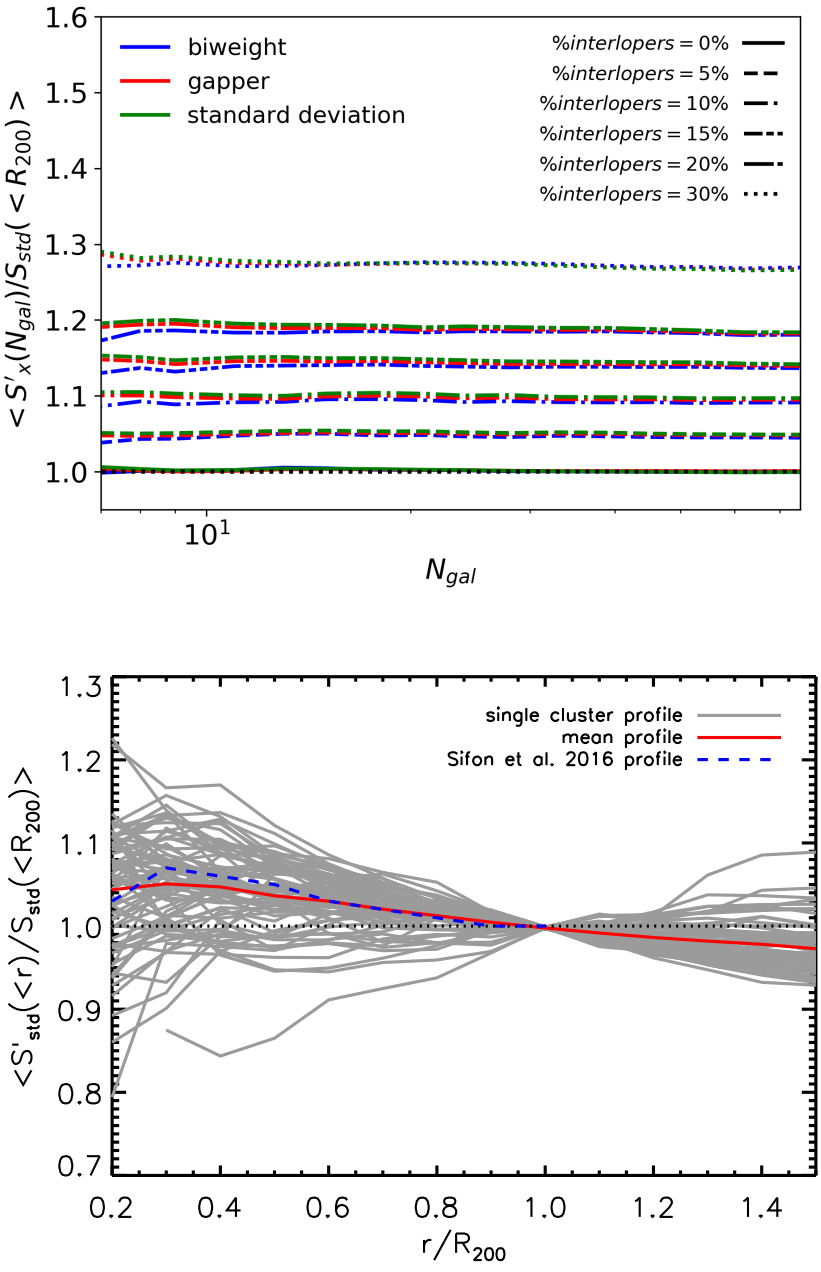

Abbildung 4.13: Adapted from Ferragamo et al. (2020). **Upper panel:** $B_{\rm inter}(f_{\rm inter}(\Gamma_{\rm ap}, z), N_{\rm gal})$, the interloper bias as a function the number of spectroscopic member galaxies and interloper fraction $f_{\rm inter}$ as shown in the upper panel of Fig. 4.14. **Lower panel**: The the mean profile aperture bias, $B_{\rm ap}(\Gamma_{\rm ap})$, is shown by the solid red line.

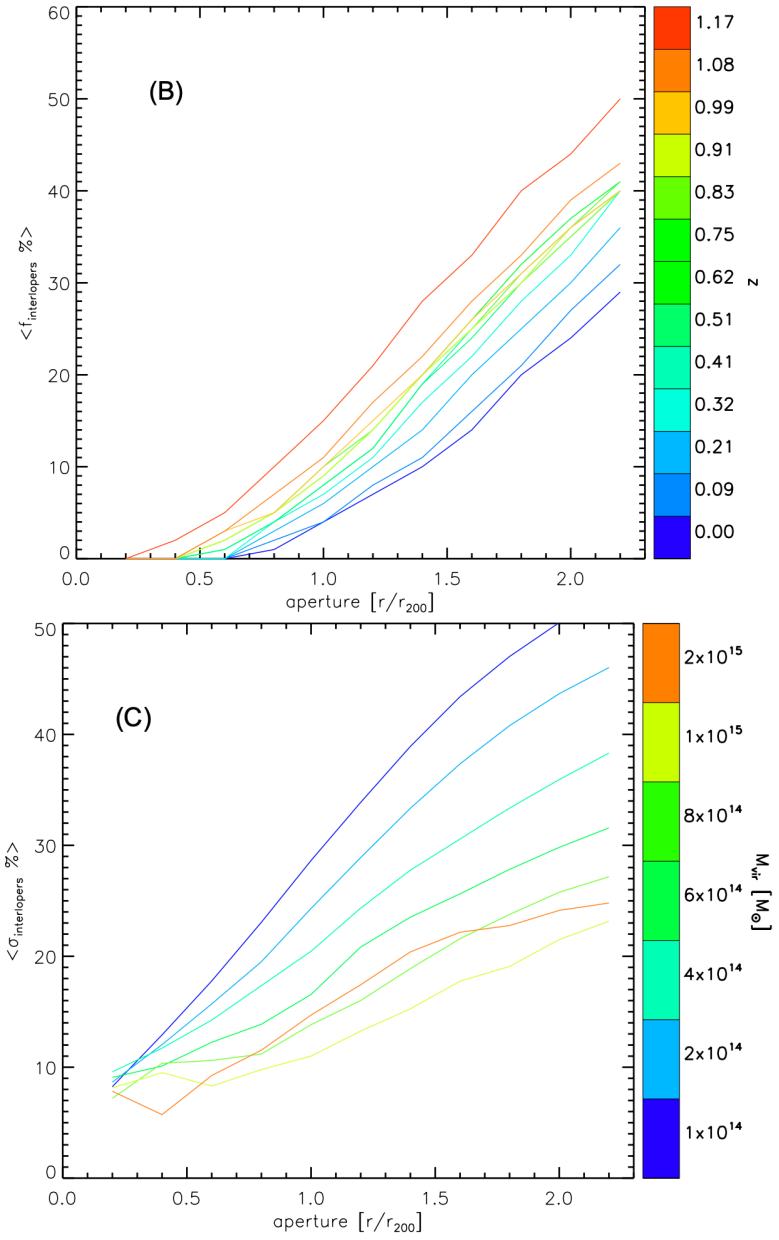

Abbildung 4.14: Adapted from (Saro et al., 2013). **Upper panel**: $f_{\rm inter}(\Gamma_{\rm ap}, z)$, the stacked mean fraction of interlopers (defined as galaxies with $R > 3R_{\rm vir}$) after red-sequence selection as a function of maximum projected separation from the cluster $R_\perp$ normalised to R_{200} colour coded by redshift (B). Combing this interloper fraction with the number of spectroscopic member galaxies and the model shown in the upper panel of Fig. 4.13 then provides, $B_{\rm inter}(f_{\rm inter}(\Gamma_{\rm ap}, z), N_{\rm gal})$, the bias due to interlopers. **Lower panel**: $\Sigma_{\rm inter}(\Gamma_{\rm ap}(\mu, z), M_{\rm vir}(\mu, z))$, the stacked mean scatter contribution from interlopers as a function of maximum projected separation from the cluster $R_\perp$ normalised to R_{200} colour coded by virial cluster mass (C).

Updated Selection terms

In order to obtain a clean X-ray selected cluster sample, a redshift dependent cut in observed velocity dispersion is applied to both the data and the modelling of the bins of observed velocity dispersion and redshift $\Delta \ln \tilde{\sigma}_i, \Delta \tilde{z}_j$ in an analogous way to the redshift dependent richness cut presented in Sect. 4.1.5. This term corresponds to the 10% sensitivity cut outlined in Finoguenov et al. (2020),

$$P(I_{\rm cut}|\tilde{\sigma}, \tilde{z}) = \theta(\tilde{\sigma} - 375(\tilde{z}/0.15)^{0.38}). \tag{4.30}$$

The second term is then analogous to $P(I_{\rm Legacy} \mid \ln \tilde{\lambda}, \tilde{z})$, but describes the 90% completeness optical of the SDSS rather than the optical selection effects due to The Legacy Surveys (since only SDSS data are used in this version of the analysis),

$$P(I_{\rm SDSS}|\tilde{\sigma}, \tilde{z}) = \theta(\tilde{\sigma} - (372 + 2.76e^{3.3\tilde{z}^2})), \tag{4.31}$$

where θ denotes a step-function in both cuts. These distributions are then combined with the CODEX selection function to model the expected number counts (Eq. 4.32) in bins of observed velocity dispersion ($\Delta \ln \tilde{\sigma}_i$) and redshift ($\Delta \tilde{\sigma}_i$).

$$\begin{aligned}\langle N(\Delta \ln \tilde{\sigma}_i, \Delta z_j)\rangle = \int_{\Delta z_j} {\rm d}z \frac{{\rm d}V}{{\rm d}z} \int_{\Delta \ln \tilde{\sigma}_i} {\rm d}\ln \tilde{\sigma} \int {\rm d}\mu \frac{{\rm d}n(\mu, z)}{{\rm d}\mu} \cdot P(\ln \tilde{\sigma} | \ln \sigma(\mu, z), z) \\ \cdot P_X(I|\mu, z, \nu = 0) \\ \cdot P(I_{\rm sdss}|\tilde{\sigma}, z) \\ \cdot P(I_{\rm cut}|\tilde{\sigma}, z).\end{aligned} \tag{4.32}$$

For simplicity covariance between the velocity dispersion and X-ray luminosity is not considered. This reduces the CODEX survey function to $P(I_{\rm X} \mid \mu, z, \nu = 0)$ (Eq. 4.16).

The binned data and best fit model are shown in Figure 4.15 for three logarithmically distributed bins of velocity dispersion over a redshift range of $[0.1, 0.3)$ with a bin-width of 0.05. The corresponding constraints are $\sigma_8 = 0.74^{+0.03}_{-0.02}$, $\Omega_{m_0} = 0.33^{+0.02}_{-0.02}$ and the velocity dispersion of the intrinsic scatter $\Sigma_0 = 0.24^{+0.02}_{-0.02}$ derived from the Markov Chain Monte Carlo (MCMC) process are shown in Fig. 4.16. As these results depend on a large number of fixed parameters which reduce the flexibility of the likelihood, a significant underestimation of the apparent uncertainties is to be expected. However, the cosmological constraints derived from velocity dispersion show a strong similarity to the constraints derived from optical richness. These constraints therefore overlap with the same cluster cosmological analysis summary shown in Fig. 4.12. The intrinsic scatter of velocity dispersion has been measured for the first time, however, compared to simulations, it is higher than the predictions made by Saro et al. (2013) and the lower than that of Munari et al. (2013). This low intrinsic scatter is also comparable to other cluster mass proxies such as X-ray luminosity, X-ray temperature, optical richness and SZ thermal energy (Capasso et al., 2019c,a; Mulroy et al., 2019; Bleem et al., 2020; Kiiveri et al., 2021) and provides good prospects for the *eROSITA* cluster follow-up program with 4MOST (Finoguenov et al., 2019b) and SDSS-V.

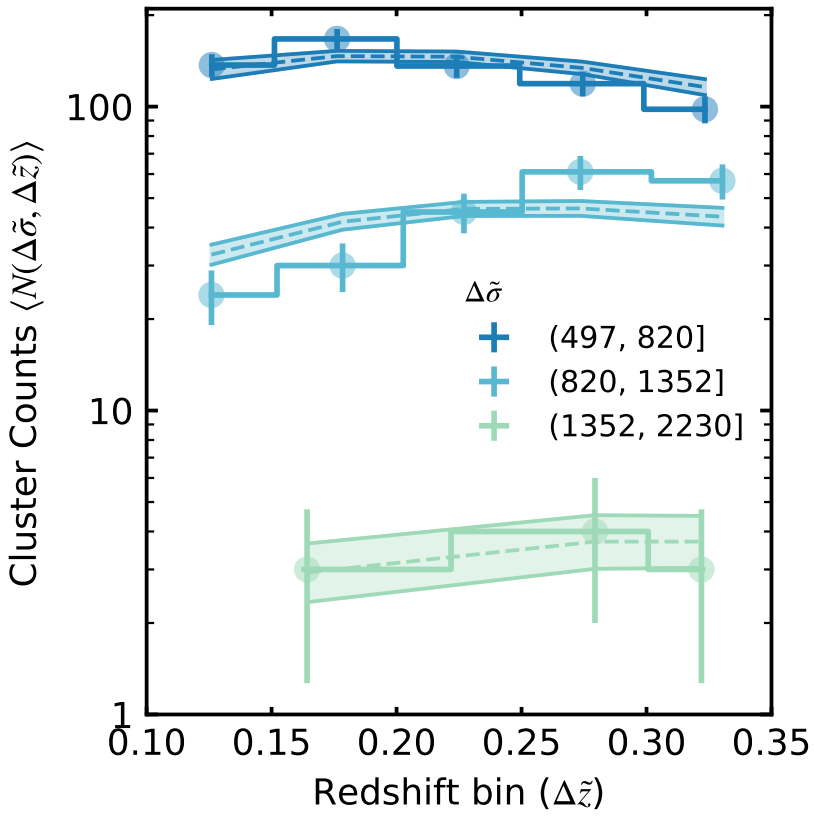

Abbildung 4.15: Adapted from Ider Chitham et al. (2020); Kirkpatrick et al. (2021). The abundance of SPIDERS clusters as a function in bins (Δ) of observed redshift ($\Delta\tilde{z}_j$) and velocity dispersion ($\Delta\sigma_i$) where $\tilde{z} \in [0.1, 0.3)$ and $\tilde{\sigma} \in [497, 2230)$. Systems are selected according to Eq. 4.30 and Eq. 4.31. These cuts on the observable space is accounted for in the cosmological likelihood function. Steps represent the observed data and shaded regions trace the expectation value of the model, centred on the median. The upper and lower limit correspond to the 15% and 85% confidence intervals of the MCMC chains.

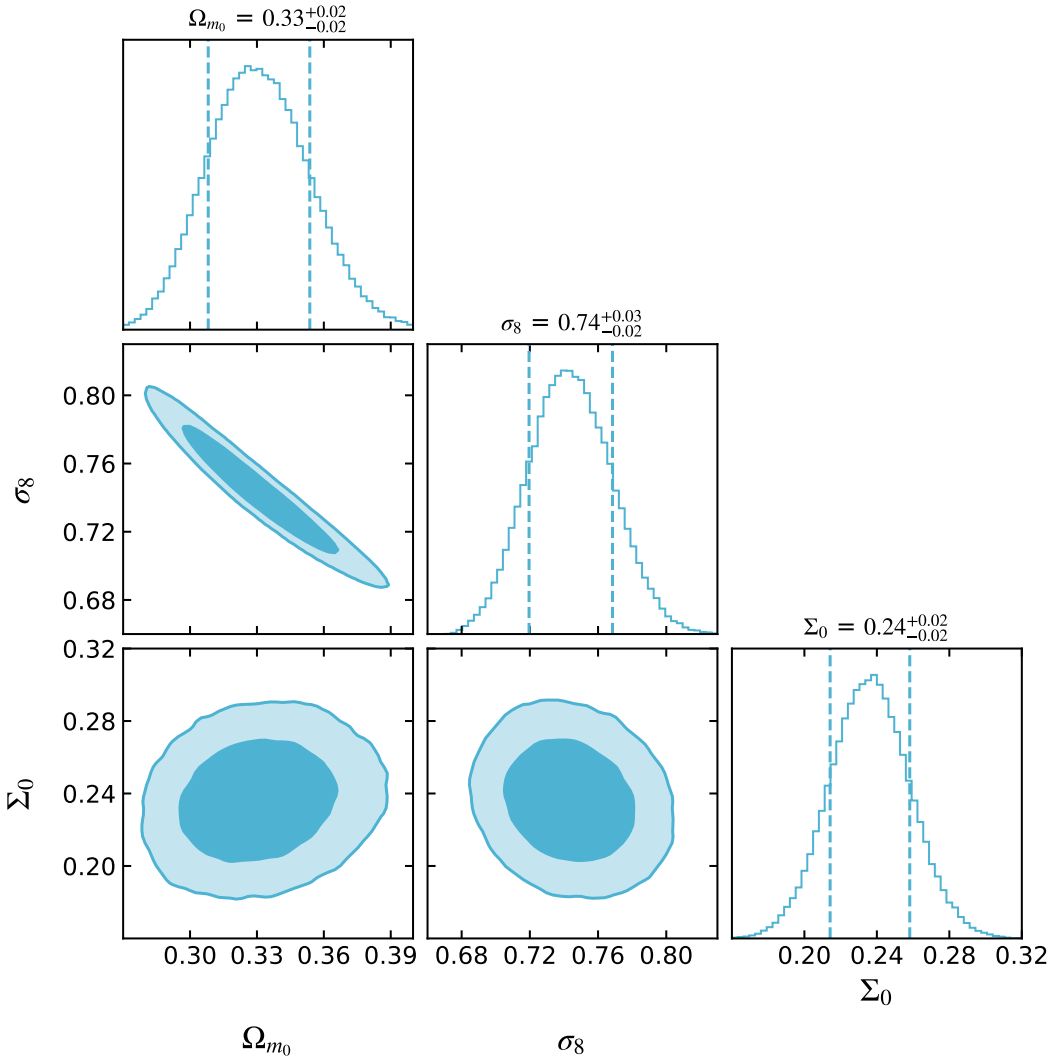

Abbildung 4.16: Adapted from Ider Chitham et al. (2020); Kirkpatrick et al. (2021). Constraints on cosmological parameters and the intrinsic scatter of the velocity dispersion constraints for SPIDERS DR16. Contours depict the 68% and 95% confidence levels where posterior distributions are obtained using the likelihood function Eq. 4.32.

Kapitel 5

Conclusion

5.1 Thesis Summary

This concluding chapter aims to summarise the scientific results as well as technical contributions that this thesis has made to the *eROSITA* cluster cosmology experiment and the field in general. These contributions include:

- An optical follow-up method that is capable of robustly measuring the photometric and spectroscopic redshifts of X-ray selected galaxy clusters as well as their optical and dynamical properties. This method is summarised in a self-contained pipeline for use by the *eROSITA*-DE consortium (soon to be publicly available), which works for any multi-band photometric survey and/or spectroscopic galaxy catalogue. It is designed to be extremely scalable i.e. it is extremely rapid, memory efficient and capable of running on an arbitrary number of cores using big data frameworks (e.g. spark, Peloton et al., 2018; Plaszczynski et al., 2019) deployed on high-performance computers. This pipeline includes everything from the generation of red-sequence galaxy models, to the ingestion and pixelisation (reduction) of terrascale photometric catalogues, to the creation of depth maps and masks (spatial and redshift masks). These data products can then be used to create volume limited cluster catalogues with a well defined optical selection function. This pipeline is an essential tool for the astrophysical and cosmological analysis of the eRASS cluster sample. An extensive compilation of spectroscopic galaxies in the literature improves the precision of cluster redshifts by a factor of 10 on average when more than three members per cluster have spectroscopic redshift information. The reliability of redshifts and optical likelihoods has also been investigated over a wide redshift range using a large compilation of SZ, X-ray and optically selected clusters from the literature.

- Spectroscopic targets for SDSS-IV (SPIDERS) over the eBOSS footprint, the 4MOST (cluster redshift survey), SDSS-V (Black-Hole Mapper) which are based on the optical follow-up of cluster samples from CODEX, eRASS and eFEDS (*eROSITA* Final Equatorial Depth Survey, Liu et al., 2021). Optically selected cluster samples have

also been generated over the *eROSITA*-DE footprint to ensure target selection is as complete as possible. Targets have also been observed using the AAT for 7 pointings of 2dF/AAOmega around the south ecliptic pole and 1 additional pointing south of the border of The Legacy Surveys. In support of the 4MOST *eROSITA* cluster redshift survey, this thesis also presents a state of the algorithm to determine optimal sky fibre positions. This is a rapid and accurate method which is scaleable to arbitrarily large spectroscopic surveys with complex geometries. It has also been successfully validated during the aforementioned AAT observations. Additionally, the accuracy of visual inspection procedure within SDSS-IV/SPIDERS has been validated using a catalogue of synthetic clusters and their galaxies to statistically determine the impact that manual inspection has on the redshift distribution of spectroscopic cluster samples.

- Two separate cosmological analyses based on the largest sample of spectroscopically confirmed X-ray selected clusters. The first uses optical richness as a mass proxy and the second uses velocity dispersion. The likelihood functions and cosmological pipeline are also validated using a simulated cluster catalogue which includes all contributions to the X-ray and optical selection functions. Constraints from richness and velocity dispersion analyses are consistent with one another and show a good agreement with respect to previous cluster count experiments in the literature (especially those based on X-ray selected clusters). The cosmological pipeline used throughout this thesis was designed such that in can be easily adapted for the eRASS cluster counts experiment. In other words, SPIDERS DR16 and the cosmological analyses presented in this thesis have effectively been a pilot program for the analysis of the next generation of spectroscopic *eROSITA* cluster surveys. This pipeline is therefore available to the *eROSITA*-DE consortium (and the public) and supports any arbitrary mass proxy (or direct mass measurements) for any cluster catalogue with configurable selection and likelihood functions.

The *eROSITA* cluster counts experiment will be the most precise cluster cosmology analysis ever conducted. Given a large sample of X-ray selected clusters, the minimum requirements for such an experiment to be successful are: a mass proxy[1] (or several), unbiased and precise cluster redshifts spanning as much cosmological volume as possible, a well defined selection function and an accurate cosmological modelling framework. In terms of the optical (and NIR) side of the multi-wavelength follow-up of X-ray selected clusters, this thesis provides all of the above. Applying the methods and tools presented in this thesis directly to the final eRASS:8 cluster catalogue is expected to produce competitive cosmological constraints that are an order of magnitude more precise than results obtained from SPIDERS DR16. This will be an unprecedented milestone for the field of cluster cosmology and when combined with the next generation of cosmological analyses from complementary probes (CMB, BAO, SNeIa etc.) we will be closer than ever before to understanding the mysterious dark components of the Universe.

[1] Mass measurements from the lensing or dynamical analysis of a subset of clusters are also very desirable.

5.2 Acknowledgements

This thesis represents an effort by both the SDSS-III and SDSS-IV collaborations. Funding for SDSS-III was provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatory of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University. This work has made extensive use the SOA/NASA Astrophysics Data System, `TOPCAT` astronomical data analysis software (Taylor, 2005) and `spark-fits` (Peloton et al., 2018; Plaszczynski et al., 2019). Data processing is carried out with the Draco, Cobra and Raven High Performance Computing Systems located at the Max Planck Computing and Data Facility[2]. The author acknowledges the following python packages: `mpi4py` (Dalcín et al., 2005, 2008; Dalcin et al., 2011), `astropy` (Astropy Collaboration et al., 2013, 2018), `numpy` (Oliphant, 2006; van der Walt et al., 2011), `scipy` (Virtanen et al., 2019), `pandas` (Team, 2020), `matplotlib` (Hunter, 2007) and `corner` (Foreman-Mackey et al., 2020).

[2]https://www.mpcdf.mpg.de/

Literaturverzeichnis

Abazajian K., Dodelson S., 2003, Phys. Rev. Lett., 91, 041301

Abbott T. M. C., et al., 2018, Phys. Rev. D, 98, 043526

Abbott T. M. C., et al., 2020, Phys. Rev. D, 102, 023509

Abell G. O., 1958, ApJS, 3, 211

Abolfathi B., et al., 2018, ApJS, 235, 42

Adami C., et al., 2018, A&A, 620, A5

Aguado-Barahona A., Barrena R., Streblyanska A., Ferragamo A., Rubiño-Martín J. A., Tramonte D., Lietzen H., 2019, A&A, 631, A148

Aguena M., et al., 2021, MNRAS,

Agulli I., Aguerri J. A. L., Sánchez-Janssen R., Dalla Vecchia C., Diaferio A., Barrena R., Dominguez Palmero L., Yu H., 2016, MNRAS, 458, 1590

Ahumada R., et al., 2020, ApJS, 249, 3

Aihara H., et al., 2011, The Astrophysical Journal Supplement Series, 193, 29

Aihara H., et al., 2019, PASJ, 71, 114

Alam S., et al., 2015, ApJS, 219, 12

Alam S., et al., 2017, MNRAS, 470, 2617

Albrecht A., et al., 2006, arXiv Astrophysics e-prints,

Allen S. W., Evrard A. E., Mantz A. B., 2011, ARA&A, 49, 409

Ambikasaran S., Foreman-Mackey D., Greengard L., Hogg D. W., O'Neil M., 2015, IEEE Transactions on Pattern Analysis and Machine Intelligence, 38

Amodeo S., et al., 2018, ApJ, 853, 36

Anderson L., et al., 2014, MNRAS, 439, 83

Arkani-Hamed N., Finkbeiner D. P., Slatyer T. R., Weiner N., 2009, Phys. Rev. D, 79, 015014

Arnaud M., Pratt G. W., Piffaretti R., Böhringer H., Croston J. H., Pointecouteau E., 2010, A&A, 517, A92

Astropy Collaboration et al., 2013, A&A, 558, A33

Astropy Collaboration et al., 2018, AJ, 156, 123

Ata M., et al., 2018, MNRAS, 473, 4773

Babcock H. W., 1939, Lick Observatory Bulletin, 498, 41

Bahcall N. A., 2015, Proceedings of the National Academy of Sciences, 112, 12243

Bahcall N. A., Cen R., 1992, ApJ, 398, L81

Bahcall N. A., Cen R., 1993, ApJ, 407, L49

Bahcall N. A., Lubin L. M., Dorman V., 1995, ApJ, 447, L81

Bahcall N. A., et al., 2003, ApJ, 585, 182

Balogh M. L., et al., 2014, MNRAS, 443, 2679

Balogh M. L., et al., 2020, MNRAS,

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

Basu-Zych A. R., et al., 2020, MNRAS, 498, 1651

Bautista J. E., et al., 2021, MNRAS, 500, 736

Baxter E. J., Rozo E., Jain B., Rykoff E., Wechsler R. H., 2016, MNRAS, 463, 205

Bayliss M. B., Hennawi J. F., Gladders M. D., Koester B. P., Sharon K., Dahle H., Oguri M., 2011, ApJS, 193, 8

Bayliss M. B., et al., 2016, ApJS, 227, 3

Bayliss M. B., et al., 2017, ApJ, 837, 88

Becker M. R., et al., 2007, ApJ, 669, 905

Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32

Bellagamba F., Roncarelli M., Maturi M., Moscardini L., 2018, MNRAS, 473, 5221

Bennett C. L., et al., 2013, ApJS, 208, 20

Bergmann A. G., Petrosian V., Lynds R., 1990, ApJ, 350, 23

Bertin E., 2011, in Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds, Astronomical Society of the Pacific Conference Series Vol. 442, Astronomical Data Analysis Software and Systems XX. p. 435

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

Bertone G., Hooper D., 2018, Reviews of Modern Physics, 90, 045002

Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279

Beutler F., et al., 2011, MNRAS, 416, 3017

Bezanson R., et al., 2018, ApJ, 858, 60

Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122

Bird S., Cholis I., Muñoz J. B., Ali-Haïmoud Y., Kamionkowski M., Kovetz E. D., Raccanelli A., Riess A. G., 2016, Phys. Rev. Lett., 116, 201301

Birkinshaw M., 1999, Phys. Rep., 310, 97

Biviano A., Murante G., Borgani S., Diaferio A., Dolag K., Girardi M., 2006, A&A, 456, 23

Blanton M. R., et al., 2017, AJ, 154, 28

Blas D., Lesgourgues J., Tram T., 2011, Journal of Cosmology and Astro-Particle Physics, 2011, 034

Bleem L. E., et al., 2015, ApJS, 216, 27

Bleem L. E., et al., 2020, ApJS, 247, 25

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bocquet S., Saro A., Dolag K., Mohr J. J., 2016, MNRAS, 456, 2361

Bocquet S., et al., 2019, ApJ, 878, 55

Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5

Böhringer H., et al., 2000, ApJS, 129, 435

Böhringer H., et al., 2004, A&A, 425, 367

Böhringer H., Chon G., Collins C. A., 2014, A&A, 570, A31

Bolton A. S., et al., 2012, AJ, 144, 144

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Borgani S., Guzzo L., 2001, Nature, 409, 39

Borm K., Reiprich T. H., Mohammed I., Lovisari L., 2014, A&A, 567, A65

Bradshaw E. J., et al., 2013, MNRAS, 433, 194

Bruzual G., 2007, in Vallenari A., Tantalo R., Portinari L., Moretti A., eds, Astronomical Society of the Pacific Conference Series Vol. 374, From Stars to Galaxies: Building the Pieces to Build Up the Universe. p. 303 (arXiv:astro-ph/0702091)

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Burke C., Hilton M., Collins C., 2015, MNRAS, 449, 2353

Caminha G. B., et al., 2017, A&A, 600, A90

Caminha G. B., et al., 2019, A&A, 632, A36

Cannon R., et al., 2006, MNRAS, 372, 425

Capasso R., et al., 2019a, arXiv e-prints, p. arXiv:1910.04773

Capasso R., et al., 2019b, MNRAS, 482, 1043

Capasso R., et al., 2019c, MNRAS, 486, 1594

Capozziello S., de Laurentis M., 2011, Phys. Rep., 509, 167

Carlberg R. G., Yee H. K. C., Ellingson E., 1997, ApJ, 478, 462

Carr B., Kühnel F., Sandstad M., 2016, Phys. Rev. D, 94, 083504

Carrasco M., et al., 2017, ApJ, 834, 210

Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499

Castagné D., Soucail G., Pointecouteau E., Cappi A., Maurogordato S., Benoist C., Ferrari C., 2012, A&A, 548, A18

Castignani G., Pandey-Pommier M., Hamer S. L., Combes F., Salomé P., Freundlich J., Jablonka P., 2020, A&A, 640, A65

Cava A., et al., 2009, A&A, 495, 707

Chabrier G., 2003, PASP, 115, 763

Chambers K. C., et al., 2016, preprint, (arXiv:1612.05560)

Child H. L., Habib S., Heitmann K., Frontiere N., Finkel H., Pope A., Morozov V., 2018, ApJ, 859, 55

Childress M. J., et al., 2017, MNRAS, 472, 273

Chimento L. P., Jakubi A. S., Pavón D., Zimdahl W., 2003, Phys. Rev. D, 67, 083513

Chisari N. E., et al., 2019, ApJS, 242, 2

Ciotti L., 1991, A&A, 249, 99

Clerc N., Sadibekova T., Pierre M., Pacaud F., Le Fèvre J.-P., Adami C., Altieri B., Valtchanov I., 2012, MNRAS, 423, 3561

Clerc N., et al., 2016, MNRAS, 463, 4490

Clerc N., et al., 2018, A&A, 617, A92

Clerc N., et al., 2020, MNRAS,

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJ, 648, L109

Coe D., et al., 2012, ApJ, 757, 22

Coil A. L., et al., 2011, ApJ, 741, 8

Colless M., et al., 2001, MNRAS, 328, 1039

Comparat J., et al., 2019, MNRAS, 487, 2005

Comparat J., et al., 2020a, arXiv e-prints, p. arXiv:2008.08404

Comparat J., et al., 2020b, A&A, 636, A97

Connelly J. L., et al., 2012, ApJ, 756, 139

Connor T., Kelson D. D., Blanc G. A., Boutsia K., 2019, ApJ, 878, 66

Conroy C., Wechsler R. H., Kravtsov A. V., 2007, ApJ, 668, 826

Cooke R. J., Pettini M., Jorgenson R. A., Murphy M. T., Steidel C. C., 2014, ApJ, 781, 31

Cool R. J., et al., 2013, ApJ, 767, 118

Cooper M. C., et al., 2012, MNRAS, 425, 2116

Copeland E. J., Sami M., Tsujikawa S., 2006, International Journal of Modern Physics D, 15, 1753

Costanzi M., et al., 2019a, MNRAS, 482, 490

Costanzi M., et al., 2019b, MNRAS, 488, 4779

Costanzi M., et al., 2021, Phys. Rev. D, 103, 043522

Coupon J., Czakon N., Bosch J., Komiyama Y., Medezinski E., Miyazaki S., Oguri M., 2018, PASJ, 70, S7

Crawford S. M., Wirth G. D., Bershady M. A., Hon K., 2011, ApJ, 741, 98

Czoske O., Kneib J. P., Soucail G., Bridges T. J., Mellier Y., Cuillandre J. C., 2001, A&A, 372, 391

DESI Collaboration et al., 2016, preprint, (arXiv:1611.00036)

Dalcín L., Paz R., Storti M., 2005, Journal of Parallel and Distributed Computing, 65, 1108

Dalcín L., Paz R., Storti M., D'Elía J., 2008, Journal of Parallel and Distributed Computing, 68, 655

Dalcin L. D., Paz R. R., Kler P. A., Cosimo A., 2011, Advances in Water Resources, 34, 1124

Dalton G. B., Maddox S. J., Sutherland W. J., Efstathiou G., 1997, MNRAS, 289, 263

Danese L., de Zotti G., di Tullio G., 1980, A&A, 82, 322

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Davis M., et al., 2003, in Guhathakurta P., ed., Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4834, Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. pp 161–172 (arXiv:astro-ph/0209419), doi:10.1117/12.457897

Dawson W. A., 2013, ApJ, 772, 131

Dawson W. A., et al., 2015, ApJ, 805, 143

Dawson K. S., et al., 2016, AJ, 151, 44

Deger S., et al., 2018, ApJ, 869, 6

Dekker A., Peerbooms E., Zimmer F., Ng K. C. Y., Ando S., 2021, arXiv e-prints, p. arXiv:2103.13241

Demarco R., et al., 2007, ApJ, 663, 164

Demarco R., et al., 2010, ApJ, 711, 1185

Deshev B., et al., 2017, A&A, 607, A131

Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486

Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, IEEE Proceedings, 97, 1482

Dey A., et al., 2019, AJ, 157, 168

Diemer B., 2018, ApJS, 239, 35

Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108

Dietrich J. P., et al., 2019, MNRAS, 483, 2871

Dolag K., Gaensler B. M., Beck A. M., Beck M. C., 2015, MNRAS, 451, 4277

Doroshenko V., Ducci L., Santangelo A., Sasaki M., 2014, A&A, 567, A7

Dressler A., 1980, ApJ, 236, 351

Dressler A., Smail I., Poggianti B. M., Butcher H., Couch W. J., Ellis R. S., Oemler Augustus J., 1999, ApJS, 122, 51

Drinkwater M. J., et al., 2010, MNRAS, 401, 1429

Driver S. P., et al., 2011, MNRAS, 413, 971

Drlica-Wagner A., et al., 2018, ApJS, 235, 33

Duarte M., Mamon G. A., 2015, MNRAS, 453, 3848

Dwelly T., et al., 2017, MNRAS, 469, 1065

Ebeling H., Edge A. C., Bohringer H., Allen S. W., Crawford C. S., Fabian A. C., Voges W., Huchra J. P., 1998, MNRAS, 301, 881

Ebeling H., Edge A. C., Henry J. P., 2001, ApJ, 553, 668

Ebeling H., Edge A. C., Mantz A., Barrett E., Henry J. P., Ma C. J., van Speybroeck L., 2010, MNRAS, 407, 83

Ebeling H., et al., 2013, MNRAS, 432, 62

Ebeling H., Ma C.-J., Barrett E., 2014, ApJS, 211, 21

Eddington A. S., 1913, MNRAS, 73, 359

Einasto J., 1965, Trudy Astrofizicheskogo Instituta Alma-Ata, 5, 87

Einasto J., Kaasik A., Saar E., 1974, Nature, 250, 309

Einstein A., 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 142–152

Einstein A., de Sitter W., 1932, Contributions from the Mount Wilson Observatory, 3, 51

Eisenhardt P. R. M., et al., 2008, ApJ, 684, 905

Eisenhardt P. R. M., et al., 2020, ApJS, 247, 69

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Famaey B., McGaugh S. S., 2012, Living Reviews in Relativity, 15, 10

Farahi A., Evrard A. E., Rozo E., Rykoff E. S., Wechsler R. H., 2016, MNRAS, 460, 3900

Farahi A., et al., 2019, arXiv e-prints,

Feng J. L., 2010, ARA&A, 48, 495

Ferragamo A., Rubiño-Martín J. A., Betancort-Rijo J., Munari E., Sartoris B., Barrena R., 2020, in European Physical Journal Web of Conferences. p. 00011 (arXiv:1911.03184), doi:10.1051/epjconf/202022800011

Ferrari C., Maurogordato S., Cappi A., Benoist C., 2003, A&A, 399, 813

Finoguenov A., et al., 2019a, arXiv e-prints, p. arXiv:1912.03262

Finoguenov A., et al., 2019b, The Messenger, 175, 39

Finoguenov A., et al., 2020, A&A, 638, A114

Fixsen D. J., 2009, ApJ, 707, 916

Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576

Flewelling H., 2018, in American Astronomical Society Meeting Abstracts #231. p. 436.01

Foëx G., Chon G., Böhringer H., 2017, A&A, 601, A145

Foley R. J., et al., 2011, ApJ, 731, 86

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306

Foreman-Mackey D., et al., 2020, dfm/corner.py: corner.py v2.1.0.rc1, doi:10.5281/zenodo.3937526, https://doi.org/10.5281/zenodo.3937526

Fossati M., et al., 2019, MNRAS, 490, 1451

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385

Frusciante N., Perenon L., 2020, Phys. Rep., 857, 1

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Fumagalli A., et al., 2021, arXiv e-prints, p. arXiv:2102.08914

Gaia Collaboration et al., 2016, A&A, 595, A2

Gaia Collaboration et al., 2018, A&A, 616, A1

Garilli B., et al., 2008, A&A, 486, 683

Garilli B., et al., 2014, A&A, 562, A23

Garilli B., et al., 2021, arXiv e-prints, p. arXiv:2101.07645

Gifford D., Kern N., Miller C. J., 2017, ApJ, 834, 204

Gil-Marín H., et al., 2020, MNRAS, 498, 2492

Giles P. A., et al., 2016, A&A, 592, A3

Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148

Gladders M. D., Yee H. K. C., 2005, ApJS, 157, 1

Gladders M. D., López-Cruz O., Yee H. K. C., Kodama T., 1998, ApJ, 501, 571

Gladders M. D., Yee H. K. C., Majumdar S., Barrientos L. F., Hoekstra H., Hall P. B., Infante L., 2007, ApJ, 655, 128

Golovich N., van Weeren R. J., Dawson W. A., Jee M. J., Wittman D., 2017, ApJ, 838, 110

Golovich N., et al., 2019, ApJ, 882, 69

Gonzalez A. H., et al., 2019, ApJS, 240, 33

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, ApJ, 622, 759

Graham A. W., Driver S. P., 2005, Publ. Astron. Soc. Australia, 22, 118

Grandis S., Mohr J. J., Dietrich J. P., Bocquet S., Saro A., Klein M., Paulus M., Capasso R., 2018, preprint, (arXiv:1810.10553)

Grandis S., et al., 2020, arXiv e-prints, p. arXiv:2002.10834

Guennou L., et al., 2014, A&A, 566, A149

Guglielmo V., et al., 2018, A&A, 620, A7

Gunn J. E., et al., 2006, AJ, 131, 2332

Guzzo L., et al., 2014, A&A, 566, A108

Halliday C., et al., 2004, A&A, 427, 397

Hansen L., Olsen L. F., Jørgensen H. E., 2002, A&A, 388, 1

Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2009, ApJ, 699, 1333

Hao J., et al., 2010, ApJS, 191, 254

Hasegawa T., et al., 2000, MNRAS, 316, 326

Hashimoto Y., Barcons X., Böhringer H., Fabian A. C., Hasinger G., Mainieri V., Brunner H., 2004, A&A, 417, 819

Hasselfield M., et al., 2013, J. Cosmology Astropart. Phys., 2013, 008

Hennig C., 2017, preprint, (arXiv:1703.09282)

Hennig C., et al., 2017, MNRAS, 467, 4015

Hikage C., Mandelbaum R., Leauthaud A., Rozo E., Rykoff E. S., 2018, MNRAS, 480, 2689

Hilton M., et al., 2020, arXiv e-prints, p. arXiv:2009.11043

Hinshaw G., et al., 2013, ApJS, 208, 19

Hinton S. R., Davis T. M., Lidman C., Glazebrook K., Lewis G. F., 2016, Astronomy and Computing, 15, 61

Hirschmann M., Dolag K., Saro A., Bachmann L., Borgani S., Burkert A., 2014, MNRAS, 442, 2304

Hofmann F., et al., 2017, A&A, 606, A118

Høg E., et al., 2000, A&A, 355, L27

Hogg D. W., 1999, ArXiv Astrophysics e-prints,

Hollowood D. L., et al., 2019, ApJS, 244, 22

Hoshino H., et al., 2015, MNRAS, 452, 998

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Huang S., et al., 2018, PASJ, 70, S6

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huchra J. P., et al., 2012, ApJS, 199, 26

Hunter J. D., 2007, Computing In Science & Engineering, 9, 90

Huterer D., Shafer D. L., 2018, Reports on Progress in Physics, 81, 016901

Ider Chitham J., et al., 2020, MNRAS, 499, 4768

Inami H., et al., 2017, A&A, 608, A2

Ishak M., 2019, Living Reviews in Relativity, 22, 1

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jimeno P., Diego J. M., Broadhurst T., De Martino I., Lazkoz R., 2018, MNRAS, 478, 638

Jones D. H., et al., 2004, MNRAS, 355, 747

Jones D. H., et al., 2009, MNRAS, 399, 683

Jørgensen I., Chiboucas K., Toft S., Bergmann M., Zirm A., Schiavon R. P., Grützbauch R., 2014, AJ, 148, 117

Jørgensen I., Chiboucas K., Berkson E., Smith O., Takamiya M., Villaume A., 2017, AJ, 154, 251

Jullo E., et al., 2019, A&A, 627, A137

Jungman G., Kamionkowski M., Griest K., 1996, Phys. Rep., 267, 195

Käfer F., Finoguenov A., Eckert D., Sanders J. S., Reiprich T. H., Nandra K., 2019, A&A, 628, A43

Kaiser N., 2004, in Oschmann Jr. J. M., ed., Proc. SPIEVol. 5489, Ground-based Telescopes. pp 11–22, doi:10.1117/12.552472

Keisler R., et al., 2011, ApJ, 743, 28

Kellogg E., Baldwin J. R., Koch D., 1975, ApJ, 199, 299

Khabibullina M. L., Verkhodanov O. V., 2009, Astrophysical Bulletin, 64, 340

Khullar G., et al., 2019, ApJ, 870, 7

Kiiveri K., et al., 2021, arXiv e-prints, p. arXiv:2101.02257

Kirby M., Rozo E., Morris R. G., Allen S. W., Costanzi M., Mantz A. B., Rykoff E. S., von der Linden A., 2019, arXiv e-prints, p. arXiv:1910.13548

Kirk B., et al., 2015, MNRAS, 449, 4010

Kirkpatrick C. C., et al., 2021, arXiv e-prints, p. arXiv:2101.04695

Klein M., et al., 2018, MNRAS, 474, 3324

Klein M., et al., 2019, MNRAS, 488, 739

Kluge M., et al., 2020, ApJS, 247, 43

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Koester B. P., et al., 2007a, ApJ, 660, 221

Koester B. P., et al., 2007b, ApJ, 660, 239

Kollmeier J. A., et al., 2017, preprint, (arXiv:1711.03234)

Kolodzig A., Gilfanov M., Hütsi G., Sunyaev R., 2013, A&A, 558, A90

Komatsu E., et al., 2011, ApJS, 192, 18

Koranyi D. M., Geller M. J., 2002, AJ, 123, 100

Koulouridis E., et al., 2016, A&A, 592, A11

Kravtsov A. V., Borgani S., 2012, ARA&A, 50, 353

Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73

Kron R. G., 1980, ApJS, 43, 305

Kuijken K., et al., 2019, A&A, 625, A2

LSST Science Collaboration et al., 2009, preprint, (arXiv:0912.0201)

Lahav O., Liddle A. R., 2019, arXiv e-prints, p. arXiv:1912.03687

Lamer G., Hoeft M., Kohnert J., Schwope A., Storm J., 2008, A&A, 487, L33

Lange R., et al., 2015, MNRAS, 447, 2603

Laureijs R., et al., 2011, arXiv e-prints, p. arXiv:1110.3193

Le Fèvre O., et al., 2005, A&A, 439, 845

Le Fèvre O., et al., 2013, A&A, 559, A14

Lee S.-K., Im M., Hyun M., Park B., Kim J.-W., Kim D., Kim Y., 2019, MNRAS, 490, 135

Lemze D., et al., 2013, ApJ, 776, 91

Lewis I. J., et al., 2002, MNRAS, 333, 279

Lidman C., et al., 2020, MNRAS, 496, 19

Lieu M., et al., 2016, A&A, 592, A4

Lilly S. J., Le Fevre O., Crampton D., Hammer F., Tresse L., 1995, ApJ, 455, 50

Lilly S. J., et al., 2007, ApJS, 172, 70

Lilly S. J., et al., 2009, ApJS, 184, 218

Liske J., et al., 2015, MNRAS, 452, 2087

Liu F. S., Mao S., Meng X. M., 2012, MNRAS, 423, 422

Liu J., et al., 2015, MNRAS, 449, 3370

Liu A., Yu H., Diaferio A., Tozzi P., Hwang H. S., Umetsu K., Okabe N., Yang L.-L., 2018, ApJ, 863, 102

Liu A., et al., 2021, arXiv e-prints, p. arXiv:2106.14518

Lloyd-Davies E. J., et al., 2011, MNRAS, 418, 14

Louis T., et al., 2017, J. Cosmology Astropart. Phys., 2017, 031

Lusso E., et al., 2020, A&A, 642, A150

Lynds R., Petrosian V., 1989, ApJ, 336, 1

Magnier E. A., et al., 2016, preprint, (arXiv:1612.05242)

Mainzer A., et al., 2014, ApJ, 792, 30

Makarov D., Prugniel P., Terekhova N., Courtois H., Vauglin I., 2014, A&A, 570, A13

Malyali A., Rau A., Nandra K., 2019, MNRAS, 489, 5413

Mamon G. A., Biviano A., Boué G., 2013, MNRAS, 429, 3079

Mana A., Giannantonio T., Weller J., Hoyle B., Hütsi G., Sartoris B., 2013, MNRAS, 434, 684

Mancone C. L., Gonzalez A. H., 2012, PASP, 124, 606

Mantz A. B., et al., 2015, MNRAS, 446, 2205

Mantz A. B., et al., 2016, MNRAS, 463, 3582

Mantz A. B., Allen S. W., Morris R. G., von der Linden A., 2018, MNRAS, 473, 3072

Marocco F., et al., 2021, ApJS, 253, 8

Marsh D. J. E., 2016, Phys. Rep., 643, 1

Masters D. C., Stern D. K., Cohen J. G., Capak P. L., Rhodes J. D., Castander F. J., Paltani S., 2017, ApJ, 841, 111

Masters D. C., et al., 2019, ApJ, 877, 81

Mather J. C., Fixsen D. J., Shafer R. A., Mosier C., Wilkinson D. T., 1999, ApJ, 512, 511

Maturi M., Bellagamba F., Radovich M., Roncarelli M., Sereno M., Moscardini L., Bardelli S., Puddu E., 2019, MNRAS, 485, 498

Maughan B. J., Giles P. A., Rines K. J., Diaferio A., Geller M. J., Van Der Pyl N., Bonamente M., 2016, MNRAS, 461, 4182

McClintock T., et al., 2018, preprint, (arXiv:1805.00039)

McClintock T., et al., 2019a, MNRAS, 482, 1352

McClintock T., et al., 2019b, ApJ, 872, 53

McDonald M., et al., 2013, ApJ, 774, 23

McDonald M., et al., 2016, ApJ, 817, 86

McDonald M., et al., 2017, ApJ, 843, 28

McLure R. J., et al., 2013, MNRAS, 428, 1088

McLure R., Pentericci L., VANDELS Team 2017, The Messenger, 167, 31

Meisner A. M., Lang D., Schlegel D. J., 2018, Research Notes of the American Astronomical Society, 2, 1

Melchior P., et al., 2017, MNRAS, 469, 4899

Merloni A., et al., 2012, preprint, (arXiv:1209.3114)

Meusinger H., Rudolf C., Stecklum B., Hoeft M., Mauersberger R., Apai D., 2020, A&A, 640, A30

Migkas K., Schellenberger G., Reiprich T. H., Pacaud F., Ramos-Ceja M. E., Lovisari L., 2020, A&A, 636, A15

Miller C. J., et al., 2005, AJ, 130, 968

Mirkazemi M., et al., 2015, ApJ, 799, 60

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

Momcheva I. G., et al., 2016, ApJS, 225, 27

Moran S. M., Ellis R. S., Treu T., Smail I., Dressler A., Coil A. L., Smith G. P., 2005, ApJ, 634, 977

Moretti A., et al., 2017, A&A, 599, A81

Morris R. A. H., Phillipps S., Jones J. B., Drinkwater M. J., Gregg M. D., Couch W. J., Parker Q. A., Smith R. M., 2007, A&A, 476, 59

Moustakas J., et al., 2021, in American Astronomical Society Meeting Abstracts. p. 527.04

Mpetha C. T., et al., 2021, MNRAS,

Mulroy S. L., et al., 2019, MNRAS, 484, 60

Munari E., Biviano A., Borgani S., Murante G., Fabjan D., 2013, MNRAS, 430, 2638

Murata R., et al., 2019, PASJ, 71, 107

Murata R., Sunayama T., Oguri M., More S., Nishizawa A. J., Nishimichi T., Osato K., 2020, PASJ, 72, 64

Murphy D. N. A., Geach J. E., Bower R. G., 2012, MNRAS, 420, 1861

Muzzin A., et al., 2009, ApJ, 698, 1934

Myles J. T., et al., 2020, arXiv e-prints, p. arXiv:2011.07070

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

Newman J. A., et al., 2013, ApJS, 208, 5

Nidever D. L., et al., 2020, arXiv e-prints, p. arXiv:2011.08868

Nilo Castellón J. L., Alonso M. V., García Lambas D., O'Mill A. L., Valotto C., Carrasco E. R., Cuevas H., Ramírez A., 2014, MNRAS, 437, 2607

Nishimichi T., et al., 2019, ApJ, 884, 29

Nishizawa A. J., et al., 2018, PASJ, 70, S24

Oemler Jr. A., 1974, ApJ, 194, 1

Oguri M., 2014, MNRAS, 444, 147

Oguri M., et al., 2018, PASJ, 70, S20

Old L., et al., 2015, MNRAS, 449, 1897

Old L., et al., 2018, MNRAS, 475, 853

Oliphant T. E., 2006, A guide to NumPy. Vol. 1, Trelgol Publishing USA

Olsen L. F., Hansen L., Jørgensen H. E., Benoist C., da Costa L., Scodeggio M., 2003, A&A, 409, 439

Olsen L. F., Benoist C., da Costa L., Hansen L., Jørgensen H. E., 2005, A&A, 435, 781

Olsen L. F., et al., 2007, A&A, 461, 81

Ondaro-Mallea L., Angulo R. E., Zennaro M., Contreras S., Aricò G., 2021, arXiv e-prints, p. arXiv:2102.08958

Onken C. A., et al., 2019, Publ. Astron. Soc. Australia, 36, e033

Ostriker J. P., Peebles P. J. E., 1973, ApJ, 186, 467

Ostriker J. P., Peebles P. J. E., Yahil A., 1974, ApJ, 193, L1

Owen F. N., Ledlow M. J., Keel W. C., 1995, AJ, 109, 14

Owers M. S., Nulsen P. E. J., Couch W. J., 2011, ApJ, 741, 122

Pacaud F., et al., 2006, MNRAS, 372, 578

Pacaud F., et al., 2016, A&A, 592, A2

Pacaud F., et al., 2018, A&A, 620, A10

Padmanabhan N., Xu X., Eisenstein D. J., Scalzo R., Cuesta A. J., Mehta K. T., Kazin E., 2012, MNRAS, 427, 2132

Palmese A., et al., 2020, MNRAS, 493, 4591

Parkinson D., et al., 2012, Phys. Rev. D, 86, 103518

Particle Data Group et al., 2020, Progress of Theoretical and Experimental Physics, 2020, 083C01

Peebles P. J. E., 1980, The large-scale structure of the universe

Peebles P. J. E., 1982, ApJ, 263, L1

Peebles P. J., Ratra B., 2003, Reviews of Modern Physics, 75, 559

Peloton J., Arnault C., Plaszczynski S., 2018, preprint, (arXiv:1804.07501)

Perlmutter S., et al., 1999, ApJ, 517, 565

Pezzotta A., et al., 2017, A&A, 604, A33

Phriksee A., Jullo E., Limousin M., Shan H., Finoguenov A., Komonjinda S., Wannawichian S., Sawangwit U., 2020, MNRAS, 491, 1643

Pierre M., et al., 2016, A&A, 592, A1

Piffaretti R., Arnaud M., Pratt G. W., Pointecouteau E., Melin J. B., 2011, A&A, 534, A109

Pillepich A., Reiprich T. H., Porciani C., Borm K., Merloni A., 2018, MNRAS, 481, 613

Pimbblet K. A., Smail I., Edge A. C., O'Hely E., Couch W. J., Zabludoff A. I., 2006, MNRAS, 366, 645

Planck Collaboration et al., 2014, A&A, 571, A29

Planck Collaboration et al., 2015a, A&A, 581, A14

Planck Collaboration et al., 2015b, A&A, 582, A29

Planck Collaboration et al., 2016a, A&A, 594, A13

Planck Collaboration et al., 2016b, A&A, 594, A24

Planck Collaboration et al., 2016c, A&A, 594, A26

Planck Collaboration et al., 2016d, A&A, 594, A27

Planck Collaboration et al., 2018, arXiv e-prints,

Planck Collaboration et al., 2020, A&A, 641, A6

Plaszczynski S., Peloton J., Arnault C., Campagne J. E., 2019, Astronomy and Computing, 28, 100305

Pospelov M., Ritz A., Voloshin M., 2008, Physics Letters B, 662, 53

Postman M., Lubin L. M., Gunn J. E., Oke J. B., Hoessel J. G., Schneider D. P., Christensen J. A., 1996, AJ, 111, 615

Postman M., Lubin L. M., Oke J. B., 2001, AJ, 122, 1125

Pranger F., Böhm A., Ferrari C., Maurogordato S., Benoist C., Höller H., Schindler S., 2014, A&A, 570, A40

Pratt G. W., Arnaud M., Biviano A., Eckert D., Ettori S., Nagai D., Okabe N., Reiprich T. H., 2019, Space Sci. Rev., 215, 25

Predehl P., et al., 2020, Nature, 588, 227

Predehl P., et al., 2021, A&A, 647, A1

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

Psychogyios A., et al., 2020, A&A, 633, A104

Raghunathan S., et al., 2019, ApJ, 872, 170

Raichoor A., et al., 2021, MNRAS, 500, 3254

Ramella M., Boschin W., Fadda D., Nonino M., 2001, A&A, 368, 776

Randall S. W., Markevitch M., Clowe D., Gonzalez A. H., Bradač M., 2008, ApJ, 679, 1173

Reichardt C. L., et al., 2012, ApJ, 755, 70

Reiprich T. H., Böhringer H., 2002, ApJ, 567, 716

Rescigno U., et al., 2020, A&A, 635, A98

Ricci M., et al., 2018, preprint, (arXiv:1807.03207)

Richard J., et al., 2021, A&A, 646, A83

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 2011, ApJ, 730, 119

Riess A. G., et al., 2016, ApJ, 826, 56

Riess A. G., et al., 2018, ApJ, 861, 126

Rines K., Geller M. J., Diaferio A., Mohr J. J., Wegner G. A., 2000, AJ, 120, 2338

Rines K. J., Geller M. J., Diaferio A., Hwang H. S., 2016, ApJ, 819, 63

Rines K. J., Geller M. J., Diaferio A., Hwang H. S., Sohn J., 2018, ApJ, 862, 172

Risaliti G., Lusso E., 2019, Nature Astronomy, 3, 272

Roberts M. S., Rots A. H., 1973, A&A, 26, 483

Robotham A. S. G., et al., 2011, MNRAS, 416, 2640

Rosenberg L. J., 2015, Proceedings of the National Academy of Sciences, 112, 12278

Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, MNRAS, 449, 835

Ross A. J., et al., 2020, MNRAS, 498, 2354

Rozo E., Rykoff E. S., 2014, ApJ, 783, 80

Rozo E., et al., 2010, ApJ, 708, 645

Rubin V. C., Ford W. Kent J., 1970, ApJ, 159, 379

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Rudnick G., et al., 2017, ApJ, 850, 181

Ruel J., et al., 2014, ApJ, 792, 45

Runge J., Yan H., 2018, ApJ, 853, 47

Rykoff E. S., et al., 2012, ApJ, 746, 178

Rykoff E. S., et al., 2014, ApJ, 785, 104

Rykoff E. S., Rozo E., Keisler R., 2015, preprint, (arXiv:1509.00870)

Rykoff E. S., et al., 2016, ApJS, 224, 1

Sadibekova T., Pierre M., Clerc N., Faccioli L., Gastaud R., Le Fevre J.-P., Rozo E., Rykoff E., 2014, A&A, 571, A87

Salvati L., Douspis M., Aghanim N., 2020, arXiv e-prints, p. arXiv:2005.10204

Salvato M., Ilbert O., Hoyle B., 2019, Nature Astronomy, 3, 212

Saro A., Mohr J. J., Bazin G., Dolag K., 2013, ApJ, 772, 47

Saro A., et al., 2015, MNRAS, 454, 2305

Schechter P., 1976, ApJ, 203, 297

Schellenberger G., Reiprich T. H., 2017, MNRAS, 471, 1370

Schirmer M., Hildebrandt H., Kuijken K., Erben T., 2011, A&A, 532, A57

Schrabback T., et al., 2018, MNRAS, 474, 2635

Scolnic D. M., et al., 2018, ApJ, 859, 101

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sersic J. L., 1968, Atlas de Galaxias Australes

Sharp R., et al., 2006, in McLean I. S., Iye M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. p. 62690G (arXiv:astro-ph/0606137), doi:10.1117/12.671022

Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Lin H., Kirshner R. P., Schechter P. L., 1996, ApJ, 470, 172

Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978

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

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sifón C., et al., 2013, ApJ, 772, 25

Silva D. R., et al., 2016, in American Astronomical Society Meeting Abstracts #228. p. 317.02

Silverman J. D., et al., 2015, ApJS, 220, 12

Skelton R. E., et al., 2014, ApJS, 214, 24

Sluse D., et al., 2019, MNRAS, 490, 613

Smee S. A., et al., 2013, AJ, 146, 32

Smith S., 1936, ApJ, 83, 23

Smith R. J., Lucey J. R., Hudson M. J., Schlegel D. J., Davies R. L., 2000, MNRAS, 313, 469

Soares-Santos M., et al., 2011, ApJ, 727, 45

Sohn J., Chon G., Böhringer H., Geller M. J., Diaferio A., Hwang H. S., Utsumi Y., Rines K. J., 2018a, ApJ, 855, 100

Sohn J., Geller M. J., Rines K. J., Hwang H. S., Utsumi Y., Diaferio A., 2018b, ApJ, 856, 172

Sohn J., Geller M. J., Zahid H. J., Fabricant D. G., 2019a, ApJ, 872, 192

Sohn J., Geller M. J., Zahid H. J., 2019b, ApJ, 880, 142

Sohn J., Geller M. J., Hwang H. S., Fabricant D. G., Moran S. M., Utsumi Y., 2020, arXiv e-prints, p. arXiv:2010.05817

Song H., Hwang H. S., Park C., Tamura T., 2017, ApJ, 842, 88

Soucail G., Mellier Y., Fort B., Mathez G., Cailloux M., 1988, A&A, 191, L19

Springel V., 2005, MNRAS, 364, 1105

Stalder B., et al., 2013, ApJ, 763, 93

Stanford S. A., Gonzalez A. H., Brodwin M., Gettings D. P., Eisenhardt P. R. M., Stern D., Wylezalek D., 2014, ApJS, 213, 25

Story K. T., et al., 2013, ApJ, 779, 86

Stott J. P., Edge A. C., Smith G. P., Swinbank A. M., Ebeling H., 2008, MNRAS, 384, 1502

Straatman C. M. S., et al., 2018, ApJS, 239, 27

Strait V., et al., 2018, ApJ, 868, 129

Strazzullo V., et al., 2019, A&A, 622, A117

Streblyanska A., Aguado-Barahona A., Ferragamo A., Barrena R., Rubiño-Martín J. A., Tramonte D., Genova-Santos R. T., Lietzen H., 2019, A&A, 628, A13

Sunyaev R. A., Zeldovich Y. B., 1970, Ap&SS, 7, 3

Sunyaev R. A., Zeldovich Y. B., 1972, Comments on Astrophysics and Space Physics, 4, 173

Suzuki N., et al., 2012, ApJ, 746, 85

Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391

Szabo T., Pierpaoli E., Dong F., Pipino A., Gunn J., 2011, ApJ, 736, 21

Szokoly G. P., et al., 2004, ApJS, 155, 271

Tanaka M., Lidman C., Bower R. G., Demarco R., Finoguenov A., Kodama T., Nakata F., Rosati P., 2009, A&A, 507, 671

Tarrío P., Melin J.-B., Arnaud M., Pratt G. W., 2016, A&A, 591, A39

Tarrío P., Melin J.-B., Arnaud M., 2018, A&A, 614, A82

Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV. p. 29

Team T. P. D., 2020, pandas-dev/pandas: Pandas, doi:10.5281/zenodo.3509134, https://doi.org/10.5281/zenodo.3509134

Tempel E., Kruuse M., Kipper R., Tuvikene T., Sorce J. G., Stoica R. S., 2018, A&A, 618, A81

The Dark Energy Survey Collaboration 2005, arXiv e-prints, pp astro–ph/0510346

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Tinker J. L., et al., 2012, ApJ, 745, 16

Tonry J. L., et al., 2012, ApJ, 750, 99

Tran K.-V. H., Franx M., Illingworth G. D., van Dokkum P., Kelson D. D., Blakeslee J. P., Postman M., 2007, ApJ, 661, 750

Trimble V., 1987, ARA&A, 25, 425

Truemper J., 1993, Science, 260, 1769

Tukey J. W., 1958, Ann. Math. Statist., 29, 581

Turner M. S., 1999, in Gibson B. K., Axelrod R. S., Putman M. E., eds, Astronomical Society of the Pacific Conference Series Vol. 165, The Third Stromlo Symposium: The Galactic Halo. p. 431 (arXiv:astro-ph/9811454)

Umetsu K., 2020, arXiv e-prints, p. arXiv:2007.00506

Vanderlinde K., et al., 2010, ApJ, 722, 1180

Verde L., Treu T., Riess A. G., 2019, Nature Astronomy, 3, 891

Verdugo M., Ziegler B. L., Gerken B., 2008, A&A, 486, 9

Vikhlinin A., et al., 2009a, ApJ, 692, 1033

Vikhlinin A., et al., 2009b, ApJ, 692, 1060

Virtanen P., et al., 2019, arXiv e-prints, p. arXiv:1907.10121

Voges W., et al., 1999, A&A, 349, 389

Wainer H., Thissen D., 1976, Psychometrika, 41, 9

Waters C. Z., et al., 2016, preprint, (arXiv:1612.05245)

Watson W. A., Iliev I. T., D'Aloisio A., Knebe A., Shapiro P. R., Yepes G., 2013, MNRAS, 433, 1230

Weinberg S., 1989, Reviews of Modern Physics, 61, 1

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

Wen Z. L., Han J. L., 2018, MNRAS, 481, 4158

Wen Z. L., Han J. L., Liu F. S., 2010, MNRAS, 407, 533

Wen Z. L., Han J. L., Liu F. S., 2012, ApJS, 199, 34

Will C. M., 2014, Living Reviews in Relativity, 17, 4

Willis J. P., et al., 2013, MNRAS, 430, 134

Wilson G., et al., 2009, ApJ, 698, 1943

Wilson M. L., Zabludoff A. I., Ammons S. M., Momcheva I. G., Williams K. A., Keeton C. R., 2016, ApJ, 833, 194

Wojtak R., Hansen S. H., Hjorth J., 2011, Nature, 477, 567

Wolf J., et al., 2021, A&A, 647, A5

Yang M., et al., 2018, ApJS, 234, 5

Yee H. K. C., et al., 2000, ApJS, 129, 475

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

Yu H., Tozzi P., Borgani S., Rosati P., Zhu Z. H., 2011, A&A, 529, A65

Yu H., Serra A. L., Diaferio A., Baldi M., 2015, ApJ, 810, 37

Zarrouk P., et al., 2018, MNRAS, 477, 1639

Zaznobin I. A., et al., 2020, Astronomy Letters, 46, 79

Zeimann G. R., et al., 2013, ApJ, 779, 137

Zel'dovich Y. B., 1968, Soviet Physics Uspekhi, 11, 381

Zentner A. R., 2007, International Journal of Modern Physics D, 16, 763

Zhang Y.-Y., et al., 2017, A&A, 599, A138

Zhang Y., et al., 2019, MNRAS, 487, 2578

Zimdahl W., Pavón D., Chimento L. P., 2001, Physics Letters B, 521, 133

Zou H., et al., 2017, PASP, 129, 064101

Zou H., Gao J., Zhou X., Kong X., 2019, ApJS, 242, 8

Zu Y., Weinberg D. H., Rozo E., Sheldon E. S., Tinker J. L., Becker M. R., 2014, MNRAS, 439, 1628

Zubeldia Í., Challinor A., 2019, MNRAS, 489, 401

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

de Haan T., et al., 2016, ApJ, 832, 95

de Jong R. S., 2019, Nature Astronomy, 3, 574

de Jong R. S., et al., 2012, in Ground-based and Airborne Instrumentation for Astronomy IV. p. 84460T (arXiv:1206.6885), doi:10.1117/12.926239

de Jong R. S., et al., 2019, The Messenger, 175, 3

de la Torre S., et al., 2013, A&A, 557, A54

eBOSS Collaboration et al., 2020, arXiv e-prints, p. arXiv:2007.08991

van der Walt S., Colbert S. C., Varoquaux G., 2011, Computing in Science and Engineering, 13, 22

von der Linden A., et al., 2014, MNRAS, 439, 2