

---

# **Cosmological Studies with Galaxy Clusters at X-ray, optical and millimeter wavelengths**

**Sebastian Grandis**

---



München 2020



---

# **Cosmological Studies with Galaxy Clusters at X-ray, optical and millimeter wavelengths**

**Sebastian Grandis**

---

Dissertation  
an der Fakultät für Physik  
der Ludwig-Maximilians-Universität  
München

vorgelegt von  
Sebastian Grandis  
aus Rom, Italien

München, den 31.10.2019

Erstgutachter: Prof. Joseph J. Mohr

Zweitgutachter: Prof. Simon D. M. White

Tag der mündlichen Prüfung: 17.12.2019

# Inhaltsverzeichnis

<b>Danksagung</b>	<b>ix</b>
<b>Zusammenfassung</b>	<b>x</b>
<b>Abstract</b>	<b>xii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Formation of Halos . . . . .	1
1.2 The Halo Mass Function and its Cosmological Dependence . . . . .	2
1.3 Observational Signatures of Galaxy Clusters . . . . .	3
1.3.1 X-ray . . . . .	4
1.3.2 Optical and near infrared . . . . .	5
1.3.3 Millimeter regime . . . . .	6
1.4 Selection of Cluster Samples . . . . .	7
1.5 Cosmological Inference from Cluster Samples . . . . .	9
1.5.1 Observable-Mass Relation . . . . .	9
1.5.2 Predicted Number of Objects . . . . .	10
1.5.3 Mass Calibration . . . . .	12
1.5.4 Fitting . . . . .	13
1.6 Outline . . . . .	14
1.6.1 Application 1 . . . . .	14
1.6.2 Application 2 . . . . .	15
1.6.3 Application 3 . . . . .	15
<b>2 WL Mass Calibration of eROSITA Cluster Counts</b>	<b>17</b>
2.1 Experimental setup . . . . .	17
2.1.1 Creating the mock cluster catalog . . . . .	19
2.1.2 Forecasting the WL signal . . . . .	22
2.1.3 Fiducial cosmology and scaling relations . . . . .	26
2.2 Cosmology analysis method . . . . .	26
2.2.1 Cluster selection scaling relation . . . . .	27
2.2.2 Likelihood functions . . . . .	28
2.2.3 Comments on sampling and model choice . . . . .	30

---

2.2.4	Choice of priors . . . . .	31
2.3	Results . . . . .	36
2.3.1	Optimal mass calibration . . . . .	37
2.3.2	Forecasts: eROSITA+WL . . . . .	38
2.3.3	Synergies with Planck CMB . . . . .	42
2.3.4	Synergies with DESI BAO measurements . . . . .	45
2.3.5	Combining all datasets . . . . .	47
2.3.6	Inclusion of low mass clusters and groups . . . . .	48
2.4	Discussion . . . . .	50
2.4.1	Impact of WL on scaling relation parameters . . . . .	51
2.4.2	Parameter sensitivities . . . . .	52
2.4.3	Comparison to previous work . . . . .	56
2.5	Conclusions . . . . .	59
2.6	Appendix . . . . .	62
2.6.1	Comments on selection . . . . .	62
2.6.2	Comments on the count rate to mass relation . . . . .	64
2.6.3	Scaling relation form . . . . .	67
<b>3</b>	<b>Cluster Sample Validation</b>	<b>69</b>
3.1	Conceptual Framework for Cluster Cosmology Analyses . . . . .	69
3.1.1	Modeling the cluster population . . . . .	70
3.1.2	Modeling measurement uncertainties . . . . .	71
3.1.3	Modeling selection functions . . . . .	72
3.2	Validation methods . . . . .	77
3.2.1	SPT-SZ cross calibration . . . . .	78
3.2.2	Calibration with number counts . . . . .	79
3.2.3	Consistency check using two cluster samples . . . . .	80
3.3	Dataset and Priors . . . . .	85
3.3.1	Cluster samples . . . . .	85
3.3.2	Priors . . . . .	88
3.4	Application to MARD-Y3 and SPT-SZ . . . . .	90
3.4.1	Validation using SPT-SZ cross calibration . . . . .	94
3.4.2	Validation using number counts . . . . .	95
3.4.3	Validation using cluster masses . . . . .	100
3.4.4	Validation using independent cluster samples . . . . .	103
3.5	Discussion . . . . .	109
3.5.1	X-ray selection function systematics . . . . .	109
3.5.2	Outcome of the validation . . . . .	111
3.5.3	Impact of the optical incompleteness . . . . .	113
3.5.4	Implications for cosmological studies . . . . .	114
3.6	Conclusions . . . . .	116
3.7	Appendix . . . . .	118
3.7.1	X-ray flux error model . . . . .	118

3.7.2	Gallery of Multi-wavelength Cluster Images	118
<b>4</b>	<b>DES-Y1 cross calibration with SPT</b>	<b>123</b>
4.1	Cluster samples	123
4.1.1	Optically selected samples	123
4.1.2	SPT matched sample	124
4.2	Methods	126
4.2.1	General cluster population model	126
4.2.2	Observation error models	128
4.2.3	SPT cross calibration	128
4.2.4	Constraining outlier fractions and purity	129
4.2.5	Predictions	131
4.2.6	Priors	135
4.3	Results	135
4.3.1	Cross Calibration	135
4.3.2	Comparison to $\hat{\lambda} > 20$	139
4.4	Discussion	142
4.4.1	Comparison to literature	142
4.4.2	Inferred Purity	145
4.4.3	Alternative Explanations	147
4.4.4	Prospects	148
4.5	Conclusions	149
4.6	Appendix	150
4.6.1	Optical completeness of the SPT sample	150
<b>5</b>	<b>Conclusion</b>	<b>153</b>



*s. d. g.*

x

Zusammenfassung

# **Zusammenfassung:**

## **Kosmologische Untersuchungen mit Galaxienhaufen in Röntgen-, optischen und Millimeterwellenlängen**

Die Anzahl an Halos als Funktion der Masse und Rotverschiebung enthält viel kosmologische Information. Die massenreichsten Halos sind von Galaxienhaufen bevölkert, deren beobachtbare Eigenschaften mit der Masse und Rotverschiebung des beherbergenden Halos skalieren. Diese Eigenschaften erlauben außerdem die Selektion von Galaxienhaufen anhand von Erhebungen in Röntgen-, optischen und Millimeterwellenlängen. In dieser Promotionsarbeit stellen wir unter Beweis, wie man kosmologische Information aus Galaxienhaufenkatalogen extrahiert. Die wichtigsten begrenzenden Faktoren dieser Messung sind die Unsicherheit im Verhältnis zwischen den beobachteten Eigenschaften und der Halomasse und die Unsicherheit in der Modellierung der Selektion. Durch die Einführung neuer und der Weiterentwicklung etablierter Untersuchungsmethoden zeigen wir auf, wie diese Unsicherheiten empirisch kalibriert werden können. Des Weiteren stellen wir unter Beweis, wie sich empirische Validierung für die Rückschlussfolgerung kosmologischer Information aus Galaxienhaufenkatalogen durchführen lassen.



# Abstract

The number of halos as a function of mass and redshift is a powerful cosmological probe. The most massive halos are inhabited by clusters of galaxies, whose observational features scale with the host's halo mass and redshift with some scatter. These features allow us to select galaxy clusters in X-ray, optical and millimeter wavelength. We demonstrate in this thesis how to extract cosmological information from a cluster sample. The major limiting factors to this measurement are the uncertainty in the mapping between observable and mass, and the uncertainties in the modelling of the selection function. We demonstrate, introducing novel techniques and developing established ones, how to empirically calibrate these sources of systematic uncertainty. We furthermore demonstrate how to set up empirical validation tests for the cosmological inference from cluster samples.



# Kapitel 1

## Introduction

This thesis focuses on cosmological studies with galaxy clusters in the different wavelengths in which we find predominant observational signatures of these objects. Before delving into the scientific results we obtained, we aim to shortly review the following points:

- What are galaxy clusters? How many are there at any given cosmic epoch? And how does this number of clusters depend on cosmology?
- What are the most pronounced observational features of galaxy clusters? How can large samples of clusters be selected?
- How can one extract information on the composition and evolution of the Universe from cluster samples?

We caution the reader here that we do not intend to discuss all of the answers in exhaustive detail. Nevertheless, references are provided for more complete and detailed treatments. We will also assume a certain degree of familiarity with the field of observational cosmology. Therefore, we will not review the basic physical concepts of the current cosmological model.

### 1.1 Formation of Halos

While the distribution of matter in the Universe is assumed to be homogeneous and isotropic at cosmic scales, on a local level inhomogeneities are the rule rather than the exception. This can indeed be appreciated by the simple observation that the mean density around us on Earth is many orders of magnitude larger than the average density of the Cosmos. The formation of structure in the Universe is however a natural consequence of even small initial inhomogeneities. Under the attraction of their own gravity, and under the right pressure conditions, small over-densities increase in density. The initial only minimal inhomogeneous matter distribution arranges itself under the influence of its own gravitational field into a web like structure: the aptly named Cosmic Web. At the intersection of the

filamentary structure constituting the Cosmic Web, purely gravitational dissipative processes are strong enough to lead to the formation of approximately round, bound structures called halos (Lynden-Bell, 1967).

The number density of halos as a function of halo mass at any given time is called the halo mass function. Its main features are described by the model of gravitational collapse, which we will quickly review in a qualitative fashion following Press and Schechter (1974). The starting point of the gravitational collapse model is the fact that the gravitational potential on the surface of any reasonably defined volume depends exclusively on the matter contained within that volume. Considering now a spherical volume enclosing an over-dense region, the collapse of this region will decouple it from the background expansion of the Universe. In this region, space-time no longer expands like in the rest of the Universe. Instead it contracts. Classically speaking, the potential well deepens as more and more matter is accreted. The fractional over-density of the collapsing region increases with time as it attracts more and more matter, and as the surrounding Universe expands and becomes less and less dense.

At an over-density of  $\sim 200$  times the average density of the Universe the collapsing object attains *Virial Equilibrium*. Further collapse is halted purely by gravitational dissipation, such as violent relaxation (particles lose energy when falling into a quickly deepening gravitational potential) or dynamical friction (the gravitational attraction of a passing particle accumulates matter behind, which in turn slows it down). A halo forms.

In virial equilibrium the kinetic energy  $K$  of the system and gravitational binding energy  $V$  follow the relation

$$2K + V = 0. \quad (1.1)$$

The setting of a characteristic over-density  $\Delta$  for a halo allows one to define its mass  $M_\Delta$  as the mass of a sphere with  $\Delta$ -times the average density of the Universe. This in turn also defines a typical size of the halo, the radius of the aforementioned sphere. Thus

$$M_\Delta = \frac{4\pi}{3} \Delta \rho R_\Delta^3, \quad (1.2)$$

where  $\rho$  is the average density of the Universe. Specific application vary in the choice of over-density (176, 200 or 500) and the density to be used as the average density of the Universe. Some works employing the critical density  $\rho_{\text{crit}}(z) = \frac{8\pi G}{3} H^2(z)$  of the Universe at the epoch determined by the cosmological redshift  $z$ , while other use the matter density  $\rho_M = \Omega_M \rho_{\text{crit}}(z)$ . We can note here for later use that  $R \sim M^{1/3} E(z)^{-2/3}$ . Here  $H(z)$  is the expansion rate at the redshift  $z$  and  $E(z) = H(z)/H_0$  the expansion at redshift  $z$  relative to the current day expansion rate.

## 1.2 The Halo Mass Function and its Cosmological Dependence

The concept of spherical over-density mass can be used to identify halos in large N-body simulations performed with different cosmological models. From such simulations the dif-

ferential number density  $\frac{dn}{dM}\Big|_{M,z}$  of halos at a mass  $M$  in a cosmic epoch given by the redshift  $z$  is calibrated. Until the recent advent of emulators (McClintock et al., 2019b), these calibrations were performed in the spirit of the Press-Schechter formalism (Jenkins et al., 2001; Evrard et al., 2002; Tinker et al., 2008; Bocquet et al., 2016, and others). This formalism builds on the spherical collapse model. It computes the time that passes between two events: 1) the over-density starts to collapse under its own gravitational influence and no longer expands like the surrounding space, and 2) the object reaches virialization. Given this time, one can estimate how much the perturbation would have grown, if it had continued linear growth. The resulting over-density is called barrier over-density. Given this set-up, the number of virialized objects equals the number of linear over-densities that passed the barrier. Linear growth is characterized by the variance  $\sigma^2(R)$  of the over-densities around the mean over-density of zero, and follows by numerical integration of the Einstein-Boltzmann equations for any set of cosmological parameters. The fraction of over-densities that exceed the barrier can thus be easily computed if the variance is known, as can the number density of such over-densities. The link to mass is finally made by choosing the linear radius  $R$  such that it encloses a mass  $M$ .

The shape of the resulting function is rather simple: at low masses it follows a power law and it has an exponential cut off. The cut-off mass increases the later the cosmic epoch, as more and more massive objects form. Also the amplitude of the function increases with time, as the general growth of the amplitude of fluctuations pushes ever more over-densities past the barrier. As such, the halo mass function is sensitive to the amplitude of matter fluctuations and its growth, as well as to the overall density of the Universe. Finally, further cosmological sensitivity is introduced by considering the differential number of objects as a function of mass  $M$  and redshift  $z$  in a survey of solid angle  $A_{\text{survey}}$ , which can be computed

$$\frac{dN}{dM}\Big|_{M,z} = \frac{dn}{dM}\Big|_{M,z} \frac{d^3V}{dzd^2\theta}\Big|_z A_{\text{survey}}, \quad (1.3)$$

where  $\frac{d^3V}{dzd^2\theta}\Big|_z$  is the cosmological volume element. It carries strong cosmological dependence on the expansion history of the Universe (see for instance Albrecht et al., 2006). In summary then, the number of galaxy clusters probes the high tail of the cosmic inhomogeneity distribution at different scales, the growth of these inhomogeneities with redshift and the expansion of the Universe itself (for instance, see Koester et al., 2007; Vikhlinin et al., 2009b; Mantz et al., 2010; Rozo et al., 2010; Benson et al., 2013; Mantz et al., 2015; Bocquet et al., 2015; Planck Collaboration et al., 2016b; de Haan et al., 2016; Bocquet et al., 2019a).

### 1.3 Observational Signatures of Galaxy Clusters

At different masses, halos are occupied by different astrophysical objects. Halos above a mass of  $M \sim 10^{14} M_\odot$  are occupied by galaxy clusters, while lower mass halos are occupied by galaxy groups and individual galaxies. Galaxy clusters set themselves apart from other

astrophysical objects by the fact that their different observational features are dominated, to first order, by the depth of the gravitational potential (Kravtsov and Borgani, 2012, for a review). Consequently, the strength of their signatures in X-rays, in optical, near infrared and in sub-millimeter wavelengths correlates with mass (see Pratt et al., 2019, and references therein), as discussed below in more detail. Calibrating this relation between the observable and the halo mass, and the scatter around it, then allows one to constrain the number of halos as a function of mass and redshift (see Allen, Evrard, and Mantz, 2011, for a review). This is, as discussed above, a powerful cosmological probe. The major observational signatures at various wavelength that are used for the identification and mass estimation shall be reviewed in the following. Besides presenting the observational features most relevant to this work, we also derive their scaling with mass following self-similar collapse theory (Kaiser, 1986).

### 1.3.1 X-ray

Galaxy clusters are associated with deep potential wells as they live in the most massive halos. These deep potential wells cause the baryonic matter falling into them to heat to temperatures above  $T \sim 2$  keV. The resulting hot plasma, called inter cluster medium (ICM), emits thermally with a bremsstrahlung spectrum (for reviews, see Sarazin, 1988; Böhringer and Werner, 2010). The temperature of this emission is a direct tracer of the gravitational potentials depth. Its scaling with mass can be estimated from the virial theorem, by noting that  $V \sim M^2/R \sim M^{2/3}E(z)^{2/3}$ . In a hot thermalized plasma, the kinetic energy is proportional to the temperature, leading to the scaling

$$T \sim M^{2/3}E(z)^{2/3} \quad (1.4)$$

Observationally, the temperature of the gas is costly to extract, as it requires high photon counts and a good understanding of the background to perform a spectral fit. Much easier to access is the X-ray surface brightness. Given the collisional emission, the surface brightness  $S_X$  (in units of  $\text{erg s}^{-1}\text{cm}^{-2}\text{steradian}^{-1}$ ) reads

$$S_X = \frac{1}{4\pi(1+z)^4} \int dl n_e n_H \Lambda_{\text{eV}}(T, Z), \quad (1.5)$$

where  $n_e$  is the number density of free electrons,  $n_H$  the number density of ionized hydrogen,  $\Lambda_{\text{eV}}(T, Z)$  the emissivity in a chosen band as a function of temperature  $T$  and ICM metallicity  $Z$ . The integral goes over the line of sight. We shall focus here on soft X-ray bands (either  $[0.1, 2.4]\text{keV}$  or  $[0.5, 2.]\text{keV}$ ), where the emissivity is temperature independent to good approximation.

De-projecting or parametric fitting of the surface brightness profile allows one to determine the electron density profile. Integrating the electron number density out to the radius  $R_\Delta$  provides an estimate of the ICM gas mass  $M_{\Delta, \text{gas}}$ . Under the simplistic assumption that the ratio between the ICM mass and total mass follows the cosmic baryon fraction,

we arrive at the scaling

$$M_{\text{gas}} \sim M. \quad (1.6)$$

Deviations from this scaling can be interpreted as mass and redshift dependent deviations from the assumption that the baryon fraction in clusters corresponds to the cosmic fraction. Several interesting physical effects such as AGN feedback and star formation are already recognized today as playing a major role in setting the mass trend of the gas mass (Tozzi and Norman, 2001; Borgani et al., 2004; Kravtsov, Nagai, and Vikhlinin, 2005; Nagai, Kravtsov, and Vikhlinin, 2007).

Integration of the surface brightness profile in an angular aperture corresponding to  $R_\Delta$  provides the X-ray flux  $f_{\Delta, \text{X}}$  within that radius. In soft X-rays, the emissivity is independent of temperature, and thus the flux scales as  $f_{\text{X}} \sim \rho_{\text{gas}}^2 R^3 d_{\text{L}}(z)^{-2}$ , where  $d_{\text{L}}(z)$  is the luminosity distance. This leads to the scaling

$$f_{\text{X}} \sim M E(z)^2 d_{\text{L}}(z)^{-2}. \quad (1.7)$$

As in the case of the gas mass, deviations from this scaling are indicative of astrophysical processes beyond gravitational collapse and adiabatic hydrodynamics. Most prominent in the case of the flux are cooling effects in the core of clusters (Tozzi and Norman, 2001; Borgani et al., 2004; Kravtsov, Nagai, and Vikhlinin, 2005; Nagai, Kravtsov, and Vikhlinin, 2007).

At an even simpler observational level as the flux  $f_{\text{X}}$  is the count rate  $\eta$ . The conversion from flux to count rate depends on the spectral sensitivity of the instrument, usually called effective area or ancillary response function (ARF). While this introduces a joint redshift temperature dependence, as we will see in a specific application later, at first order  $\eta \sim f_{\text{X}}$ .

### 1.3.2 Optical and near infrared

In the optical and near infrared regime clusters appear as over-densities of predominantly red galaxies. The most distinct features in these wavelengths are the Brightest Central Galaxy (BCG) and the presence of a red sequence (RS) of galaxy colors (Dressler, 1984). While the BCG is of central importance in the study of the astrophysics of clusters, the presence of the red sequence is pivotal in cosmological studies.

Several physical processes suppress star formation in galaxies in clusters. As a result, quiescent galaxies occupy the red sequence, which sets itself apart through its narrow color range as a function of redshift. Calibrating this relation as a redshift dependent color filter on spectroscopic data allows one to filter photometric data. Adding a spatial filter, over-densities of galaxies with colors consistent with a given redshift can be selected. This technique provides a photometric redshift estimate, necessary for cosmological studies of clusters. The amplitude in the color-spatial filter is called richness  $\lambda$  (Rozo et al., 2009). It roughly scales like the number of galaxies in the cluster above a given magnitude. Assuming a mean stellar mass  $m_\star$  per galaxy, and a total stellar mass in the cluster  $M_\star$ , under the

simplification that the star mass fraction in clusters is constant with mass and redshift, we find the scaling

$$\lambda \sim \frac{M_\star}{m_\star} \sim M. \quad (1.8)$$

Another important observational signature of galaxy clusters which is extracted from optical photometric data is the weak gravitational lensing (WL) distortion of the shapes of background galaxies by the gravitational potential of the clusters (for a review, see Hoekstra et al., 2013). With the exception of the central region of the most massive systems, the WL distortion is of the order of a few percent. This signal has to be extracted from galaxies which intrinsically have a dispersion in ellipticities of 25%–30%. Given that the intrinsic ellipticity follows to good approximation a Gaussian random field, statistically significant WL shear signals can be extracted by radially binning large numbers of galaxy shapes, resulting in a reduced tangential shear profile  $g_t(\theta_i)$  in radial bins. The reduced tangential shear can be expressed as

$$g_t(\theta) = \frac{\gamma(\theta)}{1 - \kappa(\theta)}, \quad (1.9)$$

where  $\gamma(\theta)$  is the tangential shear and  $\kappa(\theta)$  is the convergence. Given a cylindrical projection of the matter profile  $\Sigma(R)$  both the convergence and the shear can be estimated as

$$\kappa(\theta) = \frac{\Sigma(d_A(z)\theta)}{\langle \Sigma_{\text{crit}} \rangle} \text{ and } \gamma(\theta) = \frac{\Sigma(< d_A(z)\theta) - \Sigma(d_A(z)\theta)}{\langle \Sigma_{\text{crit}} \rangle}, \quad (1.10)$$

where  $d_A(z)$  is the angular diameter distance to the cluster redshift  $z$ , and  $\langle \Sigma_{\text{crit}} \rangle$  takes account of the lens configuration between the cluster and the background sources. It therefore requires knowledge of the redshift distribution of the background galaxies.

By fitting a mass profile to the measured reduced shear, direct mass estimates can be extracted. This is made possible by the fact that the density profile of clusters follows the Navarro-Frenk-White profile (Navarro, Frenk, and White, 1996), which can be parametrized as

$$\rho(r) = \frac{\rho_0}{\frac{r}{r_s} \left(1 + \frac{r}{r_s}\right)^2}, \quad (1.11)$$

where  $\rho_0$  and  $r_s$  are free parameters. Alternatively the profile can be parameterized by the total mass  $M_\Delta$  and the concentration  $c_\Delta = R_\Delta/r_s$ . Fitting for the mass under some assumptions of the concentration provides a mass estimate called weak lensing mass  $M_{\text{WL}}$ . While it traces the halo mass, it displays bias and scatter with respect to it due to morphological variety and correlated large scale structure (Becker and Kravtsov, 2011).

### 1.3.3 Millimeter regime

Galaxy clusters have a distinct signature in the millimeter and sub-millimeter regime called the Sunyaev-Zel'dovich effect (SZe, Sunyaev and Zeldovich, 1972): Cosmic Microwave

background (CMB) photons travelling through the ICM get inverse Compton scattered to higher frequencies, leading to a frequency dependent change in the CMB temperature

$$\frac{\Delta T}{T_{\text{CMB}}} = G(\nu)y_c = G(\nu) \frac{k_B \sigma_T}{m_e c^2} \int dl n_e T_e, \quad (1.12)$$

where  $G(\nu)$  encodes the distinctive spectral feature as a function of frequency  $\nu$ , and  $y_c$  is the Comptonization parameter. Furthermore,  $k_B$  is the Boltzmann constant,  $\sigma_T$  the Thomson cross-section,  $m_e$  the electron mass, and  $c$  the speed of light.

Of interest for this work is the unbiased signal to noise  $\zeta$  with which a cluster is extracted from a CMB map. An extraction aperture  $\theta_{\text{ext}}$  is selected by maximising the Comptonization  $Y_{\text{SZ}}(> \theta_{\text{ext}})$  within that aperture with respect to the instrumental and CMB noise  $\sigma_{\text{mm}}$ . This leads to the definition

$$\zeta = \frac{Y_{\text{SZ}}(> \theta_{\text{ext}})}{\sigma_{\text{mm}} \pi \theta_{\text{ext}}^2}. \quad (1.13)$$

Assuming that the extraction aperture corresponds to the angular scale of the radius  $R_\Delta$ , we can derive the self-similar scaling

$$\zeta \sim M E^2 \quad (1.14)$$

Practically, a multi-wavelength filter that compensates the typical scale of CMB fluctuations is employed. In the low signal to noise regime in which most clusters are detected, the extraction area is a very noisy tracer of the angular scale of the clusters extent.

## 1.4 Selection of Cluster Samples

The main observational features of clusters discussed above allow one to select cluster samples based on these signatures. In the following section, we shall quickly discuss the complications arising from the selection in either observing band.

Two major concerns arise in cluster selection: the completeness of the sample, and the contamination of the sample. As will be discussed in more detail below, cluster cosmological studies assume that for any halo with given mass and redshift there is a range of possible observables. The mean observable is given by a functional form of mass, redshift and cosmology, while the scatter around that relation captures both the instrumental and observational uncertainties of the observables measurement, as well as the intrinsic heterogeneity of the cluster population. Both the mean relation and the scatter around it need to be empirically calibrated. In light of this framework, if a sample is selected above a given cut in observable, its mass incompleteness can be easily modelled by considering the distribution of mass associated to a given observable, which in turn can be obtained from the distribution of observables associated with a given mass.

The issue of contamination on the other hand is harder to handle. In principle, contamination could be modelled: the number density of contaminating noise fluctuations can

be determined from the noise properties, while astrophysical contaminants would alter the mass observable relation and the scatter around it. In practise, such modelling choices have however proven less successful than the selection of possibly pure samples. For this reason, we shall quickly outline how pure cluster samples can be constructed from observations in different wavelengths and what effects complicate such selections.

In the millimeter regime the selection of pure cluster samples is arguably the most controllable: clusters leave a distinctive spectral feature in the CMB, most noticeably as a shadow in given frequencies. Besides random noise fluctuation, there is no other way to create such a shadow. Any such signal of sufficient statistical significance is thus highly likely to be a cluster. Possible caveats in the millimeter regime are emitting sources that could fill up the shadowy signature of the SZ effect, thereby reducing it (most recently Gupta et al., 2017). This could lead to an alteration of the observable mass relation and the scatter around it. Two main modes need to be distinguished here: first, fractionally small signals with little cluster by cluster variation. These signals would simply alter the observable mass relation. Second, and more importantly, emitters that are only present in a subset of the clusters, but cancel an important part of the SZe signal if present. The latter class would lead to excess incompleteness beyond the incompleteness due to the observable mass scatter. We shall discuss in later chapters how to detect such excess incompleteness.

In the X-rays, cluster selection has to contend not only with noise fluctuations, but also with the fact that active galactic nuclei (AGN) with strong X-ray emission are more frequent than clusters by at least one order of magnitude. A catalog of statistically significant X-ray sources thus contains a majority of AGN and a minority of clusters. Crucially, however, AGN appear as point sources, while clusters have a diffuse emission. Consequently, traditionally X-ray cluster surveys selected extended sources (e.g. Vikhlinin et al., 1998; Böhringer et al., 2001; Romer et al., 2001; Clerc et al., 2014). Still, multiple blended AGN or AGN in clusters lead to contamination at a  $\sim 10\%$  level.

Both SZe and X-ray selection provide rather pure cluster samples. Yet, in these wavelengths information on the cluster redshift is hard or impossible to obtain. As a consequence, optical follow up of X-ray and SZe cluster samples is the norm. The quest of determining the cluster redshifts via the presence of red galaxies allows one to further improve the purity of a sample (for a most recent application, see Klein et al., 2018; Klein et al., 2019). Indeed, if an optical structure was not found at a given magnitude limit, this implied that the source was either a contaminant or at a redshift large enough that the cluster galaxies fall below the magnitude limit. While such studies were performed traditionally with pointed observations, recently the advent of wide and deep photometric surveys and the increasing size of X-ray and SZe candidate lists have made automated confirmation tools necessary. As we will show below, such automated optical confirmation tools allow one to construct predominantly X-ray selected cluster samples without selection based on the extent of sources.

The final form of cluster selection is direct optical selection. This selection type sets itself apart from the X-ray or SZe selection by the fact that the instrumental noise of the photometry does not impact the mass limit but only the redshift completeness. This is due to the fact that the limiting magnitude of photometric surveys allows for high signal to

noise detection of even individual galaxies below a typical redshift. If the cluster search is limited to appropriately low redshifts, in principle every halo should be detected. Indeed, the argument can be reversed to say that any over-density of red galaxies is associated to a halo of some mass, namely to the halo that hosts the brightest galaxy (this popular argument was tested on simulation by Farahi et al., 2016). The observed richness is then a result of the actual richness of the cluster and alterations due to projection effects. Indeed, the color filters used for optical selection subtend a large range of physical distances along the line of sight, which can lead to significant richness contributions by galaxies in the correlated LSS surrounding the cluster (Cohn et al., 2007; Costanzi et al., 2019). When talking about contamination of optically selected samples one should thus consider contaminants as objects with a mass that is surprisingly low compared to their richness in light of the richness–mass relation and scatter assumed.

## 1.5 Cosmological Inference from Cluster Samples

As anticipated above, the crucial ingredient for the cosmological exploitation of a cluster sample is to understand the mapping between halo mass and selection observable. Formally speaking, let  $\hat{\mathcal{O}}^i$  be the measured value of an observable, say the X-ray flux. Following Mantz et al. (2010) and Bocquet et al. (2015), we can relate this measured observable to the intrinsic observable  $\mathcal{O}$  by taking account of the instrumental noise and systematics, leading to a distribution

$$P(\hat{\mathcal{O}}^i | \mathcal{O}, z^i, \vec{\theta}^i), \quad (1.15)$$

describing the likelihood of the measured observable  $\hat{\mathcal{O}}^i$ , given the intrinsic observable  $\mathcal{O}$ , the clusters redshift  $z^i$  and the position on the sky  $\vec{\theta}^i$ , which can impact the instrumental noise via the local observing condition, for instance the local exposure time.

### 1.5.1 Observable-Mass Relation

As discussed above intrinsic observables of clusters at a given mass scatter around a given observable–mass relation. This relation is a priori not known, as it depends only to first order physics which is relatively easy to simulate. As such, a parameterized form of the mean observable mass relation is required  $\langle \ln \mathcal{O} \rangle(M, z, \vec{p}_{\mathcal{O}}, \vec{p}_c)$ , and a scatter  $\sigma_{\mathcal{O}}(M, z, \vec{p}_{\mathcal{O}}, \vec{p}_c)$ , which are not only functions of mass and redshift, but depend on the parameters of the observable–mass relation  $\vec{p}_{\mathcal{O}}$ , and on cosmological parameters  $\vec{p}_c$ . As an empirical approach we then assume the population of clusters to be described by a log-normal distribution in observable given mass and redshift

$$P(\ln \mathcal{O} | M, z, \vec{p}_{\mathcal{O}}, \vec{p}_c) = \ln \mathcal{N}(\ln \mathcal{O}; \langle \ln \mathcal{O} \rangle, \sigma_{\mathcal{O}}^2), \quad (1.16)$$

where  $\ln \mathcal{N}(x; \mu, \sigma^2)$  stands for a log-normal distribution in  $x$  with mean  $\mu$  and variance  $\sigma^2$ . The exact forms of relations used shall be discussed in the respective chapters below.

Given the mapping between measured and intrinsic observable, as well as the mapping between intrinsic observable and mass, the distribution of masses associated with any measured observable can be determined as

$$\begin{aligned} P(M|\hat{\mathcal{O}}^i, z^i, \vec{p}_{\mathcal{O}}, \vec{p}_{\mathcal{C}}, \vec{\theta}^i) &\propto P(\hat{\mathcal{O}}^i|M, z^i, \vec{p}_{\mathcal{O}}, \vec{p}_{\mathcal{C}}, \vec{\theta}^i) \frac{dN}{dM} \Big|_{M, z, \vec{p}_{\mathcal{C}}} \\ &\propto \int d\mathcal{O} P(\hat{\mathcal{O}}^i|\mathcal{O}, z^i, \vec{\theta}^i) P(\ln \mathcal{O}|M, z, \vec{p}_{\mathcal{O}}, \vec{p}_{\mathcal{C}}) \frac{dN}{dM} \Big|_{M, z, \vec{p}_{\mathcal{C}}}, \end{aligned} \quad (1.17)$$

where the proportionality constant is given by the condition that the expression averages to one if integrated in mass. The first line follows from Bayes' Theorem for the inversion of conditional probabilities, i.e.  $P(A|B) \propto P(B|A)P(A)$ , where in this case, the number of clusters as a function of mass acts as the prior for the cluster mass. It encodes the fact that lower mass halos are more frequent than higher mass halos. It is thus more probable that a low mass halo scatters high to the measured observable, than that a high mass halo scatters low. This effect is called *Eddington bias* and results in  $\langle \ln M \rangle < \hat{M}_{\mathcal{O}}^i$ : the expected value for the mass is always lower than the mass obtained by inverting the observable mass relation at the measured observable value, i.e.  $\ln \hat{\mathcal{O}}^i = \langle \ln \mathcal{O} \rangle(\hat{M}_{\mathcal{O}}^i, z)$ .

### 1.5.2 Predicted Number of Objects

The formalism of the observable–mass relation can be used to compute the differential number of objects  $\frac{dN}{d\hat{\mathcal{O}}}\Big|_{\hat{\mathcal{O}}^i, z^i}$  as a function of measured observable  $\hat{\mathcal{O}}^i$  and redshift  $z^i$ . This computation is done by transforming the differential number of halos as a function of mass into the space of measured observable using the mappings described above, and reads

$$\begin{aligned} \frac{dN}{d\hat{\mathcal{O}}}\Big|_{\hat{\mathcal{O}}^i, z^i, \vec{\theta}^i} &= P(\text{sel}|\hat{\mathcal{O}}^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}}) \int d\mathcal{O} P(\hat{\mathcal{O}}^i|\mathcal{O}, z^i, \vec{\theta}^i) \\ &\quad \int dM P(\ln \mathcal{O}|M, z, \vec{p}_{\mathcal{O}}, \vec{p}_{\mathcal{C}}) \frac{dN}{dM} \Big|_{M, z, \vec{p}_{\mathcal{C}}}, \end{aligned} \quad (1.18)$$

where  $P(\text{sel}|\hat{\mathcal{O}}^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}})$  defines the probability of detecting a cluster with measured observable  $\hat{\mathcal{O}}^i$ , redshift  $z^i$  at a position  $\vec{\theta}^i$  on the sky, which is traditionally called selection function. It can depend on parameters  $\vec{p}_{\mathcal{S}}$ , which encode the systematic uncertainty on the selection function.

The total number of clusters in a survey can be predicted as

$$N_{\text{tot}} = \int d^2\vec{\theta} \int d\hat{\mathcal{O}} \int dz \frac{dN}{d\hat{\mathcal{O}}}\Big|_{\hat{\mathcal{O}}, z, \vec{\theta}}. \quad (1.19)$$

Note that computing the total number of objects requires to evaluate the observational error model  $P(\hat{\mathcal{O}}|\mathcal{O}, z, \vec{\theta})$  and the selection function  $P(\text{sel}|\hat{\mathcal{O}}, z, \vec{\theta}, \vec{p}_{\mathcal{S}})$  at input values where no clusters have been observed.

Similarly, the differential number of clusters can also be predicted as a function of two observables  $\hat{\mathcal{O}}_1$  and  $\hat{\mathcal{O}}_2$  with their respective selection functions  $P(\text{sel}_\alpha|\hat{\mathcal{O}}_\alpha, z, \vec{\theta}, \vec{p}_{\mathcal{S}_\alpha})$  and observational error models  $P(\hat{\mathcal{O}}_\alpha|\mathcal{O}_\alpha, z, \vec{\theta})$ , as

$$\begin{aligned} \frac{d^2N}{d\hat{\mathcal{O}}_1 d\hat{\mathcal{O}}_2} \Big|_{\hat{\mathcal{O}}_1^i, \hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i} &= P(\text{sel}_1|\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_1}) \int d\mathcal{O}_1 P(\hat{\mathcal{O}}_1^i|\mathcal{O}_1, z^i, \vec{\theta}^i) \\ &\quad P(\text{sel}_2|\hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) \int d\mathcal{O}_2 P(\hat{\mathcal{O}}_2^i|\mathcal{O}_2, z^i, \vec{\theta}^i) \\ &\quad \int dM P(\ln \mathcal{O}_1, \ln \mathcal{O}_2|M, z, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c) \frac{dN}{dM} \Big|_{M, z, \vec{p}_c}, \end{aligned} \quad (1.20)$$

where we introduce the multi-variate scaling relation

$$P(\ln \mathcal{O}_1, \ln \mathcal{O}_2|M, z, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c) = \ln \mathcal{N} \left( \begin{bmatrix} \ln \mathcal{O}_1 \\ \ln \mathcal{O}_2 \end{bmatrix}; \begin{bmatrix} \ln \langle \mathcal{O}_1 \rangle \\ \ln \langle \mathcal{O}_2 \rangle \end{bmatrix}, \begin{bmatrix} \sigma_{\mathcal{O}_1}^2 & \sigma_{\mathcal{O}_1} \sigma_{\mathcal{O}_2} \rho \\ \sigma_{\mathcal{O}_1} \sigma_{\mathcal{O}_2} \rho & \sigma_{\mathcal{O}_2}^2 \end{bmatrix} \right), \quad (1.21)$$

where  $\ln \mathcal{N}(\vec{x}; \vec{\mu}, \mathbf{C})$  is a multi-variate log-normal in  $\vec{x}$  with mean  $\vec{\mu}$  and covariance  $\mathbf{C}$ . Note that in this formalism we introduce the correlation coefficient  $\rho = \sigma_{\mathcal{O}}(M, z, \vec{p}_{\mathcal{O}}, \vec{p}_c)$  between the scatters of the two observables, which, in principle, could be a function of mass, redshift, additional scaling relation parameters and even cosmology. In practise, however, it is usually assumed to be an unknown constant that is fitted for. From a physical perspective, correlated scatter has the potential to reveal interesting aspects of cluster astrophysics, highlighting if the processes driving the scatter in one observable also contributes to the scatter in the other observable.

One can easily obtain the differential number of objects in only one observable, despite the selection on two observables, by marginalizing the equation above over the second observable, i.e.

$$\begin{aligned} \frac{dN}{d\hat{\mathcal{O}}_1} \Big|_{\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i} &= \int d\hat{\mathcal{O}}_2 \frac{d^2N}{d\hat{\mathcal{O}}_1 d\hat{\mathcal{O}}_2} \Big|_{\hat{\mathcal{O}}_1^i, \hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i} = \\ &= P(\text{sel}_1|\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_1}) \int d\mathcal{O}_1 P(\hat{\mathcal{O}}_1^i|\mathcal{O}_1, z^i, \vec{\theta}^i) \\ &\quad \int d\mathcal{O}_2 P(\text{sel}_2|\mathcal{O}_2, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) \\ &\quad \int dM P(\ln \mathcal{O}_1, \ln \mathcal{O}_2|M, z, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c) \frac{dN}{dM} \Big|_{M, z, \vec{p}_c}, \end{aligned} \quad (1.22)$$

with  $P(\text{sel}_2|\mathcal{O}_2, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) = \int d\hat{\mathcal{O}}_2 P(\text{sel}_2|\hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) P(\hat{\mathcal{O}}_2^i|\mathcal{O}_2, z^i, \vec{\theta}^i)$

In the case of no selection based on the second observable, that is  $P(\text{sel}_2|\mathcal{O}_2, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) = 1$ , we find that the expression above becomes the same as derived for the single observable case.

### 1.5.3 Mass Calibration

The final element used in the description of cluster populations is the conditional probability of one observable  $\hat{\mathcal{O}}_2$  given another  $\hat{\mathcal{O}}_1$  at given redshift and position on the sky. This distribution is of great use if priors of the observable mass relation for one observable exist and one seeks to transfer that mass calibration to another observable. The most obvious case is weak lensing mass calibration, where the distribution of tangential shear profiles for a given observable is constructed. Mathematically, the aforementioned distribution is

$$P(\hat{\mathcal{O}}_2|\hat{\mathcal{O}}_1, z^i, \theta^i, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c, \vec{p}_{\mathcal{S}_2}) \propto P(\text{sel}_2|\hat{\mathcal{O}}_2, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) \int d\mathcal{O}_2 P(\hat{\mathcal{O}}_2|\mathcal{O}_2, z^i, \vec{\theta}^i) \\ P(\mathcal{O}_2|\hat{\mathcal{O}}_1^i, z^i, \theta^i, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c, \vec{p}_{\mathcal{S}_2}), \quad (1.23)$$

the first and the second term on the left hand side are given above, while the third term describes the expected values of the intrinsic observables  $\hat{\mathcal{O}}_2$  given the measured first observable  $\hat{\mathcal{O}}_1$ , the redshift  $z^i$ . Crucially, this prediction can only be undertaken for a set of scaling relation parameters  $(\vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2})$  for both observables. Consequently, evaluating equation 1.24 measures the level of agreement between the prediction for  $\hat{\mathcal{O}}_2$  based on the measured  $\hat{\mathcal{O}}_1^i$  and a set of scaling relation parameters, and the actual measured value  $\hat{\mathcal{O}}_2^i$ .

The prediction for the range of intrinsic observables  $\mathcal{O}_2$  given the measured observable  $\hat{\mathcal{O}}_1$  is constructed as follows

$$P(\mathcal{O}_2|\hat{\mathcal{O}}_1, z^i, \theta^i, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c, \vec{p}_{\mathcal{S}_2}) \propto \int d\mathcal{O}_1 P(\hat{\mathcal{O}}_1^i|\mathcal{O}_1, z^i, \vec{\theta}^i) \\ \int dM P(\ln \mathcal{O}_1, \ln \mathcal{O}_2|M, z, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{O}_2}, \vec{p}_c) \frac{dN}{dM} \Big|_{M, z, \vec{p}_c}. \quad (1.24)$$

Equation 1.24 needs to be normalized in such a way as to be a probability distribution in  $\hat{\mathcal{O}}_2$ . The presence of the selection function in the second observable ensures proper accounting for the Malmquist Bias: close to the selection threshold  $\text{sel}_2$ , not the entire range of  $\hat{\mathcal{O}}_2$ 's associated to  $\hat{\mathcal{O}}_1^i$  is detected. Only the objects that scatter high make it past the selection. Note that this kind of bias is only relevant if there is an actual selection on  $\hat{\mathcal{O}}_2$ . If the second observable is instead obtained by a follow up, even of a random sub-sample, this effect does not appear. Mathematically, we would then have  $P(\text{sel}_2|\hat{\mathcal{O}}_2, z^i, \vec{\theta}^i, \vec{p}_{\mathcal{S}_2}) = 1$ .

Inspection of the equations above reveals the Bayesian nature of this treatment of the cluster population. Indeed,

$$P(\hat{\mathcal{O}}_2^i|\hat{\mathcal{O}}_1^i, z^i, \theta^i) = \frac{P(\hat{\mathcal{O}}_2^i, \hat{\mathcal{O}}_1^i, z^i, \theta^i)}{P(\hat{\mathcal{O}}_1^i, z^i, \theta^i)} \\ \text{with } P(\hat{\mathcal{O}}_2^i, \hat{\mathcal{O}}_1^i, z^i, \theta^i) = \frac{d^2N}{d\hat{\mathcal{O}}_1 d\hat{\mathcal{O}}_2} \Big|_{\hat{\mathcal{O}}_1^i, \hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i} \\ \text{and } P(\hat{\mathcal{O}}_1^i, z^i, \theta^i) = \frac{dN}{d\hat{\mathcal{O}}_1} \Big|_{\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i}, \quad (1.25)$$

which follows exactly the definition of conditional probability.

### 1.5.4 Fitting

The population model outlined above has several free parameters:

- the cosmological parameters  $\vec{p}_C$ ,
- the parameters of the observable mass relation used, together with the associated scatters and correlation coefficients,  $\vec{p}_{\mathcal{O}_\alpha}$ , and
- the parameters encoding the systematic uncertainty in the selection functions  $\vec{p}_{\mathcal{S}_\alpha}$ .

Constraints on these parameters have to be derived from the data with the use of likelihoods. The basic concept behind this is: given a catalog of objects with properties  $z^i, \hat{\mathcal{O}}_1^i$  and  $\hat{\mathcal{O}}_2^i$  for a random subset, for all or for a subset with well a defined selection function, how probable is the population distribution, given the population model defined by a set of the aforementioned parameters. How to turn the likelihood into a constraint on the parameters is left to the preference of the reader between Frequentism and Bayesianism. Observational cosmologists are in the majority Bayesians, and so the likelihoods we present below will be sampled with priors to obtain posteriors on the parameters.

The most important likelihood in cluster cosmological studies is the Poisson Likelihood. The likelihood of finding a catalog with properties  $z^i, \hat{\mathcal{O}}_1^i$  as a function of cosmological parameters, scaling relation parameters and selection function parameters is given by

$$\ln \mathcal{L}_{1d\ nc}(\vec{p}_C, \vec{p}_{\mathcal{O}_1}, \vec{p}_{\mathcal{S}_1}) = \sum_i \ln \frac{dN}{d\hat{\mathcal{O}}_1} \Big|_{\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i} - N_{\text{tot}}. \quad (1.26)$$

Given a catalog with two observables and two selections, this naturally extends to

$$\ln \mathcal{L}_{2d\ nc}(\vec{p}_C, \vec{p}_{\mathcal{O}_\alpha}, \vec{p}_{\mathcal{S}_\alpha}) = \sum_i \ln \frac{d^2N}{d\hat{\mathcal{O}}_1 d\hat{\mathcal{O}}_2} \Big|_{\hat{\mathcal{O}}_1^i, \hat{\mathcal{O}}_2^i, z^i, \vec{\theta}^i} - N_{\text{tot}}. \quad (1.27)$$

Consider now the relation we established between the mass calibration probability and the number of objects in equation 1.25. Plugging it into equation 1.27 gives

$$\begin{aligned} \ln \mathcal{L}_{2d\ nc}(\vec{p}_C, \vec{p}_{\mathcal{O}_\alpha}, \vec{p}_{\mathcal{S}_\alpha}) &= \sum_i \ln \frac{dN}{d\hat{\mathcal{O}}_1} \Big|_{\hat{\mathcal{O}}_1^i, z^i, \vec{\theta}^i} - N_{\text{tot}} + \sum_i \ln P(\hat{\mathcal{O}}_2^i | \hat{\mathcal{O}}_1^i, z^i, \theta^i) \\ &= \ln \mathcal{L}_{2d\ nc}(\vec{p}_C, \vec{p}_{\mathcal{O}_\alpha}, \vec{p}_{\mathcal{S}_\alpha}) + \ln \mathcal{L}_{\text{mssclbr}}(\vec{p}_C, \vec{p}_{\mathcal{O}_\alpha}, \vec{p}_{\mathcal{S}_\alpha}), \end{aligned} \quad (1.28)$$

where we defined the mass calibration likelihood as

$$\ln \mathcal{L}_{\text{mssclbr}}(\vec{p}_C, \vec{p}_{\mathcal{O}_\alpha}, \vec{p}_{\mathcal{S}_\alpha}) = \sum_i \ln P(\hat{\mathcal{O}}_2^i | \hat{\mathcal{O}}_1^i, z^i, \theta^i), \quad (1.29)$$

as originally demonstrated by Bocquet et al. (2015).

The latter needs not to include all objects for which we have measured  $\hat{\mathcal{O}}_1$ , but is also well-defined for a random sub-sample or a sub-sample with a well-defined selection. We shall see in practice in the later chapters, that most of these parameters are degenerate with each other, and some external priors are required to extract cosmological information.

## 1.6 Outline

In this introduction, we outlined the physical picture of how halos form. This leads to the important concept that their number as a function of mass and redshift in a survey is a sensitive function of cosmological parameters. We then discussed that the most massive halos are inhabited by a specific type of physical object: galaxy clusters. Given the large halo mass, the observational signatures of these objects in different wavelengths scale with mass in a way that, to first order, can be derived from the virial equilibrium condition and adiabatic hydrodynamics, and to second order still follows a simple relation, which, however, has to be empirically calibrated. We then summarized how these signatures can be used to selected samples of clusters. Finally, we discussed the model employed to simultaneously extract cosmological information and to empirically calibrate the observable–mass relation. We shall now move to three applications of the model. The following material, with exception of the conclusion, draws directly from Grandis et al. (2019), Grandis et al. (in prep.[a]), and Grandis et al. (in prep.[b]).

### 1.6.1 Application 1

In the first application (Grandis et al., 2019) we forecast the impact of weak lensing (WL) cluster mass calibration on the cosmological constraints from the X-ray selected galaxy cluster counts in the upcoming eROSITA survey. We employ a prototype cosmology pipeline to analyze mock cluster catalogs. Each cluster is sampled from the mass function in a fiducial cosmology and given an eROSITA count rate and redshift, where count rates are modeled using the eROSITA effective area, a typical exposure time, Poisson noise and the scatter and form of the observed X-ray luminosity– and temperature–mass–redshift relations. A subset of clusters have mock shear profiles to mimic either those from DES and HSC or from the future Euclid and LSST surveys. Using a count rate selection, we generate a baseline cluster cosmology catalog that contains 13k clusters over  $14,892 \text{ deg}^2$  of extragalactic sky. Low mass groups are excluded using raised count rate thresholds at low redshift. Forecast parameter uncertainties for  $\Omega_M$ ,  $\sigma_8$  and  $w$  are 0.023 (0.016; 0.014), 0.017 (0.012; 0.010), and 0.085 (0.074; 0.071), respectively, when adopting DES+HSC WL (Euclid; LSST), while marginalizing over the sum of the neutrino masses. A degeneracy between the distance–redshift relation and the parameters of the observable–mass scaling relation limits the impact of the WL calibration on the  $w$  constraints, but with BAO measurements from DESI an improved determination of  $w$  to 0.043 becomes possible. With Planck CMB priors,  $\Omega_M$  ( $\sigma_8$ ) can be determined to 0.005 (0.007), and the summed neutrino mass limited to  $\sum m_\nu < 0.241 \text{ eV}$  (at 95%). If systematics on the group mass scale can be controlled, the eROSITA group and cluster sample with 43k objects and LSST WL could constrain  $\Omega_M$  and  $\sigma_8$  to 0.007 and  $w$  to 0.050.

### 1.6.2 Application 2

In the second application (Grandis et al., in prep.[a]) we construct and validate the selection function of the MARD-Y3 sample. This sample was selected through optical follow-up of the 2nd ROSAT faint source catalog (2RXS) with Dark Energy Survey year 3 (DES-Y3) data. The selection function is modeled by combining an empirically constructed X-ray selection function with an incompleteness model for the optical cleaning. We validate the joint selection function by testing the consistency of the constraints on the X-ray flux–mass and richness–mass scaling relation parameters derived from different sources of mass information: (1) cross-calibration using SPT-SZ clusters, (2) calibration using number counts in X-ray, in optical and in both X-ray and optical while marginalizing over cosmological parameters, and (3) other published analyses. We find that the constraints on the scaling relation from the number counts and SPT-SZ cross-calibration agree, indicating that our modeling of the selection function is adequate. Furthermore, we apply a largely cosmology independent method to validate selection functions via the computation of the probability of finding each cluster in the SPT-SZ sample in the MARD-Y3 sample and vice-versa. This test reveals no evidence for MARD-Y3 contamination or SPT-SZ incompleteness. However, we determine an outlier fraction from the scaling relation of  $\sim 10\%$ . We present evidence that these outlier fraction is likely explained by clusters with lower than expected SPT-SZ signals rather than higher than expected X-ray signals. Finally, we discuss the prospects of the techniques presented here to limit systematic selection effects in future cluster cosmological studies.

### 1.6.3 Application 3

In the third application (Grandis et al., in prep.[b]) we perform a cross validation of the cluster catalog selected by the red-sequence Matched-filter Probabilistic Percolation algorithm (RM) in the Dark Energy Survey year 1 (DES-Y1) data by matching it with the Sunyaev-Zel'dovich effect (SZe) selected cluster catalog from the South Pole Telescope (SPT) SZ survey. Using the mass information of the SZe signal, we calibrate the richness–mass relation above richness  $\hat{\lambda} > 40$  using a Bayesian cluster population model. We find a mass trend consistent with unity, no significant redshift evolution and an intrinsic scatter in richness of  $\sigma_\lambda = 0.22 \pm 0.06$ . By considering two error models, we explore the impact of projection effects on the richness–mass relation and the scatter around it, confirming that it is smaller than current systematic uncertainties. By comparing the SPT detection probability of RM objects derived from the mass information and the actual occurrence of detection for the RM- $(\hat{\lambda} > 40)$  sample, we find a purity consistent with unity and outlier fraction from the assumed scatter model consistent with 0. Extrapolating this relation to lower richness  $\hat{\lambda} > 20$  over-predicts the stacked weak lensing mass by 30%, at more than 2 sigma significance when accounting for systematic and statistical uncertainties. At the current level of systematic uncertainties, the number counts of RM objects are consistent with our richness–mass relation and cosmological constraints from SPT number counts. We discuss possible sources for the tension between ICM based richness–mass relation and

current stacked weak lensing measurements. While current sensitivities in SZe and X-rays do not allow us to discriminate between the different hypotheses, we outline how upcoming surveys like SPT-3G and eROSITA make this mass regime testable by multi-wavelength studies.

# Kapitel 2

## Application 1: Impact of Weak Lensing Mass Calibration on eROSITA Galaxy Cluster Cosmological Studies – a Forecast

### 2.1 Experimental setup

To constrain the impact of direct mass calibration through WL tangential shear measurements on eROSITA cluster cosmology, we create an eROSITA mock cluster catalog. The actual eROSITA cluster candidate catalog will be extracted from the eROSITA X-ray sky survey using specially designed detection and characterization tools (Brunner et al., 2018).

Each candidate source will be assigned a detection significance, an extent significance, an X-ray count rate and uncertainty, and other more physical parameters such as the flux within various observing bands (Merloni et al., 2012). For a subset of this sample, precise X-ray temperatures and rough X-ray redshifts will also be available (Borm et al., 2014; Hofmann et al., 2017).

This X-ray cluster candidate catalog will then be studied in the optical to identify one or more optical counterparts (assigning a probability to each) and to estimate a photometric redshift. A special purpose Multi-Component-Matched-Filter (MCMF) optical followup tool (Klein et al., 2018) has been designed for eROSITA cluster analysis and has been tested on available X-ray and SZE catalogs. It has been shown in RASS+DES analyses that one can reliably obtain both cluster and group redshifts over the relevant ranges of redshift (Klein et al., 2019), and thus for the analysis undertaken here we assume redshifts are available for all the eROSITA clusters.

The MCMF tool also allows one to quantify the probability of chance superposition between X-ray cluster candidates and optical counterparts, using the statistics of optical systems along random lines of sight together with estimates of the initial contamination in the X-ray cluster candidate catalog. Synthetic sky simulations by Clerc et al. (2018)

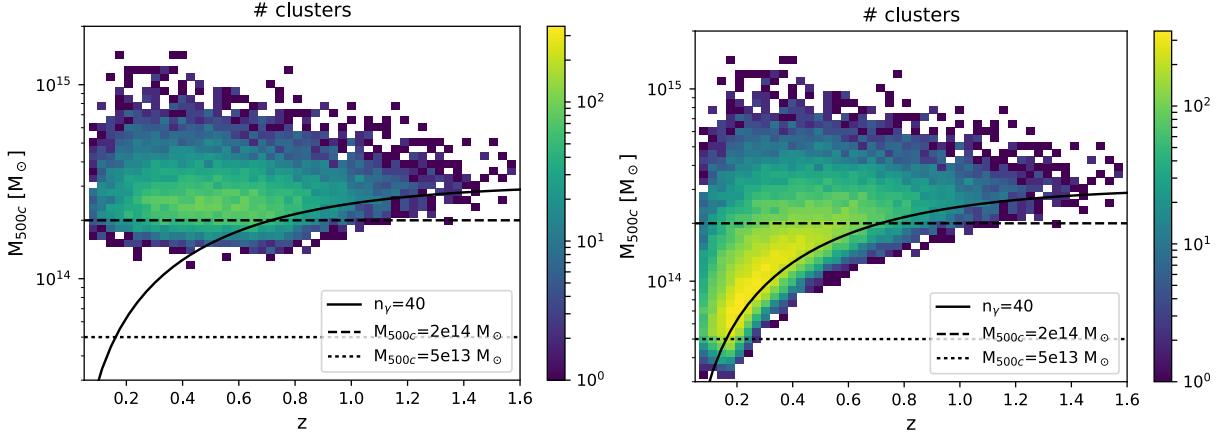


Abbildung 2.1: Distribution in halo mass  $M_{500c}$  and cluster redshift  $z$  of the mock, X-ray selected cluster catalogs used in this analysis. *Left:* Above redshift  $\sim 0.7$ , the 13k cluster baseline sample is selected using the fiducial count rate cut  $\eta = 2.5 \times 10^{-2} \text{ cts s}^{-1}$  that corresponds approximately to 40 photons at the median exposure time and a signal to noise  $\xi_{\text{det}} > 7$ . Below that redshift the observable cut is pushed upward to mimic a mass exclusion at  $M_{500c} \sim 2 \times 10^{14} M_{\odot}$ . Due to intrinsic and observational scatter between halo mass and the observable count rate, the cuts in observable used to create these samples appear smoothed in halo mass-redshift space. *Right:* The 43k sample that includes groups is selected similarly but the count rate cut is adjusted to mimic a mass exclusion at  $M_{500c} \sim 5 \times 10^{13} M_{\odot}$ .

have shown that the initial X-ray cluster candidate list selected on both detection and extent significance will be contaminated at the 10% level, consistent with experience in X-ray selection from archival ROSAT PSPC data that have a similar angular resolution to eROSITA (Vikhlinin et al., 1998). After processing with MCMF the resulting eROSITA X-ray cluster catalog is expected to have contamination at the sub-percent level. Therefore, we do not include contamination in the mock catalogs produced for this study.

For the WL mass calibration we will be using shear and photometric redshift catalogs from wide field, deep extragalactic surveys, including DES and HSC in the near term and Euclid and LSST on the longer term. The label “Euclid” refers to the nominal requirements for Euclid (Laureijs et al., 2011), although these requirements will realistically be met when combining Euclid with LSST, where the LSST data would be used for the photometric redshifts. We also explore the impact of LSST WL alone, where we adopt the requirements described in the following references (LSST DESC, 2012; LSST DESC et al., 2018). There is also the promise of CMB lensing as another method of mass calibration that is expected to be especially helpful for the highest redshift end of our cluster sample, but in our current analysis we do not model the impact of CMB lensing.

Our strategy in the analysis that follows is to adopt direct, cosmology independent cluster observables, including the cluster (1) X-ray detection significance or count rate, (2) photometric redshift, (3) WL tangential shear profile and (4) shear source redshift distributions for use in the cosmological analysis of the cluster sample. A benefit of using the count rate rather than the physical flux is that uncertainties in effective area and

the temperature dependence of the conversion from count rate to physical flux do not contribute to cosmological uncertainties.

Empirically mapping these observables to mass as a function of redshift and testing consistency of observed and theoretical cluster distributions as a function of cosmological parameters is described in Section 2.2. Below, in Section 2.1.1, we describe how the mock cluster catalog is generated and how the X-ray and optical cluster properties are assigned. In Section 2.1.2 we describe how we model the shear profiles that are produced for an appropriate subset of the mock eROSITA cluster sample. We discuss briefly our choice of fiducial cosmology and input X-ray scaling relations in Section 2.1.3.

### 2.1.1 Creating the mock cluster catalog

To create the X-ray catalog, we perform the following calculations.

1. For our choice of input cosmology (see Table 2.1 and Section 2.1.3), we compute the number of expected clusters as a function of halo mass  $M_{500c}$  and redshift  $z$  using the halo mass function (Tinker et al., 2008). We then draw a Poisson realization of the number of expected clusters, obtaining a mass selected cluster sample with  $M_{500c} > 1.3 \times 10^{13} M_{\odot}$  and  $0.05 < z < 1.8$ . For this calculation we assume a survey solid angle of  $\text{Area}_{\text{DE}} = 0.361 \times 4\pi$ , corresponding to regions of the western galactic hemisphere with a galactic hydrogen column  $N_{\text{H}} < 10^{21} \text{ cm}^{-2}$  (Kalberla et al., 2005). This corresponds approximately to a galactic latitude cut of  $|b| > 20$  deg. We adopt the cluster true redshift as the photometric redshift, because the MCMF optical followup tool has been demonstrated to achieve photometric redshift uncertainties with the DES dataset with an accuracy of  $\sigma_z/(z+1) \lesssim 0.01$  (Klein et al., 2018; Klein et al., 2019) out to redshifts  $z \sim 1.1$ . Photometric redshift uncertainties at this level are small enough to play no role in the cosmological analysis of the eROSITA cluster counts.
2. We use the scaling between X-ray luminosity  $L_{[0.5-2]\text{keV}}$  ( $L_X$  hereafter) in the rest frame  $0.5 - 2$  keV band and halo mass

$$\frac{L_X}{L_0} = e^{\ln A_L} \left( \frac{M_{500c}}{M_0} \right)^{B_L} \left( \frac{E(z)}{E_0} \right)^2 \left( \frac{1+z}{1+z_0} \right)^{\gamma_L} e^{\Delta_L}, \quad (2.1)$$

that was extracted from a large sample of SPT selected clusters with pointed XMM-*Newton* observations (Bulbul et al., 2019). In this relation  $E(z) = H(z)/H_0$  encodes the expansion history of the universe and is used to calculate the impact of changes in the critical density of the Universe ( $\rho_{\text{crit}} \propto E^2(z)$ ),  $\ln A_L$ ,  $B_L$  and  $\gamma_L$  are the amplitude, the mass trend and the non-self-similar redshift trend parameters of the luminosity–mass scaling relation, and  $\Delta_L \sim \mathcal{N}(0, \sigma_L^2)$  is a random number drawn from a Gaussian with standard deviation  $\sigma_L$ , which models the log-normal intrinsic scatter of the relation.

The Bulbul et al. (2019) X-ray scaling relations are derived from the Sunyaev-Zel'dovich effect (SZE) selected cluster sample from the SPT-SZ 2500 deg<sup>2</sup> survey (Carlstrom et al., 2011; Bleem et al., 2015) that have available XMM-*Newton* observations. This is a sample of 59 clusters with  $0.2 \leq z \leq 1.5$  and masses  $M_{500c} > 3 \times 10^{14} M_{\odot}$ . These halo masses have been calibrated separately in a cosmological analysis (de Haan et al., 2016) and exhibit a characteristic uncertainty of  $\sim 20\%$  (statistical) and  $\sim 15\%$  (systematic). The scaling relation parameter uncertainties from Bulbul et al. (2019) include both statistical and systematic uncertainties.

We also utilize the temperature mass relation

$$\frac{T}{T_0} = e^{\ln A_T} \left( \frac{M_{500c}}{M_0} \right)^{B_T} \left( \frac{E(z)}{E_0} \right)^{\frac{2}{3}} \left( \frac{1+z}{1+z_0} \right)^{\gamma_T} e^{\Delta_T}, \quad (2.2)$$

from the same analysis (Bulbul et al., 2019), where the parameters ( $\ln A_T, B_T, \gamma_T$ ) have the same meaning as in the luminosity scaling relation, with  $\Delta_T \sim \mathcal{N}(0, \sigma_T^2)$  for the scatter  $\sigma_T$ . The only difference is the scaling with the critical density, derived from self similar collapse theory.

Following these relations, we attribute to each cluster an X-ray luminosity  $L_X$  and a temperature  $T$ , randomly applying the respective intrinsic log normal scatter and assuming that the two scatters are uncorrelated.

3. Given the cluster rest frame 0.5-2 keV luminosity  $L_X$  and its redshift  $z$ , we compute the rest frame 0.5-2 keV flux

$$f_X = \frac{L_X}{4\pi d_L^2(z)}, \quad (2.3)$$

where  $d_L(z)$  is the luminosity distance.

4. For each cluster we calculate the X-ray spectrum assuming an APEC plasma emission model (Smith et al., 2001) with temperature  $T$  and metallicity  $Z = 0.3 Z_{\odot}$ <sup>1</sup>. This spectrum is normalized to the cluster rest frame 0.5-2 keV flux.
5. We compute the eROSITA count rate  $\eta$  for each cluster by shifting the spectrum to the observed frame and by averaging it with the eROSITA Ancillary Response Function (hereafter ARF) in the observed frame 0.5-2 keV band<sup>2</sup>. For simplicity, we do not follow the variation in neutral hydrogen column across the eROSITA-DE field. In fact, we ignore the impact of Galactic absorption altogether in our count rate calculation, which for the median neutral hydrogen column density in our footprint,  $N_H = 3 \times 10^{20} \text{ cm}^{-2}$  would lead on average to 5% lower rates.

<sup>1</sup>For simplicity, we do not apply any scatter to the metallicity, and assume it is constant as a function of redshift, as recent measurements of the metallicity of SPT selected clusters suggest (McDonald et al., 2016). We assume the solar abundances model of Anders and Grevesse (1989)

<sup>2</sup>Of the seven eROSITA cameras, two have a 100 nm Al and 200 nm Pl filter, while the remaining five have a 200 nm Al and 200 nm Pl filter (Predehl et al., 2010; Merloni et al., 2012). Consequently, the total ARF is the sum of two (100 nm Al + 200 nm Pl)-ARFs and five (200 nm Al + 200 nm Pl)-ARFs.

6. To model the measurement uncertainty on the rate, we draw a Poisson realization of the expected rate  $\hat{\eta} = \eta \pm \sqrt{\eta/t_{\text{exp}}}$ , where  $t_{\text{exp}} = 1600$  s is the expected median exposure time of the 4 year eROSITA survey (Pillepich, Porciani, and Reiprich, 2012). With this we account for the Poisson noise in the rate measurement. The count rate uncertainty for each cluster will be included in the real eROSITA cluster catalogs.
7. Finally, we select our baseline cluster sample using the count rate  $\eta > 2.5 \times 10^{-2} \text{ ct s}^{-1}$  (corresponding for our median exposure to  $\hat{n}_\gamma > 40$ ). For reference, given the background expectations, survey PSF and clusters modeled as  $\beta$  models with core radii that are 20% of the virial radius  $r_{500}$ , this selection threshold corresponds approximately to a cut in detection significance of  $\xi_{\text{det}} > 7$ , irrespective of the cluster redshift. Simple mock observations (see discussion in Appendix ??) indicate that at this threshold and above the extent likelihood for the eROSITA sample is  $\xi_{\text{ext}} > 2.5$ , enabling an initial eROSITA cluster candidate list after X-ray selection (but prior to optical followup) that is contaminated at the  $\sim 10\%$  level. At low redshift ( $z < 0.7$ ), we raise the detection threshold above the nominal level in such a way as to exclude most clusters with masses  $M_{500c} \lesssim 2 \times 10^{14} M_\odot$  at each redshift. We create a second sample to examine the impact of lower mass clusters and groups (see Section 2.3.6) by adjusting the low redshift count rate cut so that systems with masses  $M_{500c} \lesssim 5 \times 10^{13} M_\odot$  are excluded at each redshift. We discuss the X-ray selection in more detail in Appendix ???. The reasons for excluding lower mass systems are discussed below (cf. Section 2.3.6).

The procedure described above provides us with a baseline cosmology catalog of  $\sim 13k$  clusters. Their distribution in halo mass<sup>3</sup> and redshift is shown in the left panel of Fig. 2.1. They span a redshift range  $z \in (0.05, 1.6)$ . The total number of clusters and their redshift range are mainly impacted by the choice of the input cosmology, the observed luminosity mass relation, and the choice of cut in eROSITA count rate for selection. The sample has a median redshift  $\bar{z} = 0.51$  and median halo mass of  $\bar{M}_{500c} = 2.5 \times 10^{14} M_\odot$ . This sample extends to high redshift with 3% of the sample, corresponding to 420 clusters, at  $z > 1$ .

The sample of 43k objects with the count rate cut that only excludes lower mass systems with  $M_{500c} \leq 5 \times 10^{13} M_\odot$  is shown in Fig. 2.1 (right). The bulk of the additional low mass systems in this sample appear at redshifts  $z \leq 0.7$ . As with the overall number of clusters, the median mass and redshift depend on the observable cut used to exclude low mass objects, with these being  $\bar{z} = 0.30$ , and  $\bar{M}_{500c} = 1.4 \times 10^{14} M_\odot$ . We discuss the implications of lowering the mass limit in Section 2.3.6.

The number of objects in this  $\xi_{\text{det}} > 7$  group dominated sample is in good agreement with the numbers presented in previous discussions of the eROSITA cluster sample (Merloni et al., 2012; Pillepich, Porciani, and Reiprich, 2012; Pillepich et al., 2018). Importantly,

<sup>3</sup>We use this binning in mass just to visualize our sample, the number counts analysis will be performed on a fixed grid of observed rate  $\hat{\eta}$  and redshift, as specified in Section 2.2.2. The corresponding mass grid depends on the cosmological and the scaling relation parameters, and is thus recomputed every time the likelihood function is called on a specific set of parameters.

there are significantly more eROSITA clusters that can be detected if one reduces the detection threshold below  $\sim 7\sigma$ . But at that level there will be little extent information for each X-ray source, and so the candidate sample will be highly contaminated by AGN. Interestingly, Klein et al. (2018) have demonstrated that for the RASS faint source catalog where the survey PSF was so poor that little extent information is available, it is possible to filter out the non-cluster sources to produce low contamination cluster catalogs. The price for this filtering is that one introduces incompleteness for those systems that contain few galaxies (i.e., low mass clusters and groups at each redshift; see Klein et al., 2019).

### 2.1.2 Forecasting the WL signal

We adopt the cosmology independent tangential reduced shear profile  $\hat{g}_t(\theta_i)$  in radial bins  $\theta_i$  around the cluster as the observable for cluster WL mass calibration. A crucial complementary observable is the redshift distribution of the source galaxies  $N(z_s, z_{\text{cl}})$  behind the galaxy cluster, where  $z_s$  is the source redshift, and  $z_{\text{cl}}$  the cluster redshift. Assuming that the galaxy cluster mass profile is consistent with a Navarro-Frenk-White model (Navarro, Frenk, and White, 1996, hereafter NFW), these two observables can be combined into a measurement of the halo mass.

Although, in theory, WL mass calibration provides a direct mass measurement, in practice we refer to the mass resulting from an NFW fit to the shear profile as the WL mass  $M_{\text{WL}}$ . Following Becker and Kravtsov (2011), the WL mass is related to the halo mass by

$$M_{\text{WL}} = b_{\text{WL}} M_{200c} e^{\Delta_{\text{WL}}}, \quad (2.4)$$

with  $\Delta_{\text{WL}} \sim \mathcal{N}(0, \sigma_{\text{WL}}^2)$ , where  $\sigma_{\text{WL}}$  is the intrinsic log- normal scatter between WL mass and halo mass, induced by the morphological deviation of observed galaxy cluster mass profiles from the NFW profile, and  $b_{\text{WL}}$  is the WL mass bias describing the characteristic bias in the WL mass compared to the halo mass. This bias encodes several theoretical and observational systematics, as discussed below in Section 2.2.4.

Given that DES, HSC, Euclid and LSST will not overlap completely with the German eROSITA sky, only a fraction  $f_{\text{WL}}$  of the galaxy clusters of our X-ray mock catalog will have WL information available. Comparing the survey footprints, we estimate  $f_{\text{WL}} = 0.3$  for DES,  $f_{\text{WL}} = 0.05$  for HSC,  $f_{\text{WL}} = 0.5$  for Euclid, and  $f_{\text{WL}} = 0.62$  for LSST. For the LSST case we also assume that the northern celestial hemisphere portion of the German eROSITA sky with  $0^\circ < \delta < 30^\circ$  will be observed. For this northern extension of LSST, we adopt  $f_{\text{WL}} = 0.2$  and treat it as if it has the equivalent of DES depth. Therefore, we assign a WL mass only to a corresponding fraction of the eROSITA clusters in our mock catalogs, by drawing from equation (2.4).

Besides the WL mass and the cluster redshift, the background source distribution of the survey  $N(z_s)$  in redshift and the background source density  $n_\epsilon$  are necessary to predict the WL signal. For DES, we project  $n_\epsilon = 10 \text{ arcmin}^{-2}$  and utilize the redshift distribution presented in Stern et al. (2019), whose median redshift is  $z_{s,\text{m}} = 0.74$ . These parameters are derived from the Science Verification Data and their extrapolation to Y5 data will depend

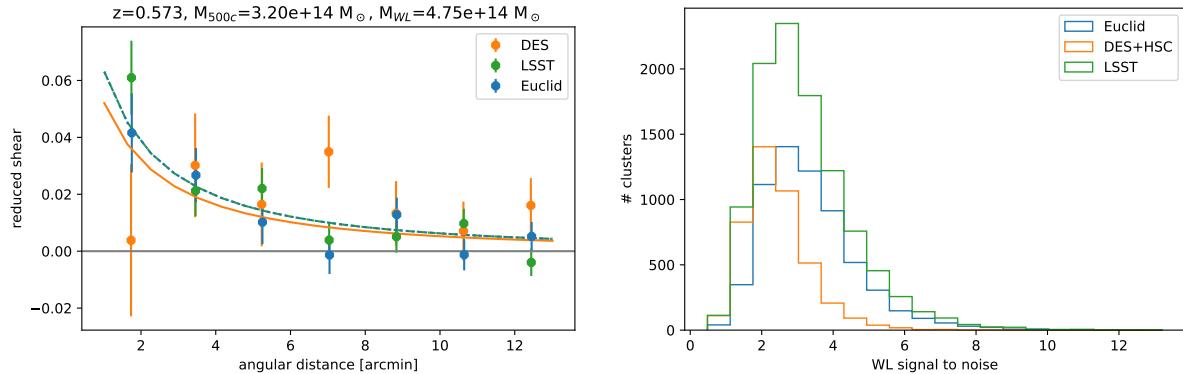


Abbildung 2.2: *Left*: Example of a shear profile in DES (orange), Euclid (blue) and LSST (green) data quality for a cluster. We show both the measured shear profile (dots with error bars) and the prediction (line). For all data quality cases, the measurement uncertainty is larger than the actual signal. *Right*: Distribution of WL signal to noise for DES+HSC (orange), Euclid (blue) and LSST (green), computed for each single cluster from the measured shear profile and the covariance matrix. While Euclid and LSST provide both more objects and higher signal to noise, objects with a clear WL mass measurement (e.g.,  $S/N > 5$ ) are rare for all datasets.

on the details of the future calibration (Gruen, priv. comm.). For HSC we assume  $n_{\epsilon} = 21 \text{ arcmin}^{-2}$ , and for the redshift distribution of HSC sources we adapt the parametrization by Smail, Ellis, and Fitchett (1994) with a median redshift  $z_{s,m} = 1.1$ . For Euclid, we use  $n_{\epsilon} = 30 \text{ arcmin}^{-2}$  (Laureijs et al., 2011). For the source redshift distribution we assume the parametric form proposed by Smail, Ellis, and Fitchett (1994) and utilized by Giannantonio et al. (2012), adopting a median redshift of  $z_{s,m} = 0.9$  (Laureijs et al., 2011). For LSST we assume  $n_{\epsilon} = 40 \text{ arcmin}^{-2}$  and parametrise the source redshift distribution as  $p(z_s) = 1/(2z_0)(z_s/z_0)^2 \exp(-z_s/z_0)$  with median redshift  $z_{m,s} = 2.67z_0 = 0.8^4$ .

The actual redshift distribution behind a galaxy cluster is assumed to be the survey redshift distribution with the cut  $N(z_s < z_{\text{cl}} + 0.1) = 0$ , where  $z_{\text{cl}}$  is the cluster redshift. This cut is helpful in reducing the contamination of the background source galaxies by cluster galaxies (that are not distorted by the cluster potential). This cut also leads to a reduction of the source density  $n_{\epsilon}(z_s > z_{\text{cl}} + 0.1)$  used to infer the observational noise on the cluster shear signal.

Given a redshift distribution, the mean reduced shear signal can be estimated, following Seitz and Schneider (1997), as

$$g_t(\theta_i) = \frac{\gamma(\theta_i)}{1 - \kappa(\theta_i)} \left( 1 + \kappa(\theta_i) \frac{\langle \beta^2 \rangle}{\langle \beta \rangle^2} \right), \quad (2.5)$$

where  $\gamma(\theta_i)$  and  $\kappa(\theta_i)$  are the shear and the convergence of an NFW mass profile,  $\theta_i$  the angular bins corresponding to radii between 0.25 and 5.0 Mpc at the cluster redshift

<sup>4</sup>These specification are taken from [https://www.lsst.org/sites/default/files/docs/sciencebook/SB\\_3.pdf](https://www.lsst.org/sites/default/files/docs/sciencebook/SB_3.pdf), Section 3.7.2

in our fiducial cosmology. This has the effect that low redshift clusters will have larger angular bins than high redshift clusters in to probe the similar physical scales. Also note that the inner radius, which we probe (0.25 Mpc), is smaller than in some previous studies (0.75 Mpc in Applegate et al., 2014; Stern et al., 2019; Dietrich et al., 2019). While this will require a more precise treatment of systematic effects such a cluster member contamination, miscentering and the impact of intra-cluster light on the shape and redshift measurements, theoretical predictions for the resulting WL mass bias and WL mass scatter associated with these smaller inner radii have already been presented (Lee et al., 2018). Furthermore, Gruen et al. (2019) investigated the impact of intra-cluster light on the photometric redshift measurement of background galaxies. We therefore assume that ongoing and future studies will demonstrate the possibility of exploiting shear information at smaller cluster radii, thereby increasing the amount of extracted mass information.

Following Bartelmann (1996), the shear and the convergence can be computed analytically for any halo, given the mass, the concentration, and the source galaxy redshift distribution  $N(z_s, z_{\text{cl}})$ . Throughout this work, the concentration of any cluster will be derived from its halo mass, following the relation presented by Duffy et al. (2008). The scatter in concentration at fixed halo mass is a contributor to the bias  $b_{\text{WL}}$  and scatter  $\sigma_{\text{WL}}$  in the WL mass to halo mass relation (equation 2.4). The lensing efficiency  $\beta = d_A(z_{\text{cl}}, z_s)/d_A(z_s)$  is the ratio between the angular diameter distance  $d_A(z_{\text{cl}}, z_s)$  from the cluster to the source, and the angular diameter distance  $d_A(z_s)$  from the observer to the source. In equation (2.5) the symbol  $\langle \cdot \rangle$  denotes averaging over the source redshift distribution  $N(z_s, z_{\text{cl}})$ .

The covariance of the measurement uncertainty on the reduced shear is

$$\mathbf{C}_{i,j} = \text{Cov}[g_t(\theta_i), g_t(\theta_j)] = \frac{\sigma_\epsilon^2}{\Omega_i n_\epsilon(z_{\text{cl}})} \delta_{i,j} + (\mathbf{C}_{\text{uLSS}})_{i,j} \quad (2.6)$$

where  $\delta_{i,j} = 1$ , if  $i = j$ , and  $\delta_{i,j} = 0$  else. The first term accounts for the shape noise in each radial bin, estimated by scaling the intrinsic shape noise of the source galaxies  $\sigma_\epsilon = 0.27$  by the number of source galaxies in each radial bin, taking into account the reduction of source galaxy density  $n_\epsilon(z_{\text{cl}}) = n_\epsilon(z_s > z_{\text{cl}} + 0.1)$  and the angular area of the  $i$ -th radial bin  $\Omega_i$ . We also add a contribution coming from uncorrelated large scale structure  $(\mathbf{C}_{\text{uLSS}})_{i,j}$  (Hoekstra, 2003). We draw the measured reduced shear profile  $\hat{g}_t$  from the Gaussian multivariate distribution with mean  $g_t$  and covariance  $\mathbf{C}$ .

For each cluster with WL information, we thus save the source redshift distribution  $N(z_s, z_{\text{cl}})$ , the measured reduced shear profile  $\hat{g}_t$ , and the covariance  $C$ . We show an example for a measured reduced shear profile, both in DES, in Euclid and in LSST data quality in the left panel of Fig. 2.2.

The WL signal around individual galaxy clusters derived from wide and deep photometric surveys is typically low signal to noise. In the right panel of Fig. 2.2, we explore the distribution of WL signal to noise for the subsamples with DES+HSC WL data, Euclid WL data and LSST WL data. To this end we define the signal to noise as  $S/N = \sqrt{0.5 \hat{g}_t^T \mathbf{C}^{-1} \hat{g}_t}$ . While the Euclid and LSST data provide a higher signal to noise on average, it rarely exceeds  $S/N > 5$ . Thus, we confirm that WL mass calibration provides a low signal to noise, direct mass measurement for a large subset of our cluster catalog.

Tabelle 2.1: Input parameters for our analysis. The exact definition of the parameters listed below is given in Section 2.2.3, Section 2.2.1 and Section 2.1.2 for the cosmological parameters, the scaling relation parameters and the WL calibration parameters, respectively. *Comments:* *a)* This value is determined to match  $\sigma_8 = 0.768$  by de Haan et al. (2016). *b)* We utilize here the value corresponding to the minimal model of a Cosmological Constant causing the accelerated expansion. *c)* This is the minimal value allowed by flavor neutrino oscillations, as reviewed by Tanabashi et al. (2018).

Cosmological Parameters		
$H_0$	73.02	Riess et al. (2016)
$\omega_b$	0.02202	Cooke et al. (2014)
$\Omega_M$	0.306	de Haan et al. (2016)
$A_S$	1.5792e-9	a)
$n_s$	0.9655	Planck Collaboration et al. (2016a)
$w$	-1.00	b)
$\sum m_\nu$	0.06 eV	c)
$\Omega_K$	0.	
Luminosity–Mass–Redshift Relation		
$\ln A_L$	1.52	Bulbul et al. (2019)
$B_L$	1.95	
$\gamma_L$	-0.20	
$\sigma_L$	0.237	
Temperature–Mass–Redshift Relation		
$\ln A_T$	1.83	Bulbul et al. (2019)
$B_T$	0.849	
$\gamma_T$	-0.28	
$\sigma_T$	0.177	
WL Mass Bias and Scatter		
$b_{WL}$	0.94	Dietrich et al. (2019) &
$\sigma_{WL}$	0.24	Lee et al. (2018)

### 2.1.3 Fiducial cosmology and scaling relations

Several steps in the above outlined creation of the mock data are cosmology sensitive. Therefore, the choice of input cosmology will impact the catalog properties. As an input cosmology, we choose the best fitting  $\Omega_M$  and  $\sigma_8$  results from the most recent SPT galaxy cluster cosmology analysis (de Haan et al., 2016). We also assumed that dark energy can be described by a cosmological constant, i.e. that the dark energy equation of state parameter  $w = -1$ . Furthermore, we adopt the minimal neutrino mass allowed by flavor neutrino oscillation measurements,  $\sum m_\nu = 0.06$  eV (Tanabashi et al., 2018). The parameter values are listed in Table 2.1.

It is worth noting here that these input values for  $\Omega_M$  and  $\sigma_8$  are somewhat different (at less than  $2\sigma$  significance) from the best fit values derived from the Planck CMB anisotropy measurements (Planck Collaboration et al., 2016a). This choice is intentional, as the masses of SPT clusters derived from a mass function fit with Planck CMB priors have been shown to be systematically high by studies of their WL signal (Dietrich et al., 2019; Stern et al., 2019), their dynamical mass (Capasso et al., 2019a) and their baryon content (Chiu et al., 2018). Furthermore, the input X-ray scaling relations by Bulbul et al. (2019), adapted to determine the X-ray properties of our catalog entries, assume an SZE signature–mass–redshift scaling relation consistent with the best fit results from the SPT galaxy cluster cosmology analysis. In summary, the input values for our analysis are chosen from the latest results of the SPT galaxy cluster sample, guaranteeing consistency between the assumed cosmology and the input X-ray scaling relations that we use to construct the mock eROSITA sample. Given that SPT covers a mass range of  $M_{500c} \gtrsim 3 \times 10^{14} M_\odot$ , and a redshift range of  $z \in (0.20, 1.7)$ , adopting SPT results within the eROSITA context implies only a modest extrapolation in mass and redshift.

On the other hand, the minimal neutrino mass is slightly inconsistent with recent results from joint fits to number counts of SPT selected clusters and Planck CMB measurements (de Haan et al., 2016; Bocquet et al., 2019a), which detect the neutrino mass at 2-3 sigma. This detection is likely sourced by the slight inconsistency in the  $(\Omega_M, \sigma_8)$  plane discussed above. For the sake of this work, we adapt the minimal neutrino mass to predict improvement on the upper limits obtained, if cluster number counts and CMB measurements were in perfect agreement.

## 2.2 Cosmology analysis method

In this section we describe the method we have developed for the cosmological analysis of an eROSITA cluster sample in the presence of WL mass calibration information. This method builds upon a method developed and used for the analysis of the SPT SZE selected cluster sample (Bocquet et al., 2015; Dietrich et al., 2019; Stern et al., 2019; Bocquet et al., 2019a). We start with a description of the minimal scaling relation to describe the mapping of the selection observable to halo mass as a function of redshift (Section 2.2.1), present the likelihoods in Section 2.2.2 and discuss the likelihood sampling tool and our adopted

priors in Sections 2.2.3 and 2.2.4.

### 2.2.1 Cluster selection scaling relation

The cosmological analysis of a galaxy cluster sample requires a model for the relation between the halo mass and the observable. In this work, we take an approach which is conceptually similar to the modeling of the scaling relation used for the SPT galaxy cluster sample first presented and applied to derive cosmological constraints by Vanderlinde et al. (2010) (for further applications, see for instance Benson et al. (2013), Bocquet et al. (2015), de Haan et al. (2016), and Bocquet et al. (2019a)). We empirically calibrate a scaling relation between the selection observable, i.e. the eROSITA count rate  $\eta$ , and the halo mass and redshift. As motivated in Appendix ??, we adopt the following scaling of the count rate with mass and redshift:

$$\frac{\eta}{\eta_0} = e^{\ln A_X} \left( \frac{M_{500c}}{M_0} \right)^{B(z)} \left( \frac{E(z)}{E_0} \right)^2 \left( \frac{d_L(z)}{d_{L,0}} \right)^{-2} \left( \frac{1+z}{1+z_0} \right)^{\gamma_X} e^{\Delta_X}, \quad (2.7)$$

where the amplitude is  $A_X$ , the redshift dependent mass slope is given by

$$B(z) = B_X + B'_X \ln \left( \frac{1+z}{1+z_0} \right), \quad (2.8)$$

the redshift trend describing departures from self-similar evolution is  $\gamma_X$ , and the deviation of a particular cluster from the mean scaling relation is described as  $\Delta_X \sim \mathcal{N}(0, \sigma_X^2)$ , with scatter  $\sigma_X$  (i.e., log-normal scatter in observable at fixed halo mass). As pivot points we choose  $M_0 = 2 \times 10^{14} \text{ M}_\odot$ ,  $z_0 = 0.5$ ,  $E_0 = 1.314$ ,  $d_{L,0} = 2710 \text{ Mpc}$ , and  $\eta_0 = 0.06 \text{ cts s}^{-1}$ .

Empirical calibration of the scaling relation has some major advantages compared to trying to measure accurate physical cluster quantities such as the flux. In doing the latter, the one might suffer biases (e.g. the effect of substructures in the context of eROSITA found by Hofmann et al., 2017) or additional sources of scatter from lack of knowledge about the cluster physical state. Furthermore, any such biases might themselves have trends with mass or redshift. An alternative approach, which has been adopted with success within SPT, is to use mass calibration to empirically determine the values of the scaling relation parameters. In this approach, an unbiased solution is found assuming the correct likelihood is adopted (see Section 2.2.2) and that the form of the observable mass scaling relation that is adopted has sufficient flexibility to describe the cluster population. One can examine this using goodness of fit tests (see Bocquet et al., 2015; de Haan et al., 2016). There is now considerable evidence in the literature that empirical calibration leads to a more robust cosmological experiment.

In summary, our model for the rate mass scaling assumes that the rate is a power law in mass and redshift with log-normal intrinsic scatter that is independent of mass and redshift. Our model allows the mass slope to vary with redshift, which is required given the redshift dependence of the eROSITA counts to physical flux conversion (see discussion in Appendix ??). Natural extensions of this model to, e.g., follow mass or redshift dependent scatter are possible, but for the analysis presented here we adopt a scaling relation with the following five free parameters:  $(\ln A_X, B_X, \gamma_X, \sigma_X, B'_X)$ .

### 2.2.2 Likelihood functions

The likelihood functions we employ to analyze our mock eROSITA and WL data are hierarchical, Bayesian models, introduced in this form by Bocquet et al. (2015). The functions account self-consistently for (1) the Eddington and Malmquist bias, (2) the cosmological dependencies of both the direct mass measurements and of the cluster number counts, and (3) systematic uncertainties in the halo mass of objects observed with a particular rate and redshift. Given that we utilize a realistic mock catalog, these likelihoods constitute a prototype of the eROSITA cosmological analysis pipeline. Using this scheme, we design three likelihoods: (1) mass calibration with perfect masses, (2) mass calibration with WL observables and (3) number counts. In the following, to ensure a concise notation, we will refer to the halo mass  $M_{500c}$  as  $M$ , and specify when we mean a mass defined w.r.t. any other overdensity.

#### Mass calibration with perfect masses

The likelihood that a cluster of measured rate  $\hat{\eta}$  and redshift  $z$  has a given mass  $M$  is given by

$$P(M|\hat{\eta}, z) \propto \int d\eta P(\hat{\eta}|\eta, z) P(\eta|M, z) \frac{dN}{dM}(M, z), \quad (2.9)$$

where

1.  $P(\hat{\eta}|\eta, z)$  is the probability density function (hereafter pdf) encoding the measurement error on the rate,
2.  $P(\eta|M, z)$  is the pdf describing the scaling relation between rate and halo mass at a given redshift. We model it as a log-normal distribution with central value given by equation (2.7) with scatter  $\sigma_x$ ,
3.  $\frac{dN}{dM}(M, z)$  is the derivative of the number of clusters w.r.t. to the mass at that redshift, which is the product of the halo mass function  $\frac{dn}{dM}(M, z)$  by Tinker et al. (2008), the co-moving volume element  $\frac{dV}{dz}(z)$  and the survey solid angle  $\Omega_{\text{DE}}$ .

These quantities, with the exception of the rate measurement uncertainty kernel, depend on scaling relation parameters, mass function parameters and cosmological parameters. Also note, that equation (2.9) needs to be properly normalized to be a pdf in halo mass  $M$ .

The total log-likelihood for mass calibration with perfect masses is then given by the sum of the natural logarithms of the likelihoods of the single clusters

$$\ln \mathcal{L}_{\text{pfct}} = \sum_j \ln P(M^{(j)}|\hat{\eta}^{(j)}, z^{(j)}), \quad (2.10)$$

where  $j$  runs over all clusters whose halo mass is known. Note that the perfect mass is only accessible in the case of a mock catalogue. This likelihood is thus not applicable to real data. Nevertheless, it is a function of the scaling relation and the cosmological parameters and can be used to extract the true underlying scaling relation from a mock dataset.

### WL mass calibration

The likelihood that a cluster with measured rate  $\hat{\eta}$  and redshift  $z$  has an observed tangential shear profile  $\hat{g}_t(\theta_i)$  can be computed as

$$P(\hat{g}_t|\hat{\eta}, z) = \int dM_{WL} P(\hat{g}_t(\theta_i)|M_{WL}, z_{cl})P(M_{WL}|\hat{\eta}, z), \quad (2.11)$$

where

1. the probability of a cluster with measured rate  $\hat{\eta}$  and redshift  $z$  to have a WL mass  $M_{WL}$  is

$$P(M_{WL}|\hat{\eta}, z) \propto \int dM \int d\eta P(\hat{\eta}|\eta, z)P(M_{WL}, \eta|M, z) \frac{dN}{dM}(M, z), \quad (2.12)$$

with  $P(M_{WL}, \eta|M, z)$  being the joint pdf describing the scaling relations for the rate and the WL mass, given in equations (2.7 and 2.4), respectively,

2. the probability of a cluster of WL mass  $M_{WL}$  having an observed reduced shear profile  $\hat{g}_{t,i} = \hat{g}_t(\theta_i)$  is given by a Gaussian likelihood

$$\ln P(\hat{g}_t|M_{WL}, z) = -\frac{1}{2} \ln (2\pi \det \mathbf{C}) - \frac{1}{2} \Delta \hat{g}_t^T \mathbf{C}^{-1} \Delta \hat{g}_t, \quad (2.13)$$

with  $\Delta \hat{g}_t = \hat{g}_t - g_t$ , where  $g_t$  is the tangential shear profile computed following equation (2.5) for a cluster of mass  $M_{WL}$  and the redshift distribution  $N(z_s, z_{cl} = z)$ .

The total log-likelihood for mass calibration with WL then reads

$$\ln \mathcal{L}_{WL\ mssclbr} = \sum_j \ln P(\hat{g}_t^{(j)}|\hat{\eta}^{(j)}, z^{(j)}), \quad (2.14)$$

where  $j$  runs over all clusters with WL information.

### Number counts

We also model the observed number of clusters  $\hat{N}$  in bins of measured rate  $\hat{\eta}$  and redshift  $z$ . We predict this number by computing the expected number of clusters in each bin, given the scaling relation, halo mass function and cosmological parameters

$$N(\hat{\eta}, z) = P(\det|\hat{\eta}, z) \int dM \int d\eta P(\hat{\eta}|\eta, z)P(\eta|M, z) \frac{dN}{dM}(M, z), \quad (2.15)$$

where  $P(\det|\hat{\eta}, z)$  is a binary function parameterizing if the bin falls within the selection criteria or not. Assuming a pure rate selection might be a simplification compared to the actual cluster selection function of the forthcoming eROSITA survey (for a study of this selection function, c.f. Clerc et al. (2018)). In summary, the expected number of clusters

in observable space can be computed using the cosmology dependent halo mass function, volume– redshift relation and observable–mass relation.

The number counts likelihood for the entire sample is the sum of the Poisson log-likelihoods in the individual bins

$$\ln \mathcal{L}_{\text{nmbr cts}} = \sum_{\text{bins}} \hat{N} \ln N - N. \quad (2.16)$$

As above, this likelihood is a function of the scaling relation, halo mass function and the cosmological parameters.

## Validation

To validate these likelihoods, we create a mock which is ten times larger than the eROSITA mock (by considering the unphysical survey footprint  $\text{Area}_{\text{test}} = 10 \text{ Area}_{\text{DE}}$ ). This leads to a reduction of the statistical uncertainties that enables us to better constrain systematic biases. We analyze this mock with the number counts and the Euclid WL mass calibration likelihood. We find that all parameters are consistent with the input values within less than two sigma. Scaling this up to the normal sized mock, we conclude that our code is unbiased at or below  $\sim \frac{2}{3}$  sigma. We present for inspection a plot showing the results of the validation run as Fig. 2.13 at the end of the paper. The plot shows the marginal contours of the posterior distributions for the parameters with the input values marked.

Given that our mock catalog is a random realization of the stochastic processes modeled by the above described likelihoods, and that these likelihoods retrieve the input values even for a ten times larger mock, we take the liberty to shift best fit parameter values of the posterior samples presented in the following sections. These shifts are of the order of one sigma. Putting all posteriors to the same central value allows us to highlight the improvement of constraining power visible in the shrinking of the contours.

### 2.2.3 Comments on sampling and model choice

Various combinations of the above described likelihood functions are sampled using `pymultinest` (Buchner et al., 2014), a python wrapper of the nested sampling code `multinest` (Feroz, Hobson, and Bridges, 2009). Nested sampling was originally developed to compute the evidence, or marginal likelihood, but has the added advantage of providing a converged posterior sample in the process (Skilling, 2006).

The parameters we sample depend on the specific application. In all cases considered, we sample the parameters of the X-ray selection scaling relation:  $(\ln A_X, B_X, \gamma_X, \sigma_X, B'_X)$ . When the WL mass calibration likelihood is sampled in Section 2.3.2, also the parameters governing the WL mass scaling relation are sampled:  $(b_{\text{WL}}, \sigma_{\text{WL}})$ .

We explore two different flat cosmological models: (1)  $\nu$ - $\Lambda$ CDM, and (2)  $\nu$ - $w$ CDM. For both, we consider the following parameters:  $H_0$ , the current expansion rate of the Universe in units of  $\text{km s}^{-1} \text{ Mpc}^{-1}$ ;  $\omega_b$ , the current day co-moving density of baryons w.r.t. the critical density of the Universe;  $\Omega_M$ , the current day density of matter w.r.t. the critical

density;  $A_S$ , the amplitude of primordial curvature fluctuations;  $n_S$ , the spectral index of primordial curvature fluctuations; and the sum of neutrino masses  $\sum m_\nu$  in eV.

The cosmological model where only these parameters are allowed to vary is called  $\nu$ - $\Lambda$ CDM, because we allow for massive neutrinos of yet unknown mass, and assume that the agent of the late time accelerated expansion is a cosmological constant  $\Lambda$ .

As a more complex model  $\nu$ - $w$ CDM, we also consider the case that the late time acceleration is not caused by the cosmological constant, but by an as yet unknown form of energy, usually referred to as dark energy. The properties of dark energy are described here by a single equation of state parameter  $w$ .

For better comparison, with other Large Scale Structure experiments, in both models, we also compute  $\sigma_8$ , the root mean square of linear matter fluctuations in a spherical region of  $8 h^{-1}\text{Mpc}$  radius, as a derived quantity in each step of the chain and present the posterior distribution in this quantity rather than in the primordial power spectrum fluctuation amplitude  $A_S$ .

### 2.2.4 Choice of priors

In general, any Bayesian analysis, and more specifically `pymultinest`, requires the specification of priors for all parameters one intends to sample. In the following, we present our choice of priors. If the parameter is not mentioned below, it has a uniform prior in a range that is larger than the typical posterior uncertainties of that parameter. The prior choices are summarized in Table 2.2.

#### Current priors on scaling relation

As mentioned above— and discussed in detail in Appendix 2.6.2— the eROSITA count rate scaling relation is described by five parameters:  $(\ln A_X, B_X, \gamma_X, \sigma_X, B'_X)$ . We put Gaussian priors on these parameters. The mean values are obtained in Section ?? by determining the maximum likelihood points of the mass calibration likelihood when using perfect masses. The corresponding uncertainties in the priors are taken to match the uncertainties on the respective parameters presented in Table 5 of Bulbul et al. (2019) for the core included 0.5-2.0 keV luminosity-mass-redshift relation when fit with the scaling relation of Form II. These parameter uncertainties were extracted using a sample of 59 SPT selected galaxy clusters observed with *XMM-Newton* together with the SPT SZE-based halo masses calculated using the calibration from de Haan et al. (2016, see Table 3 results column 2).

When we extract cosmological constraints **only** with these priors (i.e., without any WL information) we consider that a “baseline” result representing a currently achievable knowledge of the parameters of the eROSITA rate-mass relation.

#### Priors on WL calibration

The priors on the parameters of the WL mass – halo mass relation reflect the understanding of both the observational and theoretical systematics of the WL mass calibration. In this

Tabelle 2.2: Priors used in our analysis.  $\mathcal{U}(a, b)$  is a uniform flat prior in the interval  $(a, b)$ ,  $\ln \mathcal{U}(a, b)$  a uniform flat prior in log space,  $\mathcal{N}(\mu, \sigma^2)$  refers to a Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$ ,  $\mathcal{N}_{>a}(\mu, \sigma^2)$  to a Gaussian distribution truncated for values smaller than  $a$ .

*Comment:* a) Numerical stability when computing the equations (2.9, 2.11, 2.12 and 2.15), requires the scatter to be larger than the sampling size of the numerical integrals.

Cosmology for Number counts w/o CMB		
$H_0$	$\mathcal{U}(40, 120)$	cf. Section 2.2.4
$\omega_b$	$\mathcal{U}(0.020, 0.024)$	
$\Omega_M$	$\mathcal{U}(0.1, 0.5)$	
$A_S$	$\ln \mathcal{U}(0.6e - 9, 2.5e - 9)$	
$n_S$	$\mathcal{U}(0.94, 1.0)$	
$\sum m_\nu [eV]$	$\mathcal{U}(0., 1.)$	
$w$	$\mathcal{U}(-1.6, -0.6)$	
Cosmology for Number counts w/ CMB		
cf. Section 2.2.4		
X-ray Selection Scaling Relation		
$\ln A_X$	$\mathcal{N}(-0.33, 0.23^2)$	cf. Appendix ??
$B_X$	$\mathcal{N}(2.00, 0.17^2)$	
$\gamma_X$	$\mathcal{N}(0.45, 0.42^2)$	
$\sigma_X$	$\mathcal{N}_{>0.1}(0.28, 0.11^2)$	a)
$B'_X$	$\mathcal{N}(0.36, 0.78^2)$	
DES/HSC WL		
$b_{WL}$	$\mathcal{N}(0.94, 0.051^2)$	cf. Section 2.2.4
$\sigma_{WL}$	$\mathcal{N}_{>0.1}(0.24, 0.02^2)$	a)
Euclid WL		
$b_{WL}$	$\mathcal{N}(0.94, 0.013^2)$	cf. Section 2.2.4
$\sigma_{WL}$	$\mathcal{N}_{>0.1}(0.24, 0.008^2)$	a)
LSST WL		
$b_{WL}$	$\mathcal{N}(0.94, 0.015^2)$	cf. Section 2.2.4
$\sigma_{WL}$	$\mathcal{N}_{>0.1}(0.24, 0.008^2)$	a)

work, we consider, the following sources of systematic uncertainty:

1. the accuracy of the shape measurement in the optical survey parameterized as the uncertainty on the multiplicative shear bias  $\delta m$ ,
2. the systematic mis-estimation of the lensing efficiency  $\langle \beta \rangle$  due to the bias in the photometric redshift estimation  $b_z$ ,
3. the uncertainty in the estimation of the contamination by cluster members  $f_{\text{cl}}$  which results from the statistical uncertainty of the photometric redshifts  $\sigma_z$  and the background galaxy selection,
4. the statistical uncertainty with which the theoretical bias and scatter of the WL mass  $\delta b_{\text{WL, sim}}$ , and  $\delta \sigma_{\text{WL, sim}}$ , respectively, can be constrained with large structure formation simulations.

The first three effects do not directly induce a bias in the mass estimation, but affect the NFW fitting procedure. To estimate their impact on the WL mass estimate, we consider a shear profile for WL mass  $3 \times 10^{14} M_{\odot}$  and  $z = 0.4$ , add the systematic shifts, and fit for the mass again. The difference in input and output masses is then taken as the WL mass systematic uncertainty induced by these effects. This technique provides an overall estimate of the systematic uncertainty level, while ignoring potential dependences on cluster redshift and mass.

For DES, we assume  $\delta m = 0.013$  (Zuntz et al., 2018). The bias on the photometric redshift estimation of the source galaxies is  $b_z = 0.02$  (Cooke et al., 2014) which, considering the source redshift distribution of DES (cf. Section 2.1.2), leads to an uncertainty on the lensing efficiency  $\delta \langle \beta \rangle = 0.02$ . For the uncertainty on the contamination, we project  $\delta f_{\text{cl}} = 0.01$  based on Dietrich et al. (2019). Taken all together, these uncertainties propagate to a WL mass uncertainty of  $\delta b_{\text{WL, obs, DES}} = 0.045$ .

The current uncertainty on the theoretical WL mass bias is  $\delta b_{\text{WL, sim, to day}} = 0.05$  in Dietrich et al. (2019), when considering the effects of halo triaxiality, morphological variety, uncertainties in the mass-concentration relation and mis-centering. Due to larger available simulations (Lee et al., 2018), a better measurement of the mis-centering distribution and an improvement of the understanding of the mass– concentration relation, for DES we project a reduction of this uncertainty by a factor 2, yielding  $\delta b_{\text{WL, sim, DES}} = 0.025$ . The same scaling is applied to the uncertainty on the scatter, yielding  $\delta \sigma_{\text{WL, DES}} = 0.02$ .

Given the level of observational uncertainty, this projection can also be read as a necessity to improve the understanding of the theoretical biases. The estimates above provide a total uncertainty of the bias of the WL mass

$$\begin{aligned} \delta b_{\text{WL, DES}} &= \sqrt{b_{\text{WL, sim, DES}}^2 + b_{\text{WL, obs, DES}}^2} \\ &= 0.051, \end{aligned} \tag{2.17}$$

and an uncertainty on the scatter of the WL mass  $\delta \sigma_{\text{WL, DES}} = 0.02$ . This amounts to a 5.1% mass uncertainty from systematic effects, which is a conservative assumption, given

that McClintock et al. (2019a) already achieved such a level of systematics control for DES cluster mass calibration. For sake of simplicity, we assume that the final level of systematics in HSC is of the same as in DES. This assumption will be inadequate for the actual analysis of the data. We postpone the discussion about the difference between the analysis methods to the respective future works.

The specifications for Euclid are given in Laureijs et al. (2011). The requirement for the shape measurement is  $\delta m = 0.001$ . For the bias on the photometric redshift estimation, the requirement is  $b_{\hat{z}} = 0.002$ , which translates into  $\delta\langle\beta\rangle = 0.0014$ . For the projection of the uncertainty on the contamination, we assume that in the case of DES it has equal contribution from (1) the number of clusters used for to characterize it and (2) the photometric redshift uncertainty. Thus, for Euclid we estimate

$$\begin{aligned}\delta f_{\text{cl, Eu}}^2 &= \frac{\delta f_{\text{cl, DES}}^2}{2} \frac{N_{\text{DES}}}{N_{\text{Eu}}} + \frac{\delta f_{\text{cl, DES}}^2}{2} \left( \frac{\sigma_{\hat{z}, \text{Eu}}}{\sigma_{\hat{z}, \text{DES}}} \right)^2 \\ &= 0.0065^2,\end{aligned}\quad (2.18)$$

where  $N_{\text{DES}} \approx 3.8\text{k}$ , and  $N_{\text{Eu}} \approx 6.4\text{k}$ , are the number of clusters with DES and Euclid shear information in our catalog (cf. Section 2.1.2),  $\sigma_{\hat{z}, \text{Eu}} = 0.06$  is the photometric redshift uncertainty for Euclid (Laureijs et al., 2011), and  $\sigma_{\hat{z}, \text{DES}} = 0.1$  is the photometric redshift uncertainty for DES (Sánchez et al., 2014). Taking all the above mentioned values together, we find  $\delta b_{\text{WL, obs, Eu}} = 0.0085$  for Euclid. To match this improvement in data quality, we project an improvement in the understanding of the theoretical biases by a factor of 5, providing  $\delta b_{\text{WL, sim, Eu}} = 0.01$ , and  $\delta \sigma_{\text{WL, Eu}} = 0.008$ . Thus, the total uncertainty on the WL mass bias for Euclid is

$$\delta b_{\text{WL, Eu}} = 0.013. \quad (2.19)$$

The specifications for LSST systematics are summarized in LSST DESC et al. (2018). The requirement for the shape measurement is  $\delta m = 0.003$ , while the requirement for the bias on the photometric redshift estimation  $b_{\hat{z}} = 0.001$ , leading to  $\delta\langle\beta\rangle = 0.0007$ . Using  $N_{\text{LSST}} \approx 11\text{k}$ , and  $\sigma_{\hat{z}, \text{LSST}} = 0.02$ , we find an uncertainty on the cluster member contamination of  $\delta f_{\text{cl, LSST}} = 0.0044$ . Summing all the above mentioned values together, we get  $\delta b_{\text{WL, obs, LSST}} = 0.011$ . We project the same understanding in theoretical systematics for LSST as for Euclid. Thus, the total uncertainty on the WL mass bias for LSST is

$$\delta b_{\text{WL, LSST}} = 0.015. \quad (2.20)$$

These values are adopted throughout this work as priors for the WL mass scaling relation parameters, as summarized in Table 2.2. We note that the effort required to theoretically constrain the WL bias and scatter parameters with this accuracy is considerable.

### Cosmological priors

When sampling the number counts likelihood, we assume flat priors on all cosmological parameters except for  $A_S$ , for which we use a flat prior in log-space, as is good practice for

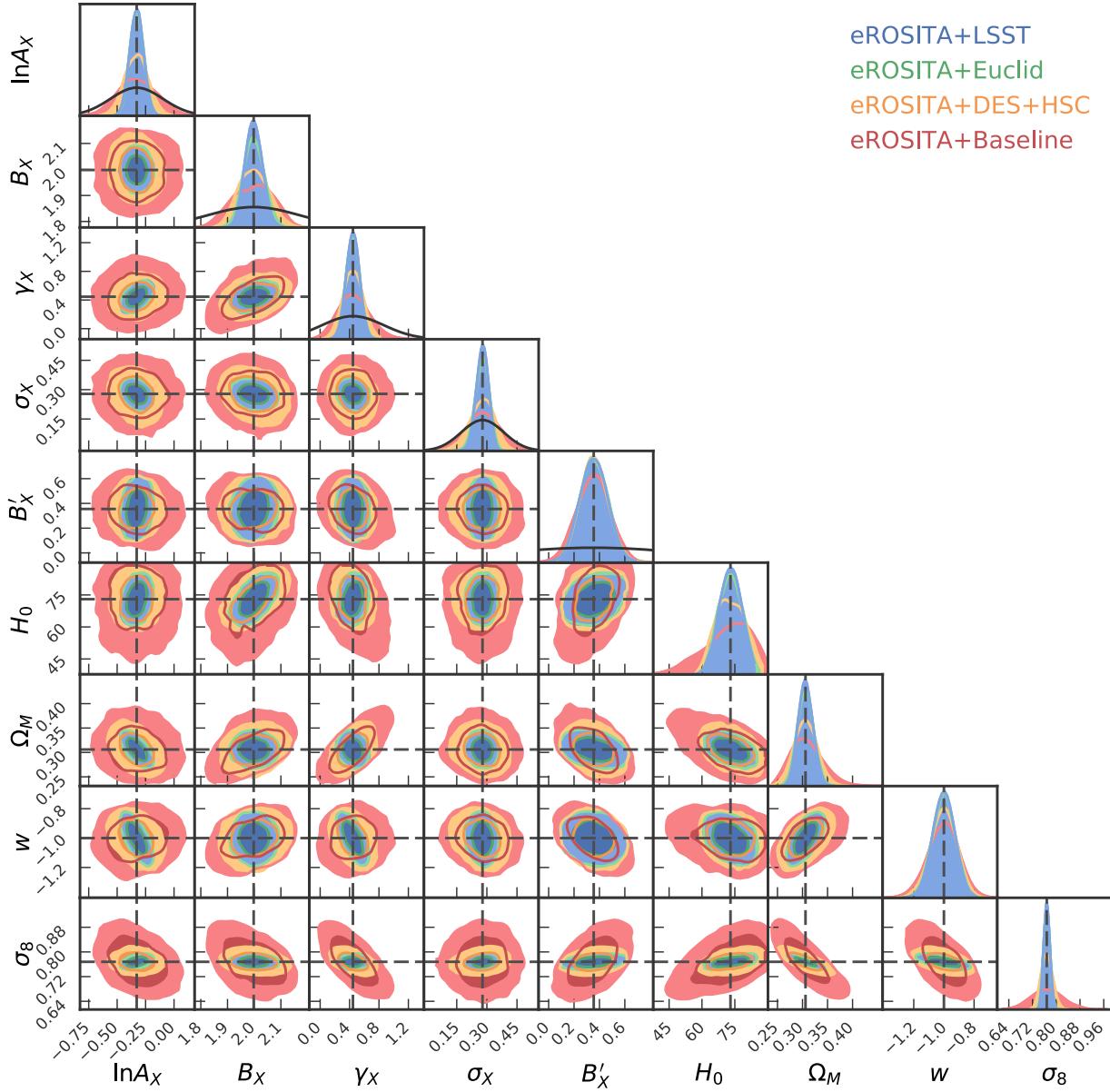


Abbildung 2.3: Predicted constraints on the scaling relation and cosmological parameters in  $w$ CDM. In red the constraints from the number counts alone (eROSITA+Baseline), in orange the constraints from number counts and DES+HSC WL calibration (eROSITA+DES+HSC), in green number counts and Euclid WL calibration (eROSITA+Euclid), and in blue number counts and LSST WL calibration (eROSITA+LSST). The median values, all statistically consistent with the input values, are shifted to the input values to better highlight the increase in constraining power.

strictly positive amplitudes. Similarly, we use priors on  $\Omega_M$ ,  $H_0$  and  $w$  that are larger than the typical uncertainties on these parameters. For  $\sum m_\nu$  we only explore the regime up to 1 eV, as current cosmological measurements, such as Planck Collaboration et al. (2016a) give upper limits on the summed neutrino mass around and below that value.

For  $\omega_b$  and  $n_S$  we use tight flat priors around the measured values of these parameters by the CMB experiments (Planck Collaboration et al., 2016a) and Big Bang Nucleosynthesis constraints derived from deuterium abundances (Cooke et al., 2014). We confirm that cluster number counts are not sensitive to these parameters within these tight ranges (Bocquet et al., 2019a). It is thus not necessary to use informative priors on these parameters, as previous studies have done (see for instance Bocquet et al., 2015; de Haan et al., 2016).

In Section 2.3.3 we will consider the synergies between eROSITA number counts and WL mass calibration, and CMB temperature and polarization anisotropy measurements, which to date provide us with a significant amount of information about the cosmological parameters. In the models of interest, where either  $w$  or  $\sum m_\nu$  are free parameters, the CMB constraints from the Planck mission (Planck Collaboration et al., 2016a) display large degeneracies between the parameters we choose to sample.<sup>5</sup> For this reason, we cannot approximate the CMB posterior as a Gaussian distribution. To capture the non-Gaussian feature, we calibrate a nearest-neighbor kernel density estimator (KDE) on the publicly available<sup>6</sup> posterior sample. We utilize Gaussian kernels and, for each model, we tune the bandwidth through cross calibration to provide maximum likelihood of the KDE on a test subsample. As discussed in Section 2.1.3, our choice of input cosmology is slightly inconsistent with the CMB constraints. As we are only interested in the reduction of the uncertainties when combining CMB and eROSITA, we shift the CMB posteriors so that they are consistent with our input values at less than one sigma. The resulting estimator reproduces the parameter uncertainties and the degeneracies accurately.

## 2.3 Results

In the following subsections we first calculate how accurately the observable–mass scaling relation parameters must be constrained to enable the best possible cosmological constraints from the sample (Section 2.3.1). Thereafter we explore the impact of the WL mass calibration on the cosmological constraints that can be extracted from an analysis of the eROSITA galaxy cluster counts (Section 2.3.2). In Section 2.3.3 we explore synergies of the eROSITA dataset with the CMB and in Section 2.3.4 we examine the impact of combining the eROSITA dataset with BAO measurements from DESI. In Section 2.3.5 we examine the constraints derived when combining with both these external data sets, and the fi-

<sup>5</sup>These degeneracies are partially due to our choice of sampling parameters. The CMB does not directly constrain  $H_0$ , which is a present day quantity. Consequently, also  $\Omega_M$  is weakly constrained. The same holds for  $w$ , which has predominantly a late time impact on the expansion rate. In contrast, co-moving densities like  $\omega_b$ , or primordial quantities like  $A_S$  and  $n_S$  are narrowed down with high precision.

<sup>6</sup><https://pla.esac.esa.int/pla/#cosmology>, where we utilized the TTTEEE\_lowTEB samples.

nal subsection focuses on the impact of an eROSITA sample where the minimum mass is allowed to fall from our baseline value of  $M_{500c} \gtrsim 2 \times 10^{14} M_{\odot}$  to  $M_{500c} \gtrsim 5 \times 10^{13} M_{\odot}$ , corresponding to a sample that is  $\sim 3.5$  times larger.

### 2.3.1 Optimal mass calibration

The number counts likelihood depends both on the scaling relation parameters, and—through the mass function, the cosmological volume and their changes with redshift—also on the cosmological parameters. Furthermore, there are significant degeneracies between the mass scale of the cluster sample (i.e., the parameters of the observable mass relation) and the cosmological parameters, as demonstrated already in the earliest studies (Haiman, Mohr, and Holder, 2001). A full self-calibration of the number counts (i.e., including no direct mass measurement information) that allows full cosmological and scaling relation freedom, results in only very weak cosmological constraints (e.g., Majumdar and Mohr, 2003; Majumdar and Mohr, 2004). Thus, before forecasting the cosmological constraints from the eROSITA sample, we estimate how accurate the mass calibration needs to be so that the information contained in the number counts is primarily resulting in the reduction of uncertainties on the cosmological parameters rather than the observable mass scaling relation parameters.

To estimate this required level of mass calibration, which we refer to as ‘optimal mass calibration’, we quantify how much the number counts constrain the scaling relation parameters when the cosmological parameters are fixed to fiducial values. In such a case, all the information contained in the number counts likelihood informs our posterior on the scaling relation parameters. If this level of information, or more, were provided by direct mass calibration, then the number counts information would predominantly constrain the cosmology. In this sense, the optimal mass calibration then provides a threshold or goal for the amount and precision of external mass calibration we should strive for in our direct mass calibration through, e.g., weak lensing.

We find that in fact the number counts alone do not contain enough information to meaningfully constrain all five scaling relation parameters even in the presence of full cosmological information. Our scaling relation parametrization includes two additional parameters beyond those explored in Majumdar and Mohr (2003), the scatter  $\sigma_X$  and the redshift evolution of the mass trend  $B'_X$ . Thus, as a next test, we examine the constraints from number counts with fixed cosmology while assuming priors only on  $B'_X$ . Interestingly, in this case we find that the constraints lead to an upper limit on the scatter of the scaling relation  $\sigma_X < 0.44$  (at 95%), which is weaker than our current knowledge of that parameter, which we infer from the scatter in the X-ray luminosity–mass relation from Bulbul et al. (2019, see discussions in Section 2.2.4 and Appendix ??). We therefore adopt this external prior on the scatter parameter and allow full freedom for all other parameters (including  $B'_X$ ). Results in this case are more interesting, providing constraints that we adopt as our estimate of optimal mass calibration. The uncertainties are  $\delta \ln A_X = 0.042$ ,  $\delta B_X = 0.024$ ,  $\delta \gamma_X = 0.053$ , and  $\delta B'_X = 0.116$ . We take this to mean that an optimal cosmological exploitation of the eROSITA cluster number counts will require that we know

the parameters of the observable mass relation to at least these levels of precision. We will discuss in the following how this can be accomplished.

### 2.3.2 Forecasts: eROSITA+WL

#### $\nu$ - $w$ CDM constraints

As a first cosmological model we investigate  $\nu$ - $w$ CDM, a flat cold Dark Matter cosmology with dark energy with constant but free equation of state parameter  $w$  and massive neutrinos. In this Section, we present the constraints on the cosmological parameters for three different cases: number counts alone combined with baseline priors on the X-ray observable mass scaling relation that we derive from the latest analysis within SPT (Bulbul et al., 2019) (eROSITA+Baseline), number counts with DES+HSC WL mass calibration (eROSITA+DES+HSC), number counts with Euclid WL mass calibration (eROSITA+Euclid), and number counts with LSST WL mass calibration (eROSITA+LSST). The respective marginal contour plot is shown in Fig. 2.3, and the corresponding uncertainties are listed in Table 2.3.

Considering the current knowledge of the X-ray scaling relation, we find that eROSITA number counts constrain  $\Omega_M$  to  $\pm 0.032$ ,  $\sigma_8$  to  $\pm 0.052$ ,  $w$  to  $\pm 0.101$ , and  $H_0$  to  $\pm 10.72$   $\text{km s}^{-1} \text{Mpc}^{-1}$ , while marginalizing over the summed neutrino mass  $\sum m_\nu < 1 \text{ eV}$  without constraining it. We also find no constraints on  $\omega_b$  and  $n_s$  within the prior ranges that we assumed.

The addition of mass information consistently reduces the uncertainties on the cosmological parameters: the knowledge on  $\Omega_M$  is improved by factors of 1.4, 2.0 and 2.3 when adding DES+HSC, Euclid, and LSST WL information, respectively; for  $\sigma_8$  the improvements are 3.1, 4.3 and 5.2, whereas for the dark energy equation of state parameter they are 1.2, 1.4 and 1.4, respectively. In summary, weak lensing calibration provides the strongest improvement of the determination of  $\sigma_8$ , followed by  $\Omega_M$ . The improvements on the dark energy equation of state parameter  $w$  are clearly weaker.

#### $\nu$ - $\Lambda$ CDM constraints

We also investigate a model in which the equation of state parameter  $w$  is kept constant:  $\nu$ - $\Lambda$ CDM. The corresponding uncertainties are shown in Table 2.3. In this model, we find that the constraints on  $\Omega_M$  and  $\sigma_8$  are 0.019 and 0.032, respectively, which is tighter than in the  $\nu$ - $w$ CDM model. However, the constraint on  $H_0$  is comparable in the two models.

We also find that the addition of WL mass information improves the constraints on  $\Omega_M$  by factors of 1.6, 2.4 and 2.9 for DES+HSC, Euclid and LSST, respectively. The determination of  $\sigma_8$  improves by factors 2.4, 4.7 and 4.7. It is especially worth highlighting how eROSITA with Euclid or LSST WL information will be able to determine  $\sigma_8$  at a sub-percent level. Nevertheless, also in this simpler model we find that eROSITA number counts do not constrain the summed neutrino mass in the sub-eV regime.

Tabelle 2.3: Forecast parameter constraints for eROSITA number counts with current, best available calibration (eROSITA+Baseline), with DES+HSC WL calibration (eROSITA+DES+HSC), with Euclid WL calibration (eROSITA+Euclid), and with LSST WL calibration (eROSITA+LSST) are presented in two different models,  $\nu$ - $w$ CDM and  $\nu$ - $\Lambda$ CDM within three different scenarios. From top to bottom they are eROSITA+WL alone, in combination with Planck CMB constraints (Pl15) and in combination with DESI BAO and Alcock-Paczynski test constraints. Also shown are the scaling relation parameter uncertainties for an optimal mass calibration. In addition to the five cosmological parameters who constraints are presented, each model includes the parameters  $n_S$  and  $\omega_b$  marginalized over weak priors (see Table 2.2). The units of the column “ $\sum m_\nu$ ” and “ $H_0$ ” are eV and  $\text{km s}^{-1} \text{ Mpc}^{-1}$ , respectively. *Comments:* a) This parameter is not constrained within the prior ranges. When reporting upper limits “<”, we refer to the 95th percentile, while lower limits “>” refer to the 5th percentile. When a parameter is kept fixed in that model, we use “\_”.

	$\Omega_M$	$\sigma_8$	$w$	$\sum m_\nu$	$H_0$	$\ln A_X$	$B_X$	$\gamma_X$	$\sigma_X$	$B'_X$
optimal mass calibration						0.042	0.024	0.053	0.116	
eROSITA + WL calibration										
$\nu$ - $w$ CDM	priors					0.23	0.17	0.42	0.11	0.78
	eROSITA+Baseline	0.032	0.052	0.101	a)	10.72	0.165	0.073	0.209	0.083
	eROSITA+DES+HSC	0.023	0.017	0.085	a)	6.449	0.099	0.053	0.121	0.062
	eROSITA+Euclid	0.016	0.012	0.074	a)	5.210	0.059	0.037	0.090	0.034
	eROSITA+LSST	0.014	0.010	0.071	a)	4.918	0.058	0.031	0.089	0.030
$\nu$ - $\Lambda$ CDM	priors	—	—	—	—	0.23	0.17	0.42	0.11	0.78
	eROSITA+Baseline	0.026	0.033	—	a)	10.18	0.157	0.069	0.192	0.078
	eROSITA+DES+HSC	0.016	0.014	—	a)	5.664	0.091	0.049	0.103	0.059
	eROSITA+Euclid	0.011	0.007	—	a)	4.691	0.040	0.035	0.065	0.033
	eROSITA+LSST	0.009	0.007	—	a)	4.691	0.039	0.032	0.058	0.029
eROSITA + WL calibration + Pl15 (TTTEEE,lowTEB)										
$\nu$ - $w$ CDM	priors (incl. CMB)	<0.393	0.063	0.242	<0.667	>62.25	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.019	0.032	0.087	<0.590	2.857	0.165	0.026	0.132	0.083
	eROSITA+DES+HSC	0.018	0.019	0.085	<0.554	2.206	0.099	0.024	0.118	0.062
	eROSITA+Euclid	0.014	0.010	0.074	<0.392	1.789	0.059	0.020	0.090	0.034
	eROSITA+LSST	0.013	0.009	0.069	<0.383	1.662	0.058	0.018	0.080	0.030
$\nu$ - $\Lambda$ CDM	priors (incl. CMB)	0.024	0.035	—	<0.514	1.723	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.016	0.018	—	<0.425	1.192	0.122	0.025	0.101	0.077
	eROSITA+DES+HSC	0.013	0.015	—	<0.401	1.067	0.086	0.023	0.098	0.060
	eROSITA+Euclid	0.011	0.007	—	<0.291	0.978	0.039	0.020	0.065	0.033
	eROSITA+LSST	0.009	0.007	—	<0.285	0.767	0.038	0.020	0.054	0.030
eROSITA + WL calibration + DESI (BAO)										
$\nu$ - $w$ CDM	priors (incl. BAO)	0.007	a)	0.086	a)	a)	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.007	0.030	0.063	a)	1.987	0.164	0.043	0.139	0.083
	eROSITA+DES+HSC	0.006	0.010	0.051	a)	1.597	0.086	0.037	0.110	0.056
	eROSITA+Euclid	0.006	0.005	0.047	a)	1.463	0.040	0.030	0.086	0.032
	eROSITA+LSST	0.006	0.005	0.043	a)	1.403	0.040	0.026	0.076	0.029
$\nu$ - $\Lambda$ CDM	priors (incl. BAO)	0.006	a)	—	a)	a)	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.006	0.015	—	a)	0.943	0.094	0.041	0.109	0.078
	eROSITA+DES+HSC	0.006	0.010	—	a)	0.925	0.074	0.040	0.077	0.055
	eROSITA+Euclid	0.006	0.005	—	a)	0.910	0.040	0.029	0.054	0.032
	eROSITA+LSST	0.006	0.005	—	a)	0.910	0.035	0.025	0.053	0.027
eROSITA + WL calibration + DESI + Pl15										
$\nu$ - $w$ CDM	priors (incl. CMB+BAO)	0.007	0.027	0.049	<0.284	1.118	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.006	0.026	0.049	<0.281	1.103	0.161	0.023	0.079	0.083
	eROSITA+DES+HSC	0.006	0.011	0.048	<0.245	1.050	0.085	0.023	0.071	0.061
	eROSITA+Euclid	0.005	0.006	0.047	<0.241	1.023	0.039	0.017	0.064	0.032
	eROSITA+LSST	0.005	0.006	0.039	<0.223	0.870	0.038	0.017	0.064	0.029
$\nu$ - $\Lambda$ CDM	priors (incl. CMB+BAO)	0.004	0.020	—	<0.256	0.255	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.004	0.016	—	<0.254	0.253	0.093	0.024	0.067	0.074
	eROSITA+DES+HSC	0.004	0.009	—	<0.218	0.251	0.072	0.021	0.062	0.051
	eROSITA+Euclid	0.003	0.004	—	<0.211	0.148	0.035	0.020	0.050	0.033
	eROSITA+LSST	0.002	0.003	—	<0.185	0.145	0.033	0.017	0.050	0.033

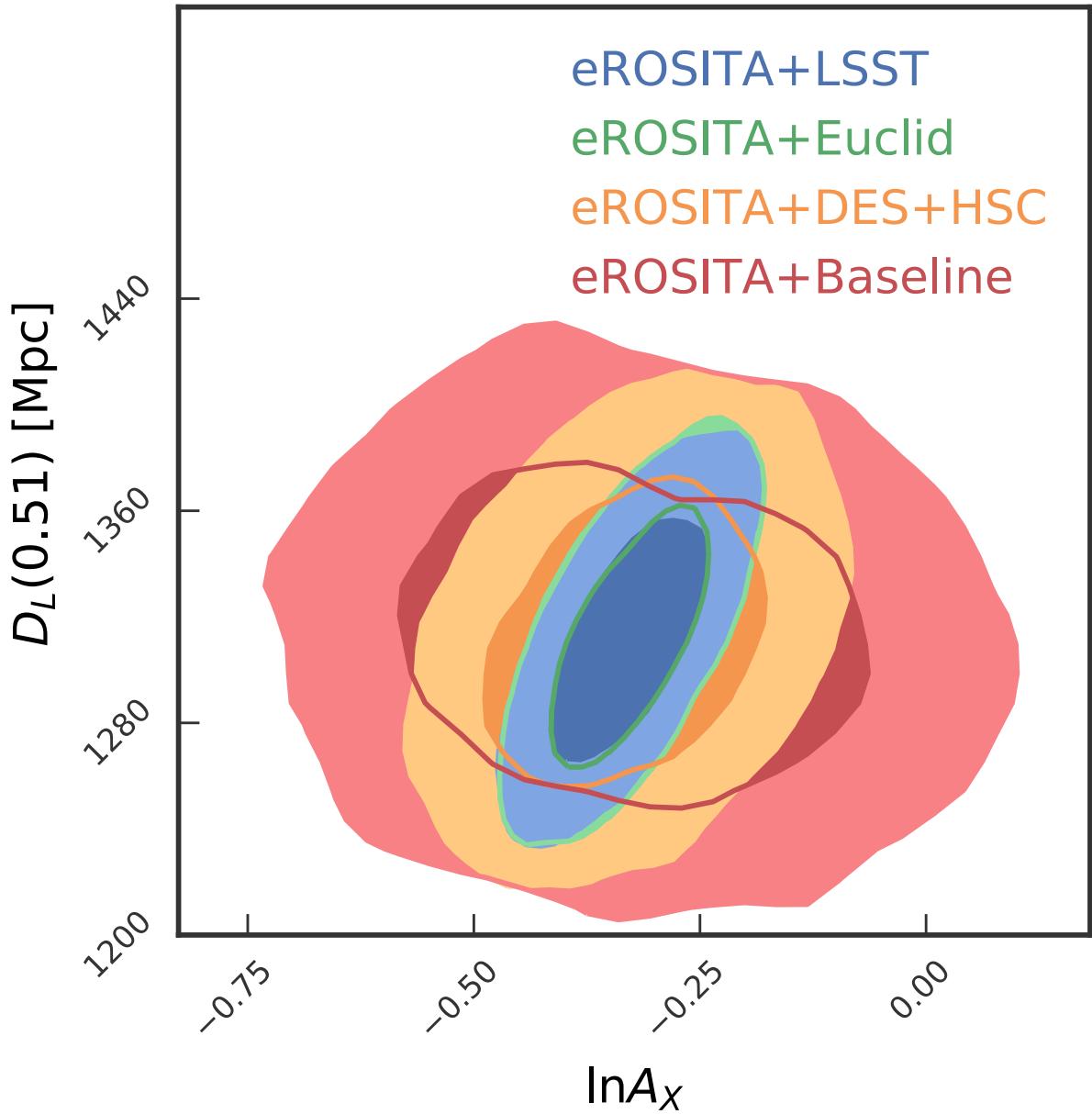


Abbildung 2.4: Two dimensional marginalized posterior sample of the amplitude of the scaling relation  $A_X$  and the luminosity distance to the median redshift of our sample  $D_L(0.51)$  in Mpc, as derived from the cosmological parameters in the posterior sample in the  $w$ CDM model. In red, orange, green and blue we present the constraints from the number counts alone (eROSITA+Baseline), from number counts and DES+HSC WL calibration (eROSITA+DES+HSC), Euclid WL calibration (eROSITA+Euclid), and LSST WL calibration (eROSITA+LSST), respectively. When no direct mass information is present, as in the case of number counts only, the two quantities are not degenerate with each other. As mass information is added, the underlying parameter degeneracy between the amplitude of the X-ray observable mass relation and the cosmological distance information emerges.

### Limiting parameter degeneracy

We have studied the causes of the weaker improvement in  $w$  when calibrating with Euclid or LSST WL, and we have discovered an interesting degeneracy due to the  $w$  sensitivity of the distance. Remember that our WL calibration dataset consists of observations of the shear profiles and the redshift distributions of the background galaxies. To turn these into masses, one needs the cosmology sensitive angular diameter distances (see discussion below equation 2.5). Moreover, our selection observable is the eROSITA count rate (similar to X-ray flux) that is related to the underlying X-ray luminosity through the luminosity distance (see equation 2.7). This leads to a degeneracy between  $w$ , governing the redshift evolution of distances, and the amplitude and redshift trend of the selection observable–mass relation.

The degeneracy between  $w$  and  $(\ln A_X, \gamma_X)$  can be easily understood by considering the parametric form of the rate mass scaling relation in equation (2.7). Ignore for a moment the distance dependence of the mass. Then for a given redshift  $z$  and rate  $\eta$ , a shift in  $w$  leads to a shift in the luminosity distance  $D_L(z)$ , and, to a minor degree, to a shift in the co-moving expansion rate  $E(z)$ . Such a shift can be compensated by a shift in  $\ln A_X$  and  $\gamma_X$ , resulting in the same mass, and consequently the same number of clusters, making it indiscernible. The distance dependence of the shear to mass mapping and the power law dependence of the rate on mass leads to a somewhat different dependence, and so the parameter degeneracy is not catastrophic.

This effect is demonstrated in Fig. 2.4, where the joint posterior of the luminosity distance to the median cluster redshift  $D_L(0.51)$  and of the amplitude of the scaling relation  $\ln A_X$  is shown. In the case of no direct mass information, when we fit the number counts with priors on the scaling relation parameters, the median distance and the amplitude are uncorrelated. As one adds more mass information, e.g., the +DES-HSC WL, and +Euclid WL or +LSST WL cases, the underlying correlation between the median distance and the amplitude becomes apparent. This degeneracy provides a limitation to improving the  $w$  constraint from the number counts by means of mass calibration. Given that it affects the halo masses directly, and not only the WL signal, we expect these degeneracies to be present also in other mass calibration methods, although to a different extent, given the different scaling of the selection observables with mass.

As a side note, these degeneracies highlight the importance of fitting for mass calibration and number counts simultaneously and self consistently. A mass calibration done at fixed cosmology would miss these correlations and lead to underestimated uncertainties on the scaling relation parameters. More worrisome, modeling mass calibration by simply adopting priors on the observable mass scaling relation parameters would miss the underlying physical degeneracies altogether (e.g., Sartoris et al., 2016; Pillepich et al., 2018).

The degeneracies between the distance redshift relation and the scaling relation parameters in the mass calibration explain why the impact of WL mass calibration is weaker in the  $\nu$ - $w$ CDM model, compared to the  $\nu$   $\Lambda$ CDM model: in the latter  $w$  is kept fixed, and the redshift evolution of distances and critical densities is controlled predominantly by a single variable:  $\Omega_M$ . With one degenerate degree of freedom less, WL mass calibration can

put tighter constraints on  $\ln A_X$  and  $\gamma_X$  in the  $\nu$ - $\Lambda$ CDM than in the  $\nu$ - $w$ CDM model.

### 2.3.3 Synergies with Planck CMB

It is customary in observational cosmology to combine the statistical power of different experiments to further constrain the cosmological parameters. An important part of these improvements is due to the fact that each experiment has distinctive parameter degeneracies that can be broken in combination with constraints from another experiment. This is especially true for CMB temperature and polarization anisotropy measurements, which constrain the cosmological parameters in the early Universe, but display important degeneracies on late time parameters such as  $\Omega_M$ ,  $\sigma_8$  and  $w$  (for a recent study applicable to current CMB measurements, see Howlett et al., 2012). We will discuss in the following the synergies between the Planck cosmological constraints from temperature and polarization anisotropy (Planck Collaboration et al., 2016a) and those from the eROSITA cluster counts analysis.

#### $\nu$ - $w$ CDM constraints

In the  $\nu$ - $w$ CDM model, the CMB suffers from the so called *geometrical degeneracy* (Efstathiou and Bond, 1999), that arises because the CMB anisotropy primarily constrains the ratio of the sound horizon at recombination and the angular diameter distance to that epoch. As a consequence, for example, the current day expansion rate  $H_0$  is degenerate with the equation of state parameter  $w$ . This uncertainty in the expansion history of the Universe leads to large uncertainties on late time properties such as  $\Omega_M$  and  $\sigma_8$ . Addition of a late time probe that constrains these quantities allows one to break the degeneracies and put tighter constraints on  $w$ . This can be nicely seen for the case of eROSITA in Fig. 2.5, where the red CMB degeneracy between  $\sigma_8$  and  $w$  is broken by the addition of cluster information. The corresponding uncertainties are shown in Table 2.3.

While in this model the CMB alone is not able to determine  $\Omega_M$ , the addition of eROSITA number counts allows a constraint of  $\pm 0.019$ . Inclusion of WL mass information further reduces the uncertainty to 0.018, 0.014 and 0.013 for DES+HSC, Euclid and LSST, respectively. The uncertainty in  $\sigma_8$  is reduced from 0.065 when considering only the CMB, to 0.032 with number counts, 0.019 with number counts and DES+HSC WL, and 0.010 with number counts and Euclid, and 0.009 with LSST WL. Noticeably, the determination of the equation of state parameter  $w$  is improved from 0.242 from CMB data alone, to 0.087 when adding number counts. Even more remarkable is the fact that WL calibrated eROSITA constraints on  $w$  are only marginally improved by the addition of CMB information.

#### Constraints on sum of the neutrino masses

We showed earlier that cluster number counts, even when they are WL calibrated, provide little information about the sum of the neutrino masses in the regime  $< 1$  eV. On the other hand, the CMB posteriors on  $\sigma_8$  and  $\Omega_M$  are strongly degenerate with the neutrino mass,

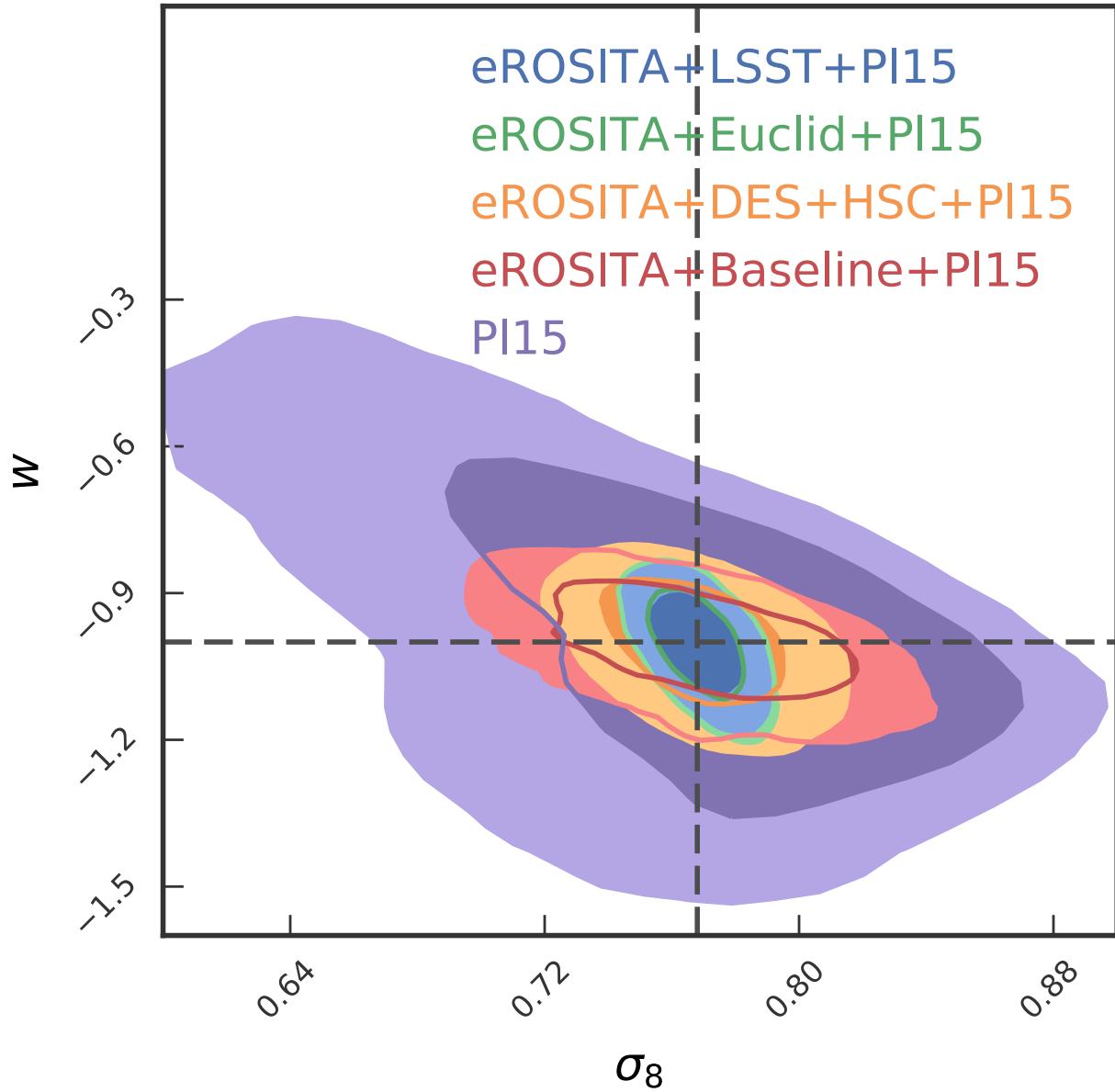


Abbildung 2.5: Marginalized posterior sample of  $\sigma_8$  and  $w$  in the  $w$ CDM model. In purple the constraints from Planck CMB alone (Pl15), in red the constraints from the number counts and Planck (eROSITA+Baseline+Pl15), in orange the constraints from the addition of DES+HSC WL calibration (eROSITA+DES+HSC+Pl15), in green for the addition Euclid WL calibration (eROSITA+Euclid+Pl15), in blue for the addition LSST WL calibration (eROSITA+LSST+Pl15). Cluster information breaks the inherent CMB degeneracies and allows to constrain the late time parameters to high precision.

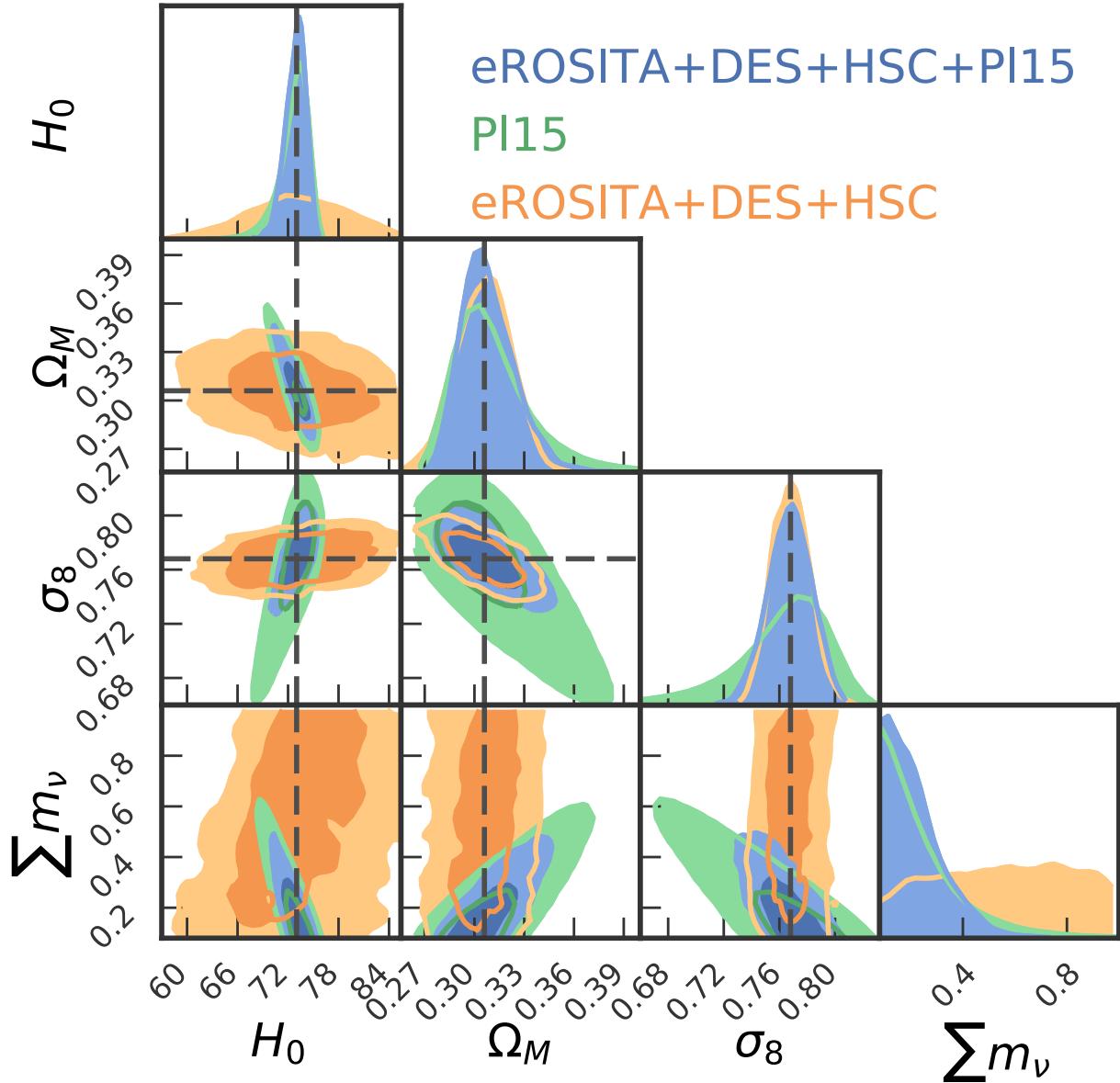


Abbildung 2.6: Marginalized posterior sample of  $H_0$ ,  $\Omega_M$ ,  $\sigma_8$  and  $\sum m_\nu$  in the  $\nu$ - $\Lambda$ CDM model. In red the constraints from Planck CMB alone (Pl15) and the constraints from eROSITA number counts and DES WL calibration without CMB priors in blue (eROSITA+DES), and with CMB priors in purple (eROSITA+DES+Pl15). By measuring  $\sigma_8$  and  $\Omega_M$  independently of the sum of neutrino masses, WL calibrated cluster number counts break the degeneracy among these parameters in the CMB posteriors.

as can be seen in Fig. 2.6. Contrary to the CMB, the number counts of galaxy clusters are only weakly affected by the sum of the neutrino mass. Recent studies have shown that the halo mass function is a function of the power spectrum of baryons and dark matter only (Costanzi et al., 2013; Castorina et al., 2014). Effectively, this means that number counts can be used to constrain the density  $\Omega_{\text{coll}}$  and fluctuation amplitude  $\sigma_{8,\text{coll}}$  of baryons and dark matter independently of the neutrino mass. If one considers matter as cold dark matter, baryons and neutrinos, as is customarily done, then  $\Omega_M = \Omega_{\text{coll}} + \Omega_\nu$  and  $\sigma_8^2 = \sigma_{8,\text{coll}}^2 + \sigma_{8,\nu}^2$ , where  $\Omega_\nu$  is the density parameter of neutrinos and  $\sigma_{8,\nu}^2$  is the amplitude of their clustering on  $8h^{-1}$  Mpc scales. The counts derived constraints on  $\Omega_{\text{coll}}$  and  $\sigma_{8,\text{coll}}$  then lead to only very weak degeneracies between the sum of the neutrino masses and  $\Omega_M$  and  $\sigma_8$ , respectively, because neutrinos constitute a tiny fraction of the total matter density and the total matter fluctuations on  $8h^{-1}$  Mpc scales. In Fig. 2.6 we can see how these very different parameter degeneracies in the CMB and cluster counts manifest themselves. Combining these weaker degeneracies arising from eROSITA+DES WL with the more pronounced degeneracies in the CMB posteriors allows us to break the latter and to better constrain the sum of the neutrino masses.

Consistently, we find that in the  $\nu$ - $\Lambda$ CDM model, the addition of CMB priors only marginally improves the constraints eROSITA will put on  $\sigma_8$  and  $\Omega_M$ . However, while the CMB alone puts an upper limit of  $\sum m_\nu < 0.514$  eV (at 95%) we determine that the combination of Planck CMB and eROSITA number counts will constrain the neutrino masses to  $< 0.425$  eV, which will improve to  $< 0.401$  eV,  $< 0.291$  eV and  $< 0.285$  eV with the addition of WL information from DES+HSC, Euclid and LSST, respectively.

### 2.3.4 Synergies with DESI BAO measurements

From the discussion in Section 2.3.2, it is apparent that the flux based X-ray selection and the distance dependent WL mass information lead to an inherent degeneracy between distances to the clusters and scaling relation parameters that ultimately limits the constraint on  $w$ . It would be desirable to utilize CMB independent constraints on the distance redshift relation, to allow for more stringent consistency checks between cluster derived constraints and CMB constraints. Some previous cosmological studies of X-ray clusters have used the distance information gleaned from the assumption of constant intracluster medium (ICM) mass fraction with redshift (Mantz et al., 2015; Schellenberger and Reiprich, 2017). While these results are encouraging, a challenge with this method is that it only provides accurate distance information if in fact the ICM mass fraction is constant at all redshifts. It has been established for decades now that the ICM mass fraction varies with cluster mass (e.g., Mohr, Mathiesen, and Evrard, 1999), but direct studies of how the ICM mass fraction varies over the redshift range of the eROSITA survey (i.e., extending beyond  $z = 1$ ) have only recently been undertaken (Lin et al., 2012; Chiu et al., 2016; Chiu et al., 2018; Bulbul et al., 2019). The evolution is consistent with constant ICM mass fraction, but the uncertainties are still large. Further study is clearly needed. Another interesting eROSITA internal prospect for better constraining the distance redshift relation is to utilize the clustering of clusters to determine the BAO scale (for a recent application,

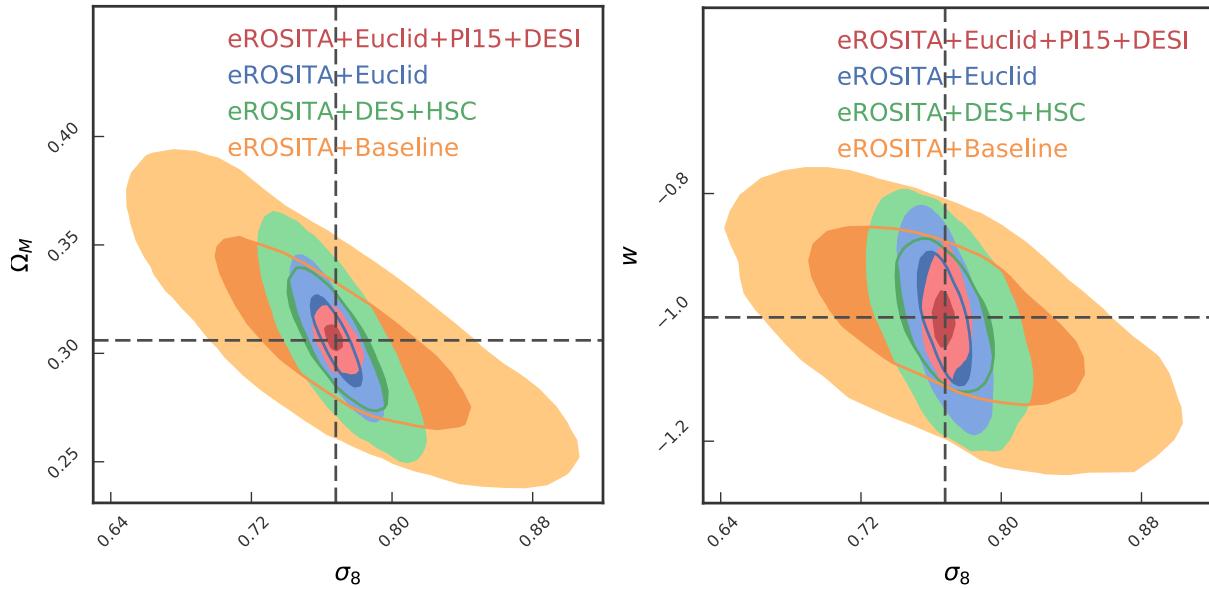


Abbildung 2.7: 2 dimensional marginal contours of the posteriors in  $(\Omega_M, \sigma_8)$  (left panel) and  $(w, \sigma_8)$  (right panel), showing the incremental improvement of constraining power when first adding WL information and second combining with external cosmological data sets (“Pl15” stands for the CMB fluctuation measurements by the Planck satellite, while “DESI” refers only to the BAO constraints). These posteriors are derived while simultaneously marginalizing over the summed neutrino mass.

see Marulli et al., 2018, and references therein.)

As an alternative, we consider constraints from other low redshift experiments, more precisely the measurement of the Baryonic Acoustic Oscillations (hereafter BAO) in future spectroscopy galaxy surveys. In this work, we consider the forecast for the constraints provided by the Dark Energy Spectroscopy Instrument<sup>7</sup> (DESI; Levi et al., 2013) as the relative error on the transversal BAO measurement  $d_A/r_s$  and the radial BAO measurement  $H(z)r_s$  as functions of redshift, where  $d_A$  is the angular diameter distance,  $H(z)$  the expansion rate, and  $r_s$  is the sound horizon. The values adopted in this work are reported in Table V of Font-Ribera et al. (2014). Furthermore, we follow the authors indications and assume that in each redshift bin, the measurement error on the two quantities are correlated with correlation coefficient  $\rho = 0.4$ . Using this information we perform an importance sampling of the posterior samples presented above and summarize the resulting uncertainties in Table 2.3.

When considering the uncertainties on the different parameters obtained by sampling these observables, we find that the BAO measurement dominates the uncertainty on  $\Omega_M$ . The addition of number counts, or number counts and WL information does not lead to major improvements on this parameter either in  $\nu$ - $w$ CDM or in  $\nu$ - $\Lambda$ CDM. However, the uncertainty on the dark energy equation of state parameter  $w$  is reduced from 0.086 in the BAO only case, to 0.065 when adding just number counts, 0.054 and 0.047 when adding DES+HSC and Euclid WL information, respectively. Remarkably, eROSITA counts with BAO priors on the expansion history outperforms eROSITA counts with CMB priors when it comes to constraining the parameters  $\Omega_M$ ,  $\sigma_8$  and  $w$ , while simultaneously marginalizing over the summed neutrino mass. The latter is unconstrained by eROSITA+BAO, even when considering WL mass information. Furthermore, eROSITA+BAO allows us to measure the Hubble constant  $H_0$  to varying degrees of precision, depending on the quality of the WL data. While these constraints never go below the present precision from other methods (see for instance Riess et al., 2016), they will provide a valuable systematics cross-check (for an example of systematics in SNe Ia that impact local  $H_0$  measurements, see, e.g., Rigault et al., 2013; Rigault et al., 2015; Rigault et al., 2018).

### 2.3.5 Combining all datasets

It is current practice in cosmology to first test consistency of constraints from different data sets as a check on systematics and to then combine the constraints as possible to provide the most precise cosmological parameter constraints possible. In the case of a forecast work like this, agreement is guaranteed by the choice of input cosmology for the mock creation, while statistical independence can be assumed for eROSITA with WL data, DESI and the CMB measurement from Planck.

We provide the results of this combination at the bottom of Table 2.3. In  $\nu$ - $w$ CDM, already the combination of Planck CMB measurements and DESI BAOs allows us to determine  $\Omega_M$  and  $w$  to 0.007 and 0.049, respectively, while simultaneously putting an upper

---

<sup>7</sup><https://www.desi.lbl.gov>

limit of  $< 0.284$  eV on the summed neutrino mass. Addition of eROSITA+Euclid WL only marginally improves these constraints to 0.005 and 0.047 for  $\Omega_M$  and  $w$ , respectively, and leads to the 95% confidence upper limit  $\sum m_\nu < 0.241$  eV. In this configuration, however, the added value of eROSITA number counts and WL mass calibration lies in the ability to constrain  $\sigma_8$ : while CMB and BAO put a constraint of 0.027, addition of eROSITA improves this to 0.026, 0.011 and 0.006, when considering the baseline mass information, DES+HSC WL, Euclid WL or LSST WL, respectively. In summary, using BAO and CMB priors together increases the constraining power of eROSITA cluster cosmology considerably, as can be seen in the shrinking of the 2 dimensional marginal contours in  $(\Omega_M, \sigma_8)$  and  $(w, \sigma_8)$  space, shown in the left and the right panel of Fig. 2.7, respectively.

### 2.3.6 Inclusion of low mass clusters and groups

In this work, we have taken the conservative approach of excluding all systems with a halo mass  $\lesssim 2 \times 10^{14} M_\odot$  by means of increasing the eROSITA cluster count rate threshold at low redshift (cf. Section 2.1.1 and Appendix ??). There are several good reasons to do so, all of them related, in one way or another, to an increase in systematic uncertainty when going to lower mass systems that are not as well studied. However, to enable comparison to previous work, and as a motivation to further investigate and control the systematic uncertainties in low mass clusters and groups, we also examine the impact of WL mass calibration on the constraining power for a cluster sample where the count rate threshold is reduced at low redshift so that only clusters with masses  $M_{500c} \lesssim 5 \times 10^{13} M_\odot$  are excluded.

#### Systematics of low mass clusters and groups

There are several important systematic concerns. For instance, Bocquet et al. (2016) find in a study using hydrodynamical structure formation simulations that for masses below  $10^{14} M_\odot$ , baryonic feedback effects reduce the halo mass function by up to 10% compared to halo mass functions extracted from dark matter only simulations. The magnitude of this effect depends on the feedback model, and therefore needs to be treated as a systematic uncertainty in the cosmological modeling. The magnitude of this uncertainty awaits further study.

Baryonic feedback effects also impact the mass profiles of clusters. Lee et al. (2018) show how active galactic nuclei feedback induces a deficit of mass in the cluster center when compared to gravity only simulations. The partial evacuation of baryons is strong enough to modify also the matter profile. Lee et al. (2018) demonstrate how this effect impacts the WL bias  $b_{WL}$  and the WL scatter  $\sigma_{WL}$ , making them mass dependent. Such effects will need to be taken into account, especially when considering lower mass systems.

Similarly, the thermodynamic structure of low mass systems, generally called groups, is more complex than for massive galaxy clusters, showing a larger impact of non gravitational physics (Eckmiller, Hudson, and Reiprich, 2011; Bharadwaj et al., 2014; Barnes et al., 2017). Lovisari, Reiprich, and Schellenberger (2015) showed that the mass slope of the luminosity mass relation is significantly steeper for groups than for clusters. Schellenberger

Tabelle 2.4: Parameter uncertainties, for number counts (eROSITA+Baseline), number counts and DES+HSC WL calibration (eROSITA+DES+HSC), number counts and Euclid WL calibration (eROSITA+Euclid), and number counts and LSST WL calibration (eROSITA+LSST) in the  $\nu$ -wCDM model when including low mass clusters. The units of the column “ $\sum m_\nu$ ” and “ $H_0$ ” are eV and  $\text{km s}^{-1} \text{Mpc}^{-1}$ , respectively. *Comments:* a) This parameter is not constrained within the prior ranges. When reporting upper limits “<”, we refer to the 95th percentile, while lower limits “>” refer to the 5th percentile. When a parameter is kept fixed in that model, we use “–”.

	$\Omega_M$	$\sigma_8$	$w$	$\sum m_\nu$	$H_0$	$\ln A_X$	$B_X$	$\gamma_X$	$\sigma_X$	$B'_X$
optimal mass calibration						0.028	0.021	0.050		0.116
eROSITA + WL										
$\nu$ -wCDM	priors					0.23	0.17	0.42	0.11	0.78
	eROSITA+Baseline	0.025	0.038	0.079	a)	8.081	0.113	0.071	0.202	0.078
	eROSITA+DES+HSC	0.012	0.012	0.069	a)	4.572	0.081	0.028	0.097	0.052
	eROSITA+Euclid	0.009	0.007	0.056	a)	3.762	0.042	0.019	0.073	0.027
	eROSITA+LSST	0.007	0.006	0.050	a)	2.707	0.042	0.016	0.068	0.023
eROSITA + WL + Pl15 (TTTEE_lowTEB)										
$\nu$ -wCDM	priors (incl. CMB)	<0.393	0.063	0.242	<0.667	>62.25	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.017	0.028	0.078	<0.580	2.745	0.131	0.026	0.128	0.083
	eROSITA+DES+HSC	0.010	0.012	0.069	<0.542	1.587	0.092	0.017	0.102	0.052
	eROSITA+Euclid	0.007	0.006	0.060	<0.381	1.401	0.046	0.013	0.076	0.021
	eROSITA+LSST	0.006	0.005	0.051	<0.365	1.317	0.045	0.012	0.065	0.021
eROSITA + WL + DESI (BAO)										
$\nu$ -wCDM	priors (incl. BAO)	0.007	a)	0.086	a)	a)	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.006	0.016	0.051	a)	1.703	0.136	0.036	0.090	0.068
	eROSITA+DES+HSC	0.006	0.009	0.048	a)	1.425	0.080	0.025	0.084	0.050
	eROSITA+Euclid	0.005	0.005	0.038	a)	1.379	0.036	0.016	0.063	0.021
	eROSITA+LSST	0.004	0.005	0.038	a)	1.303	0.036	0.014	0.061	0.021
eROSITA + WL + DESI + Pl15										
$\nu$ -wCDM	priors (incl. CMB+BAO)	0.007	0.027	0.049	<0.284	1.118	0.23	0.17	0.42	0.11
	eROSITA+Baseline	0.005	0.015	0.046	<0.279	1.114	0.134	0.022	0.079	0.067
	eROSITA+DES+HSC	0.005	0.010	0.044	<0.242	1.040	0.078	0.014	0.067	0.049
	eROSITA+Euclid	0.005	0.005	0.037	<0.237	1.015	0.039	0.012	0.058	0.021
	eROSITA+LSST	0.004	0.005	0.034	<0.224	0.790	0.039	0.010	0.053	0.021

and Reiprich (2017) demonstrate how such a break in the power law might bias the cosmological results derived from an X-ray selected cluster sample. We have thus chosen the conservative approach of excluding these systems from our primary eROSITA forecasts, thereby reducing the sensitivity of the forecast cosmological parameter constraints to these important complications at low masses.

### Improvement of the constraints

Nevertheless, the controlled environment of mock data analysis allows us to investigate how much constraining power could ideally be gained by lowering the mass limit if all the above described systematics were well understood and controlled. To this end, we select a low mass sample by imposing an observable selection with redshift that enforces  $M_{500c} \gtrsim 5 \times 10^{13} M_{\odot}$ , assuming that the scaling relation and the mass function used for the fiducial sample are still valid also at this lower mass scale. This increases the sample size to 43k clusters, with a median redshift  $\bar{z} = 0.31$  and a median halo mass of  $\bar{M}_{500c} = 1.4 \times 10^{14} M_{\odot}$ . The resulting constraints on the parameters of the  $\nu$ - $w$ CDM model are shown in Table 2.4. The constraints both on the cosmological parameters, as well as on the scaling relation parameters show a strong improvement compared to those from the higher mass sample. For eROSITA number counts we determine that the uncertainty on  $\Omega_M$ ,  $\sigma_8$  and  $w$  will be reduced by factors of 1.3, 1.4 and 1.3, respectively. When calibrating masses with DES+HSC, we find improvements of factor 1.9, 1.4 and 1.2, when considering Euclid the inclusion of low mass systems will reduce the uncertainties by 1.8, 1.7 and 1.3, while using LSST leads to reductions by 2.0, 1.7 and 1.4. In absolute terms, eROSITA including low mass systems, calibrated with Euclid will provide constraints on  $\Omega_M$ ,  $\sigma_8$  and  $w$  of 0.009, 0.007 and 0.056, respectively. We emphasize that these tight constraints can only be obtained if the aforementioned systematic effects are adequately controlled.

## 2.4 Discussion

The above presented results on the constraining power of the eROSITA cluster sample demonstrate its value as a cosmological probe. They also underline the crucial impact of WL mass calibration on the constraining power of cluster number counts. However, they also give some clear indications of how this impact manifests itself in detail.

In the following subsections we discuss first how the constraints on the scaling relation parameters are affected by the addition of better WL data, by the choice of the model and by the choice of cosmological priors, resulting in an assessment of the conditions under which we can attain an optimal mass calibration. We then determine the sensitivity of our observable to the different input parameters. Finally, we compare our prediction to the constraints from current and future experiments.

### 2.4.1 Impact of WL on scaling relation parameters

In the previous section we discussed in detail the impact of WL mass calibration on the eROSITA cosmological parameter constraints. Naturally, adding WL information will also improve the constraints on the scaling relation parameters. The resulting uncertainties are reported in Table 2.3. In the following, we will focus on two interesting aspects of these results: first, we assess under which circumstances eROSITA will be optimally calibrated; second, we comment on the constraints on the scatter in observable at fixed mass.

#### Which mass calibration is optimal?

In Section 2.3.1, we introduced the concept of the *optimal* mass calibration. Comparing the bounds on the parameter uncertainties derived there to the forecasts for DES+HSC, we find that, independent of the presence of external cosmological priors and in both models we consider, DES WL will not provide an optimal calibration of the eROSITA observable mass relation. Only the calibration of the mass slope  $B_X$  when considering CMB and BAO data is an exception to this. This is not to say that, as shown above, the inclusion of DES+HSC WL information does not improve the cosmological constraints. It is to say that some part of the information contained in the number counts is used to constrain the scaling relation parameters instead of the cosmological parameters.

The optimal nature of the Euclid or LSST mass calibration is more subtle. When the dark energy equation of state parameter is kept fixed in the  $\nu$ - $\Lambda$ CDM model, Euclid provides an optimal mass calibration on the amplitude of the scaling relation, both with and without external cosmological priors from CMB or BAO observations. However, in the  $\nu$ - $w$ CDM model without external priors, Euclid or LSST WL does not constrain the scaling relation parameters optimally. The amplitude is calibrated optimally after the inclusion of BAO data. On the other hand, including CMB priors makes an optimal calibration of the mass trend possible. In the presence of dark energy with free but constant equation of state, the redshift slope is never calibrated optimally. Nevertheless, as demonstrated in the previous section, even in the limit of sub-optimal mass calibration, the eROSITA dataset provides cosmological information complementary to these other cosmological experiments. Furthermore, the calibration of the redshift trend could be improved by complementary direct mass calibration methods. At high redshift, the most promising options would be pointed observations of high- $z$  clusters (Schrabback et al., 2018a; Schrabback et al., 2018b) and CMB-WL calibration (Baxter et al., 2015; Planck Collaboration et al., 2016b; Baxter et al., 2018).

#### Scatter in the count rate to mass relation

One may imagine that the inclusion of low scatter mass proxies in the number counts and mass calibration analysis may tighten the constraints on the scatter and thereby reduce the uncertainties on the cosmological parameters. Our present work does not seem to support this hypothesis. First, we show that even an arguably weak constraint on the scatter can be considered an optimal calibration (cf. Section 2.3.1). In other words, even

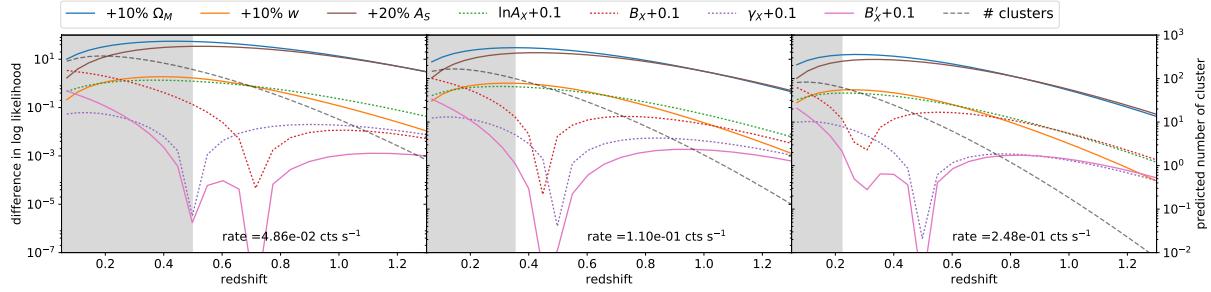


Abbildung 2.8: Sensitivity in terms of change in log likelihood of the number counts likelihood to various parameters as a function of redshift. From left to right, each panel represents a higher count rate bin. The total number of clusters for the fiducial parameter values is shown as a dashed line. The parameters are varied from the fiducial values as noted at the top of the figure. The grey area shows the redshift range where we exclude low mass clusters by raising the selection threshold. Notably, we find that the number counts likelihood is most sensitive to the parameters  $\Omega_M$  and  $A_S$  with comparable sensitivity to  $w$  and  $A_X$ .

at fixed cosmology, the number counts are unable to constrain the scatter. Consequently, our ability to constrain the cosmology using the number counts is not expected to depend strongly on the knowledge of the scatter. This can also be seen by the fact that there is little correlation between the scatter and any other parameter of interest in the  $\nu$ - $w$ CDM posterior sample, as shown in Fig. 2.3 and Fig. 2.10. We conclude that constraining the scatter to high precision, although of astrophysical interest, is not required to perform an optimal cosmological analysis.

Furthermore, our results indicate that DES+HSC, Euclid and LSST WL mass calibration will be able to determine the scatter to 0.062, 0.034, and 0.030, respectively (see Table 2.3). This may seem surprising, because WL mass calibration has large observational uncertainties and a large intrinsic scatter when compared to typical low scatter mass proxies such as the ICM mass or temperature. However, the final constraining power stems in our analysis from the large number of cluster with WL information and the relatively small prior uncertainty on the intrinsic WL scatter  $\sigma_{WL}$ . In summary, given that the knowledge of the scatter does not impact the constraints on the cosmological parameters, and that WL mass calibration is able to constrain the scatter directly, it is not clear that a dedicated scatter calibration through the inclusion of low scatter mass proxies like the ICM mass will significantly impact eROSITA cluster cosmology constraints. Further study would be required to confirm this.

### 2.4.2 Parameter sensitivities

To investigate in more detail how our observables—i.e. the number counts of clusters as a function of rate and redshift together with the WL mass calibration information—depend on the model parameters, we perform the following experiment: we vary the model

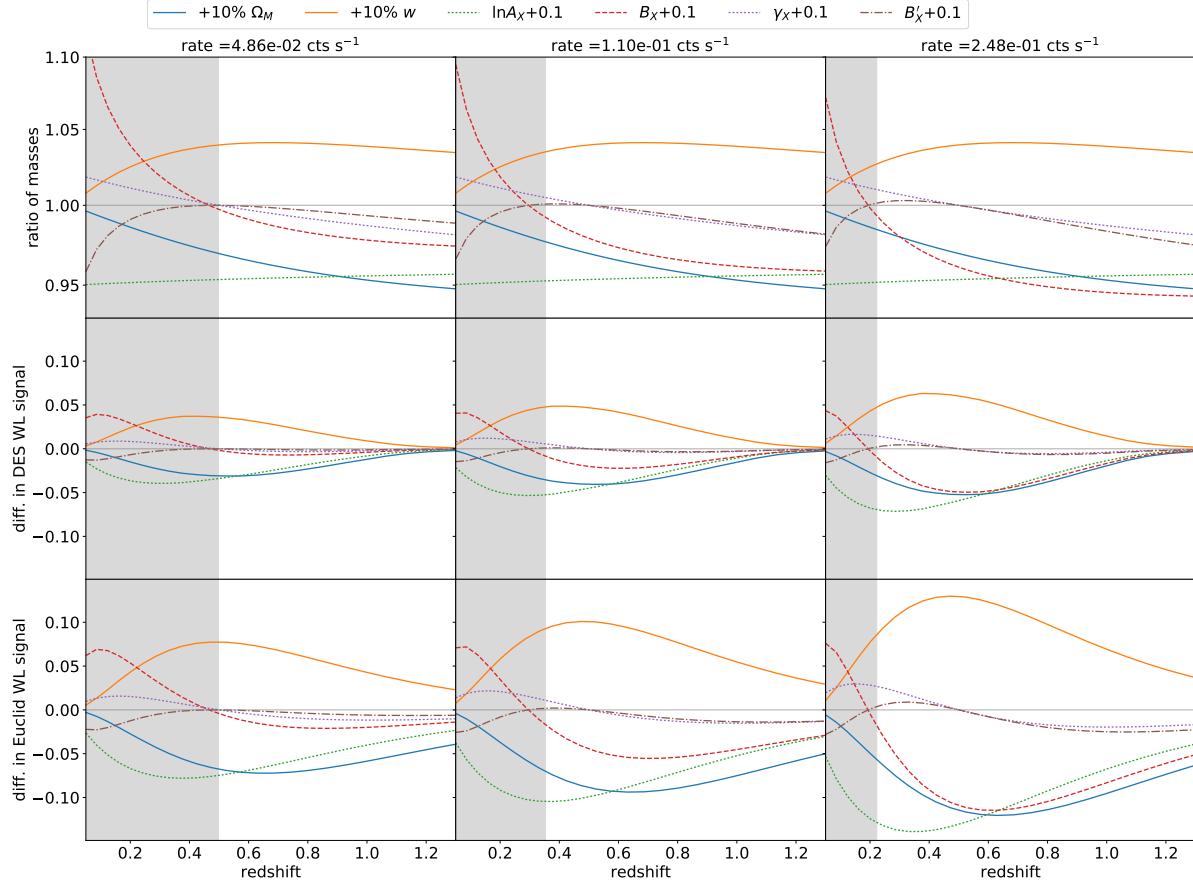


Abbildung 2.9: Sensitivity of different mass observables to the parameters considered in this work. On the x-axis, we plot the redshift and each column represents a different count rate. The parameters are varied around the input values. The grey area shows the observable range which is excluded by the approximate mass cut. From the top, the first row shows the fractional change in mass. The second and third rows show the difference in tangential shear for a single cluster, weighted by the observational WL uncertainty for a single cluster at that redshift, for DES and Euclid, respectively. We also see that both for the halo masses and for the shear signal,  $\Omega_M$  and  $w$  lead to changes comparable to the change in amplitude  $\ln A_X$  and  $\gamma_X$ . We conclude that these parameters must be degenerate with each other.

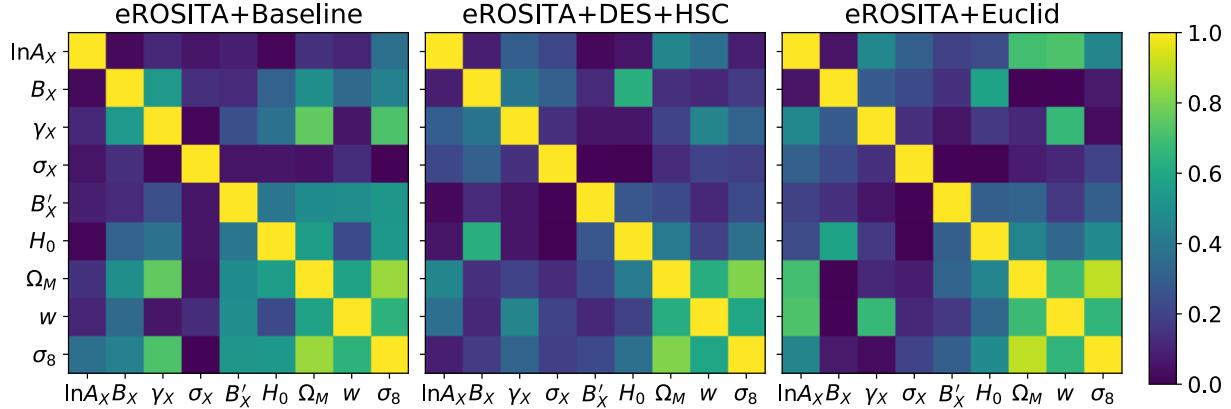


Abbildung 2.10: Absolute values of the correlation matrices of the posterior samples in the  $w$ CDM model, for number count (eROSITA+Baseline), number counts with DES+HSC WL information (eROSITA+DES+HSC) and number counts with Euclid WL information (eROSITA+Euclid). Noticeably, we find that the initial correlations between the pairs  $(\Omega_M, \sigma_8)$  and  $(B_X, \gamma_X)$  (in eROSITA+Baseline) is gradually broken by the addition of better mass information (eROSITA+DES+HSC and eROSITA+Euclid). However, the better the mass information, the clearer the inherent correlations between  $w$ ,  $\Omega_M$ ,  $A_X$  and  $\gamma_X$ . They indicate the degeneracies among these parameters stemming from the cosmology dependence of the rate mass mapping, as discussed in Section 2.4.2.

parameters one by one and examine how the number counts, the masses and the WL signals change. The results of this test are shown in Figs. 2.8 and 2.9. At three different fixed rates (increasing from left to right in the three columns), we investigate the sensitivity as a function of redshift with respect to the input parameter of the likelihood of the number counts (Fig. 2.8), as well as the masses, and the WL signals (Fig. 2.9). We grey out the part of rate-redshift space that is rejected due to our mass cut.

### Number counts

Fig. 2.8 shows the sensitivity of the number counts with respect to shifts in the input parameters. We decide here to plot the difference in log likelihood between the fiducial number counts  $N_{\text{fid}}$  and the number counts  $\tilde{N}$  if one parameter is varied. The difference in log likelihood in each bin reads

$$\delta \ln L = N_{\text{fid}} \ln \left( \frac{N_{\text{fid}}}{\tilde{N}} \right) - N_{\text{fid}} + \tilde{N}, \quad (2.21)$$

which can be simply obtained by taking the Poisson log likelihoods in that bin. We find that the number counts are most sensitive to the parameters  $\Omega_M$  and  $A_S$ . The sensitivity to the parameters  $w$ ,  $A_X$ ,  $B_X$ , and  $\gamma_X$  is much lower. This is reflected also in our results for the parameter uncertainties (Tables 2.3 and 2.4). The number counts do put tighter constraints on  $\Omega_M$  and  $\sigma_8$ , than on  $w$ , consistent with results from the first forecast studies

for large scale cluster surveys (Haiman, Mohr, and Holder, 2001; Holder, Haiman, and Mohr, 2001).

For comparison we also plot the total number of objects  $N_{\text{fid}}$  (dashed line), on a scale proportional to the difference in log likelihood. We can readily see that the difference in log likelihood is not simply proportional to the number of objects: the rarer, higher redshift, and consequently, at fixed rate, higher mass objects contribute more log likelihood per cluster than the lower redshift, lower mass systems. This trend is especially true for the constraints on  $\Omega_M$  and  $\sigma_8$  ( $A_S$ ), as noted in previous studies of cluster number counts (Haiman, Mohr, and Holder, 2001; Majumdar and Mohr, 2004). The sensitivity to 10% shifts in  $w$  and  $A_X$  are comparable. The more similar the shapes of the sensitivity curves for two parameters, the stronger the parameter degeneracy one could expect between those parameters.

### Masses and WL observables

The first row of Fig. 2.9 shows how much the masses are impacted by changes in input parameters. To this end, we plot the ratio between the input mass and the mass determined at the shifted parameters. In the range of interest for our study, the white area, we find that all parameters (except for  $A_S$ , of course, which we do not include in this figure) have a comparably large impact on the masses. Most remarkably, both shifts in  $\Omega_M$  and  $w$  change the masses associated to a given rate and redshift. This is because the rate mass relation has a strong distance dependence and also some critical density dependence. Both  $\Omega_M$  and  $w$  alter the redshift dependence of distances and critical densities.

More precisely, the shift to more positive  $w$  leads to a shift to higher masses, which mirrors the effect of changing the amplitude of the scaling relation  $\ln A_X$  and the redshift slope  $\gamma_X$ . Similarly, the redshift dependent mass shift induced by  $\Omega_M$  could be compensated by a shift in the redshift slope  $\gamma_X$  and  $\ln A_X$ . We therefore conclude that within the context of the masses corresponding to a fixed eROSITA count rate, the parameters  $w$  and  $\Omega_M$  are degenerate with combinations of  $\ln A_X$  and  $\gamma_X$ . This degeneracy impacts the predicted halo masses. The mass slope parameter  $B_X$ , however, seems to impact the masses in a distinctively different way, leading to no obvious parameter degeneracy. The same can be said for its redshift trend  $B'_X$ .

In our main experiment, we do not consider perfect halo masses, but WL signal. Therefore, we explore also the sensitivity of the WL signal for a single cluster to the input parameters. For the sake of simplicity, we do not consider the entire profile, but just assume one large radial bin spanning the fixed metric range corresponding to 0.25 – 5.0 Mpc in our fiducial cosmology. Given the constant metric size of the area considered, the WL measurement uncertainty for a single cluster due to shape noise can be computed by considering the background source density as a function of cluster redshift  $n_\epsilon(z_{\text{cl}}) = n_\epsilon(z_s > z_{\text{cl}} + 0.1)$ . In addition, the mapping from halo mass to tangential shear is non-linear and cosmology dependent. Consequently, the shear signal associated with a given rate and redshift is expected to have strong dependencies on cosmological parameters and, through the mass, also on the scaling relation parameters.

We visualize these effects in the second and third rows of Fig. 2.9 by plotting the difference between the WL signal for a single cluster in the fiducial model and the shifted model, divided by the expected magnitude of the shape noise for a single cluster. Indeed, one can readily see how the sensitivity per cluster of DES WL (second row) is generally lower, but also decreases more quickly with redshift than the sensitivity of Euclid WL (third row). This is due to the larger Euclid source galaxy sample and its extension to higher redshift as compared to DES. The trends we discuss above for the difference in halo mass do apply also to the sensitivity of the WL signal as a function of redshift.

We find the same degeneracies in the covariance matrices of our posterior samples in the  $\nu$ - $w$ CDM model for the three cases of eROSITA+baseline, eROSITA+DES+HSC, and eROSITA+Euclid, shown in Fig. 2.10. In the case of number counts alone, we find a strong correlation between the pairs  $(\Omega_M, \sigma_8)$  and  $(B_X, \gamma_X)$ . The latter degeneracy is strongly reduced by the addition of WL mass information, and is not present in the case of Euclid WL calibration. This is due to the fact that WL is quite sensitive to  $B_X$ . This is in line with improvements of both the  $(\Omega_M, \sigma_8)$  and  $B_X$  constraints when adding WL mass information. However, when  $w$  is free to vary, the degeneracies between  $w$ ,  $\ln A_X$  and  $\gamma_X$  lead to stronger correlations between these parameters for better mass information. They are most pronounced in the case of number counts with Euclid WL mass calibration.

### 2.4.3 Comparison to previous work

Finally, we compare our results to the constraints of recent and future experiments, with the intention of exploring how competitive eROSITA will be.

#### Current probes

The most up to date number counts analysis of an X-ray selected sample with WL mass calibration has been presented by Mantz et al. (2015, called Weighing the Giants, hereafter WtG). It consists of 224 clusters, 51 of which have a WL mass measurement, and 91 of which have ICM mass measurements. The analysis method is similar to the one described in this paper, with the exception that we did not consider cosmological constraints from the measurement of the ICM mass fraction. In the  $w$ CDM model (i.e. fixing the neutrino mass), when considering only X-ray and WL data, the uncertainties on  $\Omega_M$ ,  $\sigma_8$  and  $w$  are 0.036, 0.031 and 0.15, respectively. The direct comparison to our work is made difficult by the addition of the distance sensitive gas fraction measurements, which by themselves constrain  $\delta\Omega_M = 0.04$  and  $\delta w = 0.26$  (Mantz et al., 2014). This measure clearly dominates the error budget on  $\Omega_M$  and provides valuable distance information. Nevertheless, eROSITA cluster cosmology is evenly matched with WtG when considering just the number counts. It will outperform the constraining power of WtG when calibrated with DES+HSC WL information. In the case of LSST WL calibration, we project that the uncertainties on  $\Omega_M$ ,  $\sigma_8$  and  $w$  are smaller by factors 2.6, 3.1 and 2.1, respectively. These projections ignore distance information from the eROSITA clusters and AGN, which would further improve the constraints.

Another recent cluster cosmology study has been presented by de Haan et al. (2016). Therein, the cosmological constraints from 377 Sunyaev-Zeldovich selected clusters detected by the South Pole Telescope (hereafter SPT) above redshift  $> 0.25$  are determined. From the number counts alone, the dark energy equation of state parameter is constrained to a precision of  $\delta w = 0.31$ , which is a factor 3.1 worse than our prediction for the number counts from eROSITA alone. Furthermore, de Haan et al. (2016) find  $\delta\Omega_M = 0.042$  and  $\delta\sigma_8 = 0.039$ , while keeping the summed neutrino mass fixed at its minimal value. By comparison, in the baseline configuration eROSITA will improve the constrain on  $\Omega_M$  and  $\sigma_8$  by a factor 1.5 and 1.2, however while marginalizing over the summed neutrino mass. Also note that the priors used for the de Haan et al. (2016) analysis encode the mass uncertainty over which Bulbul et al. (2019) marginalized when deriving the uncertainties on the X-ray scaling relation parameters we employ as our eROSITA+Baseline.

When the SPT number counts are combined with the CMB constraints from Planck, de Haan et al. (2016) report constraints on  $\sigma_8$  and  $w$  of 0.045 and 0.17 respectively. We find that eROSITA number counts alone, in combination with Planck, will do better by a factor 2.8 on  $\sigma_8$  and a factor 2.0 on the equation of state parameter  $w$ , while additionally marginalizing over the summed neutrino mass. These numbers improve even more, if we consider the WL mass calibration by DES+HSC, Euclid and LSST.

Comparing our forecasts on the improvement of the upper limit on the summed neutrino mass to previous results from the combination of Planck CMB measurements with either SPT cluster number counts or WtG is complicated by several factors. First, we consider the full mission results for Planck (Planck Collaboration et al., 2016a), while SPT (de Haan et al., 2016) used the half mission data (Planck Collaboration et al., 2014) in addition to BAO data, and WtG (Mantz et al., 2015) additionally added ground based CMB measurements and supernova data. SPT reports the measurement  $\sum m_\nu = 0.14 \pm 0.08$  eV, which is impacted to some degree by the statistically insignificant shift between their constraint and the CMB constraints in the  $(\Omega_M, \sigma_8)$  plane. Comparison to this result is complicated by our choice to use the minimal neutrino mass as input value. On the other hand, WtG reports  $\sum m_\nu \leq 0.22$  at 95% confidence, which is comparable with our result from eROSITA number counts, DES+HSC WL, Planck CMB and DESI BAO.

The latest cosmological constraints from measurements of the Large Scale Structure (LSS) of the Universe were presented by the Abbott et al. (2018) for the first year of observations (Y1), where the joint constraints from the cosmic shear and photometric galaxy angular auto- and crosscorrelation functions are derived. In the  $\nu$ - $w$ CDM model, the uncertainties on  $\Omega_M$ ,  $\sigma_8$  and  $w$  are 0.036, 0.028 and 0.21, respectively. This is better than the constraints from eROSITA number counts alone, except for the dark energy equation of state parameter, which will be constrained better by eROSITA. However, utilizing DES+HSC to calibrate the cluster masses, we forecast that eROSITA will outperform the DES-Y1 analysis. In combination with Planck CMB data, DES-Y1 puts a 95% upper limit of 0.62 eV on the sum of the neutrino masses, whereas we forecast an upper limit of 0.424 (0.401) when combining eROSITA number counts (and DES+HSC WL calibration) with Planck data. Considering that our DES WL analysis assumes year 5 data, it will be interesting to see whether the DES Y5 LSS measurements or eROSITA with DES WL

calibration will provide the tighter cosmological constraints.

As can be seen from Table 2.3, eROSITA will clearly outperform Planck CMB measurements on several cosmological parameters. In the  $\nu$ - $\Lambda$ CDM model, eROSITA with WL mass information will outperform Planck on the parameters  $\Omega_M$  and  $\sigma_8$ , and in the  $\nu$ - $w$ CDM eROSITA with WL case will also outperform Planck on the equation of state parameter  $w$ . However, for constraints on the sum of the neutrino mass, Planck alone offers much more than eROSITA alone. Given, however, that eROSITA and Planck extract their constraints at low redshift and high redshift, respectively, the true benefit of these two experiments lies in assessing the mutual consistency and thereby probing whether our evolutionary model of the Universe is correct. If this is the case, their joint constraints will tightly constrain the cosmological model, and provide improved constraints on the sum of neutrino masses.

### Previous forecasts for eROSITA

This work elaborates further on the forecast of the eROSITA cosmological constraints first presented in Merloni et al. (2012), and subsequently discussed in more detail in P18. The direct comparison to the latter is complicated by several diverging assumptions, including that we only consider the German half of the sky. Perhaps the most significant difference is their approach of using Fisher matrix estimation and modeling mass calibration as simply being independent priors on the various scaling relation parameters, whereas we have developed a working prototype for the eROSITA cosmology pipeline and used it to analyze a mock sample with shear profiles in a self-consistent manner.

Other differences include their use of different input scaling relations from older work at lower redshift and different fiducial cosmological parameters. P18 includes constraints from the angular clustering of eROSITA clusters, although these constraints are subdominant in comparison to counts except for parameters associated with non-Gaussianity in the initial density fluctuations (see Pillepich, Porciani, and Reiprich, 2012). In our analysis, we marginalize over the sum of the neutrino mass as well as relatively weak priors on  $\omega_b$  and  $n_s$ .

Following what P18 call the *pessimistic* case with an approximate limiting mass of  $5 \times 10^{13} M_\odot h^{-1}$ , they predict 89 k clusters, which is in good agreement with our forecast of 43 k clusters when including clusters down to masses of  $5 \times 10^{13} M_\odot$ . Under the assumption of a 0.1 % amplitude prior, 14 % mass slope prior and 42 % redshift slope prior, they forecast a constraint of 0.017, 0.014 and 0.059 on  $\sigma_8$ ,  $\Omega_M$  and  $w$ , respectively. P18 also consider an *optimistic* case, in which clusters down to masses of  $1 \times 10^{13} M_\odot h^{-1}$  are used under the assumption of 4 times better priors on the scaling relation parameters. For this case, the constraints on  $\sigma_8$ ,  $\Omega_M$  and  $w$  are 0.011, 0.008 and 0.037, respectively.

A quantitative comparison to our work is complicated by the fact that we find a constraint on the amplitude of the scaling relation (through direct modeling of the WL calibration from Euclid or LSST) that is worse than their *pessimistic* case, but our constraint on the mass and redshift trends is better than their optimistic case. Consistently, we predict tighter constraints of  $\sigma_8$  and  $\Omega_M$ , which are sensitive to the mass and redshift trends of the scaling relation, while we predict lower precision on  $w$ , which we demonstrate to

be degenerate with the amplitude of the scaling relation through the amplitude distance degeneracy. Important here is the realization that the observed shear profiles map into cluster mass constraints in a distance dependent fashion (this is true for all direct mass constraints; Majumdar and Mohr, 2003). It is not straightforward to capture this crucial subtlety by simply adopting priors on observable mass scaling relation parameters.

### Euclid cosmological forecasts

The Euclid survey will not only provide shear catalogs to calibrate the masses of clusters, but will also allow the direct detection of galaxy clusters via their red galaxies (Sartoris et al., 2016), and the measurement of the auto- and cross-correlation of red galaxies and cosmic shear (Giannantonio et al., 2012). For the optically selected Euclid cluster sample, Sartoris et al. (2016) forecast  $2 \times 10^6$  galaxy clusters with limiting mass of  $7 \times 10^{13} M_\odot$  up to redshift  $z = 2$ , yielding constraints on  $\Omega_M$ ,  $\sigma_8$ , and  $w$  of 0.0019 (0.0011), 0.0032 (0.0014), and 0.037 (0.034), respectively, when assuming no knowledge on the scaling relation parameter (perfect knowledge of the scaling relation parameters). Under these assumptions, the number counts and the angular clustering of Euclid selected clusters would outperform eROSITA cluster cosmology. Nevertheless, cross comparisons between the X-ray based eROSITA selection and the optically based Euclid cluster selection will provide chances to validate the resulting cluster samples.

Giannantonio et al. (2012) forecast that the auto- and cross-correlations between red galaxies and cosmic shear in the Euclid survey will provide constraints on  $\Omega_M$ ,  $\sigma_8$ , and  $w$  of 0.005, 0.033 and 0.050, respectively. Such a precision on  $\sigma_8$  would be achieved by the baseline eROSITA+Euclid analysis, too. However, to achieve similar precisions in  $\Omega_M$  and  $w$ , it would be necessary to consider eROSITA detected clusters down to masses of  $5 \times 10^{13} M_\odot$ .

## 2.5 Conclusions

In this work, we study the impact of WL mass calibration on the cosmological constraints from an eROSITA cluster cosmology analysis. To this end, we create a mock eROSITA catalog of galaxy clusters. We assign luminosities and ICM temperatures to each cluster using the latest measurements of the X-ray scaling relations over the relevant redshift range (Bulbul et al., 2019). Considering the eROSITA ARF, we then compute the eROSITA count rate for all clusters in this sample. We apply a selection on the eROSITA count rate, corresponding to a  $\sim 6\sigma$  detection limit given current background estimates, to define a sample for a cosmological forecast. This detection limit ensures both high likelihood of existence and angular extent, and – through raising the detection threshold at low redshift – also excludes low mass objects at low redshift. We assume all cluster redshifts are measured photometrically using red sequence galaxies (see discussion in, e.g. Klein et al., 2018; Klein et al., 2019). We forecast that in the  $14,892 \text{ deg}^2$  of the low Galactic extinction sky accessible to the eROSITA-DE collaboration, when raising the detection

threshold at low redshift to exclude clusters with  $M_{500c} \lesssim 2 \times 10^{14} M_{\odot}$ , we predict that eROSITA will detect 13k clusters. This baseline cosmology sample has a median mass of  $\bar{M}_{500c} = 2.5 \times 10^{14} M_{\odot}$  and a median redshift of  $\bar{z} = 0.51$ . For the case where we adjust the low redshift detection threshold to exclude clusters with  $M_{500c} \lesssim 5 \times 10^{13} M_{\odot}$ , we predict 43k clusters. This sample has a median mass  $\bar{M}_{500c} = 1.4 \times 10^{14} M_{\odot}$ , and a median redshift  $\bar{z} = 0.31$ . Both samples extend to high redshift with  $\sim 400$  clusters at  $z > 1$ .

We then analyze these mock samples using a prototype of the eROSITA cluster cosmology code that is an extension of the code initially developed for SPT cluster cosmology analyses (Bocquet et al., 2015; de Haan et al., 2016; Bocquet et al., 2019a). This code employs a Bayesian framework for simultaneously evaluating the likelihoods of cosmological and scaling relation parameters given the distribution of clusters in observable and redshift together with any direct mass measurement information. The scaling relation between the selection observable (eROSITA count rate) and the mass and redshift is parametrized as a power law with log-normal intrinsic scatter. Final parameter constraints are marginalized over the uncertainties (systematic and statistical) in the parameters of the mass-observable scaling relation.

We first estimate the optimal level of mass calibration necessary for the number counts of eROSITA clusters to mainly inform the constraints on the cosmological parameters. This requires a calibration of the amplitude of the mass observable relation at 4.2%, the mass trend of the scaling relation at 2.4%, and the redshift trend at 5.3%. These numbers are derived using current knowledge of the scatter around the mass luminosity relation. Furthermore, we determine that the mass trend of the rate mass relation has to be allowed to vary with redshift to enable the recovery of unbiased cosmological results.

We then examine cosmological constraints in three different cluster mass calibration contexts: (1) using ‘baseline’ constraints existing today that are taken from the recent SPT analysis of the X-ray luminosity and temperature mass relations (Bulbul et al., 2019), (2) using WL information from the DES+HSC survey and (3) using WL information from the future Euclid and LSST survey. For the subset of the two catalogs that overlap the DES, HSC, Euclid or LSST survey footprints, we produce tangential shear profiles with appropriate characteristics for these surveys. We also estimate the level of systematic mass uncertainties in the WL masses that would result from the data quality of these two surveys and from theoretical uncertainties in the impact of mis-centering and mis-fitting the shear profiles. We adopt mass uncertainties of 5.1%, 1.3% and 1.5% for DES+HSC, Euclid, and LSST, respectively. These levels of systematic mass uncertainty will require that our understanding of the theoretical mass bias from simulations be improved by factors of 2 and 5 for DES+HSC and Euclid/LSST, respectively, in comparison to current work (Dietrich et al., 2019). We note that achieving these improvements will require a significant investment of effort.

Throughout this work, we allow the summed neutrino mass to vary. All results are thus marginalized over the summed neutrino mass. In the  $\nu$ - $w$ CDM model, we forecast that eROSITA number counts will constrain the density of matter in the Universe  $\Omega_M$  to 0.032, the amplitude of fluctuation  $\sigma_8$  to 0.052, and the equation of state parameter of the dark energy  $w$  to 0.101. Calibrating the masses of eROSITA clusters with DES+HSC (Euclid;

LSST) WL will reduce these uncertainties to 0.023 (0.016; 0.014), 0.017 (0.012; 0.010), and 0.085 (0.074; 0.071), respectively. We also find that eROSITA clusters alone will not provide appreciable constraints on the sum of the neutrino masses.

eROSITA number counts will be able to break several degeneracies in current CMB constraints, especially on late time parameters such as  $\Omega_M$ ,  $\sigma_8$  and  $w$ . In combination with Planck constraints from the measurement of the angular auto- and crosscorrelation functions of CMB temperature and polarization anisotropies, we determine that eROSITA will constrain these parameters to 0.019, 0.032 and 0.087 when adopting ‘baseline’ priors on the scaling relation parameters. These uncertainties shrink to 0.018 (0.014; 0.013), 0.019 (0.010; 0.009) and 0.085 (0.074; 0.069) when calibrating the masses with DES+HSC (Euclid; LSST) WL information.

When considering the  $\nu$ - $\Lambda$ CDM model, the upper limit on the neutrino mass of 0.514 eV from CMB alone can be improved to a constraint of 0.425 eV when utilizing number counts with the ‘baseline’ priors, 0.404 eV when also considering DES WL calibration, and to 0.291 eV when calibrating with Euclid WL, and 0.285 eV when calibrating with LSST WL.

We find that the constraining power of eROSITA cluster cosmology, even when calibrated with high quality shear profiles, is limited by a degeneracy between the scaling relation parameters and the cosmological distance to the clusters. This degeneracy arises, because the luminosity distance is necessary to transform observed count rates into luminosities, whose absolute and redshift dependent scaling with mass needs to be fitted simultaneously with the cosmological parameters that alter the redshift distance relation. This leads to the assessment that even the Euclid or LSST WL mass calibration will, by itself, not reach what we have defined as optimal levels in the  $\nu$ - $w$ CDM model.

However, we demonstrate that, with the inclusion of BAO measurements that constrain the redshift distance relation, the Euclid or LSST WL dataset can be used to calibrate cluster masses at an optimal level. Considering DESI-like BAO measurements, we project that eROSITA with Euclid WL mass calibration will constrain  $\sigma_8$  to 0.005 and  $w$  to 0.047, while the uncertainty on  $\Omega_M$  will be dominated by the BAO measurement.

Furthermore, we investigate the impact of lowering the mass limit to  $M_{500c} \gtrsim 5 \times 10^{13} M_\odot$ . Given the larger number of low mass clusters or groups, the eROSITA counts with Euclid WL can optimistically be used to determine  $\Omega_M$  to 0.009,  $\sigma_8$  to 0.007, and  $w$  to 0.056, if these low mass systems are simple extrapolations of the high mass systems. The expected additional complexity of these low mass systems would have to be modeled, and this additional modeling would likely weaken the cosmological constraints.

In summary, WL mass calibration from DES+HSC, Euclid, and LSST will significantly improve cosmological constraints from eROSITA cluster number counts, enabling a precise and independent cross-check of constraints from other measurements. The constraining power on  $w$  suffers from an inherent degeneracy between the distance redshift relation and the scaling relation between the X-ray observable, mass and redshift. This degeneracy can be lifted by inclusion of other cosmological measurements, such as BAO or CMB measurements. In turn eROSITA cluster cosmology can break degeneracies in these other observations, underscoring the synergies between different cosmological experiments.

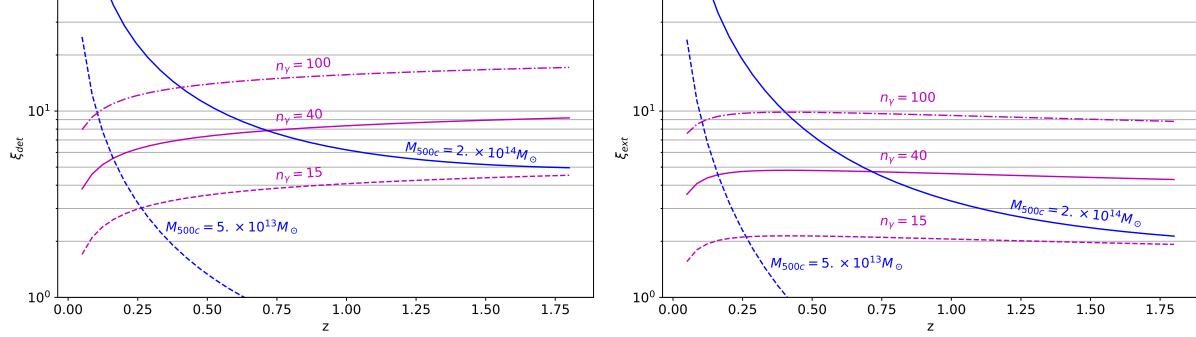


Abbildung 2.11: Estimated significance of detection (left panel) and significance of extent (right panel) as functions of redshift for a cluster with 15, 40, and 100 photons, and clusters of halo mass  $M_{500c} = 5 \times 10^{13} M_{\odot}$ , and  $2 \times 10^{14} M_{\odot}$  for median exposure time and background brightness. We find that 40 (15, 100) photon counts corresponds, at least, to a  $8\sigma$  ( $3\sigma$ ,  $\gtrsim 10\sigma$ ) detection and a  $3.5\sigma$  ( $2\sigma$ ,  $9\sigma$ ) significance of extent, rather independently of the cluster redshift.

## 2.6 Appendix

### 2.6.1 Comments on selection

In this work, we assume two selection criteria for our X-ray cluster sample:

- a cut in measured number of photons  $\hat{n}_{\gamma} > 40$  cts, which for the median eROSITA field with an exposure time of 1.6 ks translates into a measured rate cut  $\hat{\eta} > 2.5 \times 10^{-2}$  cts  $s^{-1}$ ,
- a cut in the observed mass

$$M_{\text{obs}} = M_0 e^{-\frac{\ln A_X}{B(z)}} \left( \frac{\hat{\eta}}{\eta_0} \right)^{\frac{1}{B(z)}} \left( \frac{E(z)}{E_0} \right)^{-\frac{2}{B(z)}} \left( \frac{D_L(z)}{D_L(z_0)} \right)^{\frac{2}{B(z)}} \left( \frac{1+z}{1+z_0} \right)^{-\frac{\gamma(z)}{B(z)}}, \quad (2.22)$$

which is derived from the rate-mass scaling relation equation (2.7). It is evaluated for the fiducial cosmology and the fiducial scaling relation parameters derived in Appendix 2.6.2. The cut  $M_{\text{obs}} > 2 \times 10^{14} M_{\odot}$ , or  $M_{\text{obs}} > 5 \times 10^{13} M_{\odot}$  is thus provided by a function of redshift, which is independent of cosmology and of the scaling relation parameters, and leads to the above mentioned cuts  $M_{500c} \gtrsim 2 \times 10^{14} M_{\odot}$  or  $M_{500c} \gtrsim 5 \times 10^{13} M_{\odot}$ . The low mass cut is thus implemented as a redshift dependent cut in observables (as for instance also in Vikhlinin et al., 2009a; Pillepich, Porciani, and Reiprich, 2012).

The cut at  $\hat{n}_{\gamma} > 40$  cts is justified by the following considerations on X-ray cluster detection. Detection of galaxy clusters hinges on the assumption that galaxy clusters are extended sources in the extragalactic X-ray sky, as discussed, for instance, by Vikhlinin

et al. (1998) in the case of the ROSAT PSPC, Pacaud et al. (2006) in the case of XMM-Newton, and Clerc et al. (2018) in the context of eROSITA. For this reason, their extraction is usually divided into two steps: first all X-ray sources are identified, then, among the identified source, those who are extended are selected. As outlined in Clerc et al. (2018), eROSITA will follow a similar procedure.

Following Pacaud et al. (2006), consider an X-ray image with a number of photons in each pixel  $i$  given by  $\hat{n}_i$ , and define the following three likelihoods:

- the likelihood that the image is simply background with background brightness  $\mu_{\text{bkg}}$ , which reads

$$\ln \mathcal{L}_{\text{bkg}} = \sum_i \hat{n}_i \ln(\mu_{\text{bkg}} A_i) - \mu_{\text{bkg}} A_i, \quad (2.23)$$

where  $A_i$  is the area of each pixel  $i$ ;

- the likelihood of being a point source centered in  $\mathbf{x}_{\text{ps}}$  with total number of photons  $n_{\text{ps}}$ , given by

$$\ln \mathcal{L}_{\text{ps}} = \sum_i \hat{n}_i \ln(n_{\text{ps}} \text{PFS}(\mathbf{x}_{\text{ps}})_i + \mu_{\text{bkg}} A_i) - n_{\text{ps}} \text{PFS}(\mathbf{x}_{\text{ps}})_i - \mu_{\text{bkg}} A_i \quad (2.24)$$

where  $\text{PFS}(\mathbf{x})_i$  is the value in the pixel  $i$  of the average survey point spread function (PSF) centered in  $\mathbf{x}$ ;

- the likelihood of being a cluster with total number of photons  $n_\gamma$ , modeled as the convolution of a  $\beta$ -model (Cavaliere and Fusco-Femiano, 1976) with the PSF, which, for a cluster position  $\mathbf{x}$  and cluster core radius  $\theta_c$ , reads

$$\ln \mathcal{L}_{\text{cl}} = \sum_i \hat{n}_i \ln(n_\gamma S(\mathbf{x}; \theta_c)_i + \mu_{\text{bkg}} A_i) - n_\gamma S(\mathbf{x}; \theta_c)_i - \mu_{\text{bkg}} A_i, \quad (2.25)$$

where  $S(\mathbf{x}; \theta_c)_i$  stands for the value of the PSF convolved beta-profile with center  $\mathbf{x}$  and core radius  $\theta_c$ .

For the median eROSITA field, we expect  $\mu_{\text{bkg}} = 3.6 \text{ cts arcmin}^{-2}$  (Clerc et al., 2018). To estimate the approximate significances of clusters with  $n_\gamma = (15, 40, 100)$  at different redshifts, we use the rate- mass relation derived in this work (c.f. Section ??). Furthermore, we assume that the core radius is given by  $\theta_c = 0.2\theta_{500c}$ , where  $\theta_{500c}$  is the angular extent of the radius inclosing an over density 500 times the critical density of the Universe. We take the PSF to be a gaussian with half energy width of 24 arcsec. Assuming  $\beta = 2/3$ , we create an X-ray 10-by-10 arcmin image of the expected number of photons  $\hat{n}$ , by computing  $\hat{n}_i = n_\gamma S(\mathbf{x}; \theta_c)_i + \mu_{\text{bkg}} A_i$ . We intentionally do not draw a Poisson realization of the model, in order to capture the mean behavior of the extraction procedure.

On this image, the three likelihoods are then maximized by varying  $(n_{\text{ps}}, \mathbf{x}_{\text{ps}})$ , and  $(n_\gamma, \theta_c, \mathbf{x})$ , respectively. We shall denote the maximum likelihood  $\ln \hat{\mathcal{L}}_\alpha$  for  $\alpha \in (\text{bkg}, \text{ps},$

cl). To create an analogy to the SZE case<sup>8</sup>, which can be interpreted as ‘sigmas’, both for detection  $\xi_{\text{det}} = \sqrt{2(\ln \hat{\mathcal{L}}_{\text{cl}} - \ln \hat{\mathcal{L}}_{\text{bkg}})}$ , and for extent  $\xi_{\text{ext}} = \sqrt{2(\ln \hat{\mathcal{L}}_{\text{cl}} - \ln \hat{\mathcal{L}}_{\text{ps}})}$ . Note that for well detected point sources, the best fit core radius  $\theta_c \approx 0$ , such that  $\ln \hat{\mathcal{L}}_{\text{cl}} \approx \ln \hat{\mathcal{L}}_{\text{ps}}$  and  $\xi_{\text{ext}} \approx 0$ , while  $\xi_{\text{det}} \gg 0$ . In contrast, for well detected extended sources,  $\xi_{\text{det}} \gg 0$  and  $\xi_{\text{ext}} \gg 0$ . For practical purposes, the region in which the likelihood is maximized is often pre-selected.

We repeat this exercise for different redshifts. The results are shown in Fig. 2.11. We find that a source with 40 (15, 100) photons will be detected, at least, at  $8\sigma$  ( $3\sigma$ ,  $10\sigma$ ) for redshifts above 0.5. Furthermore, we determine that these clusters will have a significance of extent of  $3.5\sigma$  ( $2\sigma$ ,  $9\sigma$ ). The significance of extension at, e.g., 15 or 40 photons, is rather low. Extent acts as a secondary selection on a sample which contains approximately 10% clusters and 90% point sources. Considering that  $3.5\sigma$  ( $2\sigma$ ) corresponds to a p-value of  $2.3 \times 10^{-4}$  ( $2.3 \times 10^{-2}$ ), the extent cut would create a cluster sample with an approximate contamination of 0.2% (17%) by point sources after the X-ray selection. With the optical followup of a tool like MCMF, any X-ray cluster candidate without an associated overdensity of red galaxies can be easily removed from the sample (Klein et al., 2018; Klein et al., 2019).

Fig. 2.11 also shows the redshift evolution of the significances for a cluster of fixed mass. The characterization of this evolution, especially its dependence on distance both through the flux and the angular extent, might be worth further investigation, as its knowledge would allow us to use the significance as a the primary X-ray observable. Such a study is currently limited by the somewhat simplistic assumption that the core radius is a multiple of the virial radius  $\theta_{500c}$ . Furthermore, a study of the measurement uncertainty on the significance would be necessary.

### 2.6.2 Comments on the count rate to mass relation

To obtain unbiased cosmological results, the parametric form of the scaling relation needs to provide accurate mass predictions over the mass and redshift range of interest for the sample considered. This is necessary, as systematic biases in mass lead to a systematic misestimation of the abundance of clusters in a given observable redshift range. These misestimations of cluster abundance will then be compensated by shifts in the cosmological parameters, which would generically lead to systematic biases in the best fitting cosmological parameters. Therefore, it is of paramount importance to utilize a parametric form for the scaling relation that has adequate freedom to describe the sample being modeled.

<sup>8</sup>In the SZE case, the likelihoods above take the form of  $\chi^2$  thanks to the Gaussian nature of the noise. Their maximization w.r.t. to the amplitude of the templates simplifies the problem to a maximization of the signal to noise as a function of scale and position. Maximizing the signal to noise is thus formally equivalent to maximizing the likelihood. As can readily be seen, this simplification does not apply to the Poissonian case of X-ray images. In both cases, however, we can define the maximum signal to noise, or significance, as  $\xi = \sqrt{2 \ln \hat{\mathcal{L}}}$ .

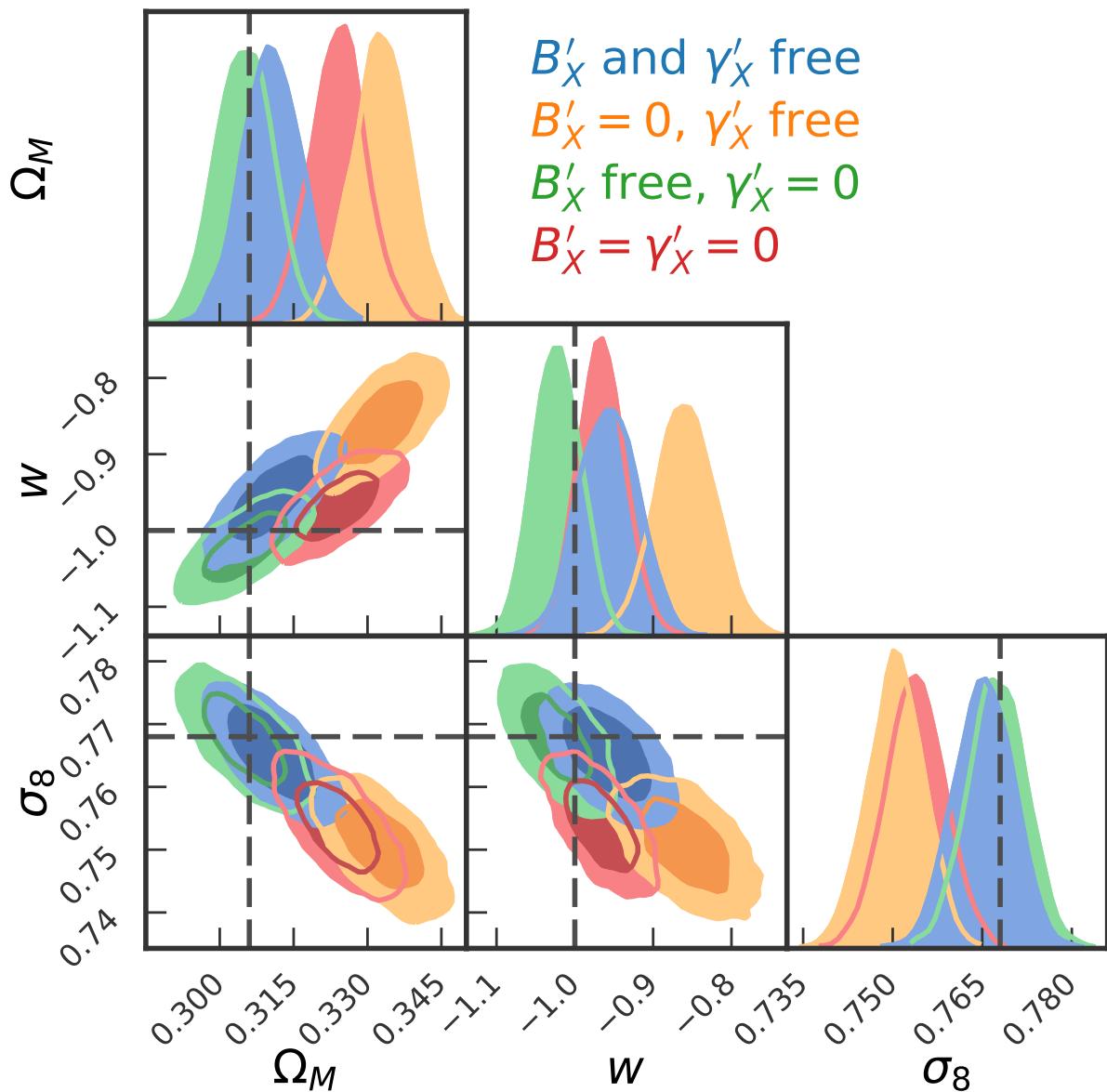


Abbildung 2.12: Cosmological constraints derived from a ten times larger sample in four cases: free  $B'_X$  and  $\gamma'_X$ , free  $B'_X$  and  $\gamma'_X = 0$ ,  $B'_X = 0$  and free  $\gamma'_X$ , and  $B'_X = \gamma'_X = 0$ . Noticeable shift in the inferred cosmological parameters occur once  $B'_X = 0$ , whereas the case of free  $B'_X$  and  $\gamma'_X = 0$  is very similar to the maximal case.

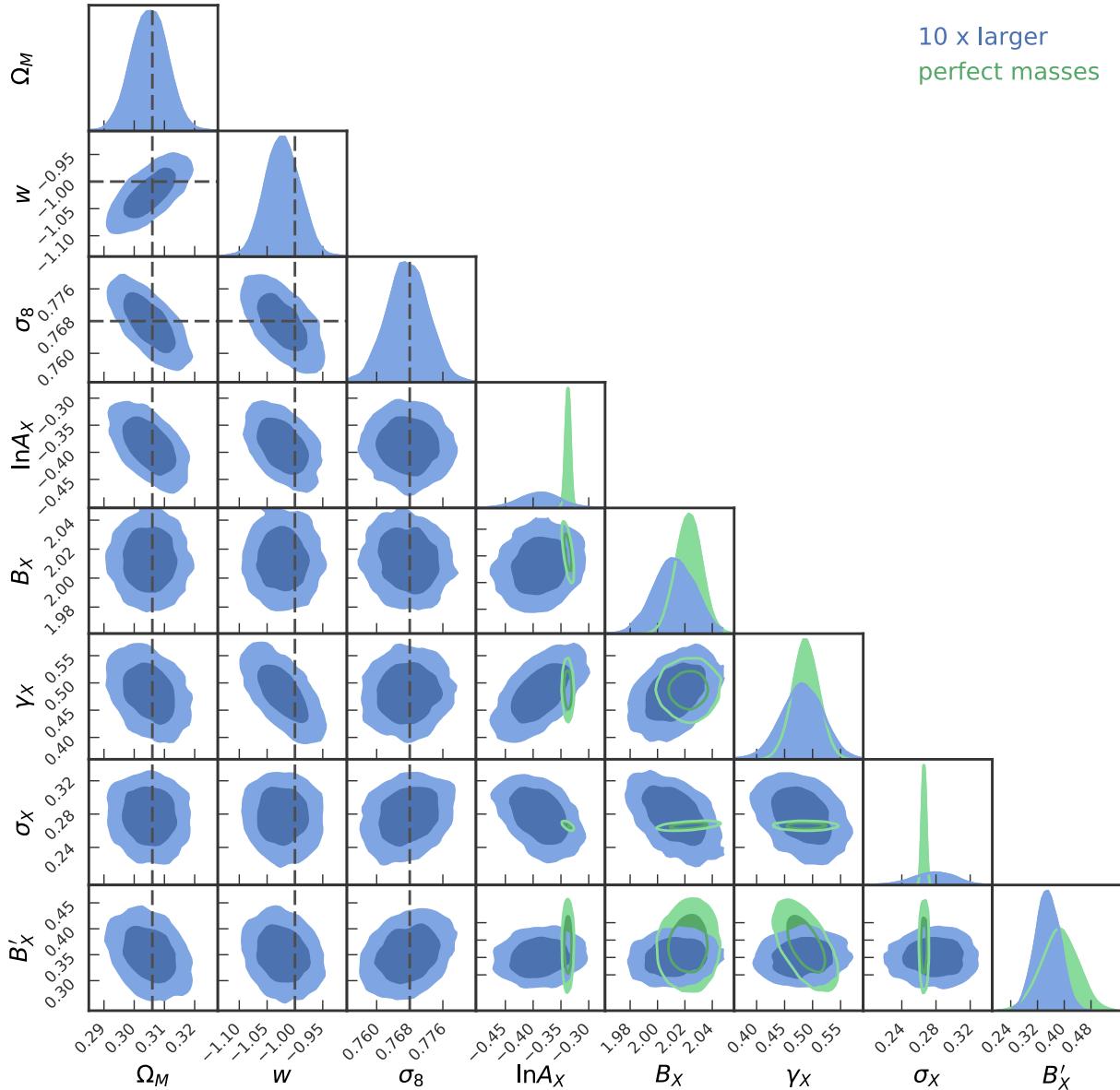


Abbildung 2.13: Posterior constraints of the number counts and the Euclid WL mass calibration of the 10 times larger validation mock (blue), and the constraints on the scaling relation parameters from the perfect masses mass calibration (green). All values are consistent within two sigma, indicating that our likelihoods are unbiased at a level which is small compared to the Poisson noise in our mock eROSITA sample.

### 2.6.3 Scaling relation form

It is worth noting that even though the observed X-ray luminosity to mass and temperature to mass relations show no evidence supporting a redshift dependent mass trend or a changing power law in the redshift trend (Bulbul et al., 2019), the mass and redshift dependences of the eROSITA count rate to flux conversion could introduce these additional trends into the count rate to mass scaling relation (equation 2.7). Here we explore whether these additional freedoms are needed in the forecasts we perform.

The ratio between the eROSITA count rate of a cluster and its rest frame 0.5-2 keV flux, which we shall call  $K = \eta/f_X$ , parameterizes the response of the eROSITA cameras to a specific spectral form and the k-correction necessary to account for the transformation from rest frame 0.5-2 keV to the observed frame. As such, one would expect it to be both redshift and temperature dependent. The rate can be written as

$$\eta = K f_X = K \frac{L_X}{4\pi d_L^2(z)}. \quad (2.26)$$

Thus, conceptually, the rate mass redshift scaling equals the luminosity mass redshift scaling, with the addition of the luminosity distance dependence and the mass and redshift trends of the flux to rate conversion.

$K$  has a noticeable but weak redshift dependence, and this redshift dependence is different for clusters of different temperature. The redshift dependence leads to the difference between the redshift trends of the rate and luminosity to mass relations (i.e., parameters  $\gamma_X$  and  $\gamma_L$ ), while the temperature dependence of this redshift trend combined with the temperature to mass redshift scaling leads to a non vanishing redshift variation in the mass trend of the rate scaling relation.

The baseline scaling relation for our analysis (equation 2.7) already has a redshift dependent mass trend parameter  $B'_X$  (equation 2.8). Here we introduce a generalization to equation (2.7) by allowing a redshift trend  $\gamma'_X$  in the redshift slope

$$\gamma(z) = \gamma_X + \gamma'_X \ln \left( \frac{1+z}{1+z_0} \right), \quad (2.27)$$

which allows for the power law index in redshift to change with redshift.

We examine the importance of this additional redshift dependence by fitting the scaling relation described by the full parameter set  $(\ln A_X, B_X, \gamma_X, \sigma_X, B'_X, \gamma'_X)$  using both true masses and Euclid WL mass constraints and allowing the parameter values to vary within large priors. Such an analysis is straightforward in the controlled regime of a mock catalog analysis where true masses and therefore the underlying form of the scaling known. In analyzing the real eROSITA dataset, one must use direct mass constraints like those from WL or dynamical masses to carry out a goodness of fit test for any proposed scaling relation form (as done in the SPT analyses; Bocquet et al., 2015; de Haan et al., 2016).

We analyze the number counts with perfect and with Euclid WL constraints for the 10 times larger validation mock (cf. Section 2.2.2) using effectively four different mass observable relations. As a baseline we sample both  $B'_X$  and  $\gamma'_X$ , the two parameters that

Tabelle 2.5: Results for the scaling relation parameters when using the perfect halo mass  $M_{500c}$  to calibrate the mass observable relation. These values are used as fiducial values for the scaling relation parameters in the rest of the work.

$\ln A_X$	$B_X$	$\gamma_X$	$\sigma_X$	$B'_X$
-0.328	1.997	0.446	0.278	0.355

allow for redshift variation of the mass and redshift trends. We then also consider the cases where either or both of the two extra parameters are set to zero. The resulting constraints on the cosmological parameters of interest are shown in Fig. 2.12. We find that the run with  $\gamma'_X = 0$  has constraints comparable to the baseline case where both  $B'_X$  and  $\gamma'_X$  are free. On the other hand, the run with  $B'_X = 0$  produces biased cosmological parameters. This finding is consistent with the conclusions of Pillepich, Porciani, and Reiprich (2012), who—although not using a rate mass scaling with empirical calibration, noted nonetheless that the effective rate mass scaling they derived had a mass slope which varied with redshift. We thus conclude that the parameter  $B'_X$  needs to be sampled, while the parameter  $\gamma'_X$  can be fixed to zero without biasing the cosmological inference at a level that is important, given the statistical uncertainties. In principle this parameter could also be left free to float without any bias implications, but adding more free parameters than needed in the mass observable relation tends to reduce the precision of the constraints on all parameters.

### Fiducial parameter values

To determine the fiducial values for the parameters of the scaling between rate and mass, we sample the mass calibration likelihood with perfect masses. We adopt the form justified in the previous section, where  $\gamma'_X = 0$  (see equation 2.7). The best fit scaling relation parameters when perfect masses are used are reported in Table 2.5. They are used as fiducial values in several occasions during this work. The value of  $\ln A_X$  can be rescaled arbitrarily by changing the pivot points  $\eta_0$ ,  $M_0$  and  $z_0$ .  $B_X$  and  $\sigma_X$  take values very similar to the mass slope of the luminosity mass relation  $B_L$  and the scatter around that relation  $\sigma_L$ , respectively. For the values presented in Table ??, we have sampled the mass calibration likelihood for perfect masses discussed in Eq. (2.10). The value for  $\gamma_X$  is larger than the redshift slope of the luminosity–mass scaling. Also, there is a clear preference for redshift evolution of the mass trend.

# Kapitel 3

## Application 2: Validation of Selection Function, Sample Contamination and Mass Calibration in Galaxy Cluster Samples

### 3.1 Conceptual Framework for Cluster Cosmology Analyses

In the following section we will present in mathematical details the model of the cluster population used in this work to describe the properties of the cluster samples. This discussion follows the Bayesian hierarchical framework established by Bocquet et al. (2015). The cluster population is modeled by a forward modeling approach which transforms the differential number of clusters as a function of halo mass  $M_{500c}$ <sup>1</sup> and redshift  $z$  to the space of observed cluster properties, such as the measured X-ray flux  $\hat{f}_X$ , the measured richness  $\hat{\lambda}$  and the measured SZE signal to noise  $\xi$ . This transformation is performed in two steps. First, utilising scaling relations with intrinsic scatter to compute the number in intrinsic flux, richness and SZE signal to noise. These relations have several free parameters such as amplitudes, mass and redshift trends, intrinsic scatter around the mean relation, and correlation coefficients among the intrinsic scatter on different observables. Constraining these free parameters is the aim of this work, as these constraints characterise the systematic uncertainty in the observable–mass relations. A second modeling step we apply models of the measurement uncertainty to construct the number density in observed cluster properties. We also present the modeling of the selection function and of the likelihood used to infer the parameters governing the scaling relations.

---

<sup>1</sup> $M_{500,c}$  is the mass enclosed in a spherical over density with average density 500 times the critical density of the Universe.

### 3.1.1 Modeling the cluster population

The starting point of our modeling of the cluster population is the differential number as function of halo mass  $M_{500c}$  and redshift  $z$ , given by

$$\frac{dN}{dM} \Big|_{M,z} = \frac{dn}{dM} \Big|_{M,z} \frac{d^2V}{dzd\Omega} dzd\Omega, \quad (3.1)$$

where  $\frac{dn}{dM} \Big|_{M,z}$  is the halo mass function describing the differential number density of halos at mass  $M$  and redshift  $z$ , as presented by Tinker et al. (2008)<sup>2</sup>;  $\frac{d^2V}{dzd\Omega} dzd\Omega$  is the cosmological volume subtended by the redshift bin  $dz$  and the survey angular footprint  $d\Omega$ .

The mapping from halo mass to intrinsic cluster properties is modeled by scaling relations, which are characterised by a mean relation with free parameters and a correlated scatter. The mean intrinsic relations we use read as

$$\langle f_X \rangle = \frac{L_0 A_X}{4\pi d_L^2(z)} \left( \frac{M h}{M_{0,X}} \right)^{B_X} \left( \frac{E(z)}{E(z_{0,X})} \right)^2 \left( \frac{1+z}{1+z_{0,X}} \right)^{C_X} \quad (3.2)$$

for the X-ray flux<sup>3</sup>,

$$\langle \lambda \rangle = A_\lambda \left( \frac{M h}{M_{0,\lambda}} \right)^{B_\lambda} \left( \frac{E(z)}{E(z_{0,\lambda})} \right)^{C_\lambda} \quad (3.3)$$

for the richness, and

$$\langle \zeta \rangle = A_{\text{SZ}} \left( \frac{M h}{M_{0,\text{SZ}}} \right)^{B_{\text{SZ}}} \left( \frac{E(z)}{E(z_{0,\text{SZ}})} \right)^{C_{\text{SZ}}} \quad (3.4)$$

for the SZE signal.  $h$  is numerical value the present day expansion rate if expressed in units of  $100 \text{m s}^{-1} \text{Mpc}^{-1}$ , and  $E(z)$  the ratio between the expansion rate at redshift  $z$  and the current day expansion rate. The form of the redshift evolution adopted in equations (3.3) and (3.4) has explicit cosmological dependence in the redshift evolution that is not well motivated (see discussion in Bulbul et al., 2019, hereafter Bu19), but we nevertheless adopt these forms for consistency with previous studies (e.g. S15). The pivot points in mass  $M_{0,X} = 6.35 \times 10^{14} h M_\odot h^{-1}$ ,  $M_{0,\lambda} = 3. \times 10^{14} M_\odot h^{-1} = M_{0,\text{SZ}}$ , in luminosity  $L_0 = 10^{44} \text{erg s}^{-1}$ , and in redshift  $z_{0,X} = 0.45$ ,  $z_{0,\lambda} = 0.6 = z_{0,\text{SZ}}$  are constants in our analysis. In contrast, the parameters  $A_N$ ,  $B_N$  and  $C_N$  for  $N \in (X, \lambda, \text{SZ})$  are free parameters of the likelihoods described below. These parameters encode the systematic uncertainty in the mass derived from each observable.

The inherent stochasticity in the cluster populations is modeled by assuming that the intrinsic observable scatters log-normally around the mean intrinsic relation<sup>4</sup>. Consequently, given mass  $M$  and redshift  $z$ , the probability for the intrinsic cluster observables ( $f_X$ ,

<sup>2</sup>We present here the entire formalism with distributions of  $M$  and of observables. In practice, we work with the natural logarithm of mass and of observables.

<sup>3</sup>The flux in this form makes explicit the cosmological dependencies due to distances and to self-similar evolution while allowing for departures from that self-similar evolution.

<sup>4</sup>The notation utilised here is imprecise. The scaling relation describes the mean of the natural logarithm of the intrinsic observable, not the natural logarithm of the mean, as suggested by the notation. Not interpreting  $\langle \cdot \rangle$  as an average ensures a fully consistent notation.

$\lambda, \zeta$ ) is given by

$$P(f_X, \lambda, \zeta | M, z) = \frac{1}{\sqrt{(2\pi)^3 \det \mathbf{C}}} \frac{1}{f_X \lambda \zeta} \exp \left\{ -\frac{1}{2} \mathbf{\Delta x}^T \mathbf{C}^{-1} \mathbf{\Delta x} \right\}, \quad (3.5)$$

with

$$\mathbf{\Delta x}^T = (\ln f_X - \ln \langle f_X \rangle, \ln \lambda - \ln \langle \lambda \rangle, \ln \zeta - \ln \langle \zeta \rangle) \quad (3.6)$$

and

$$\mathbf{C} = \begin{bmatrix} \sigma_X^2 & \sigma_X \sigma_\lambda \rho_{X,\lambda} & \sigma_X \sigma_{\text{SZ}} \rho_{X,\text{SZ}} \\ \sigma_X \sigma_\lambda \rho_{X,\lambda} & \sigma_\lambda^2 & \sigma_\lambda \sigma_{\text{SZ}} \rho_{\lambda,\text{SZ}} \\ \sigma_X \sigma_{\text{SZ}} \rho_{X,\text{SZ}} & \sigma_\lambda \sigma_{\text{SZ}} \rho_{\lambda,\text{SZ}} & \sigma_{\text{SZ}}^2 \end{bmatrix}, \quad (3.7)$$

where  $\sigma_{\aleph}$  for  $\aleph \in (X, \lambda, \text{SZ})$  encodes the magnitude of the intrinsic log-normal scatter in the respective observable, while the correlation coefficients  $\rho_{\aleph,\aleph'}$  encode the degree of correlation between the intrinsic scatters on the respective observables. The scatter parameters and the correlation coefficients are free parameters of our analysis. The probability density above is written as a probability density in the natural logarithm of the intrinsic observables. However, for sake of brevity, we omit the natural logarithm in the notation.

The differential number of objects as a function of intrinsic observables can be computed by applying the stochastic mapping from mass to intrinsic observables to the differential number of cluster as a function of mass, i.e

$$\frac{d^3 N}{d f_X d \lambda d \zeta} \Big|_{f_X, \lambda, \zeta, z} = \int dM P(f_X, \lambda, \zeta | M, z) \frac{dN}{dM} \Big|_{M, z}. \quad (3.8)$$

In some parts of our subsequent analysis, we do not require the distribution in SZE signal to noise. The differential number of objects as a function of intrinsic X-ray flux and richness can be obtained either by marginalising equation (3.8) over the intrinsic SZE signal  $\zeta$ , or by defining  $P(f_X, \lambda | M, z)$  just for the X-ray and optical observable by omitting the SZE part,

$$\begin{aligned} \frac{d^2 N}{d f_X d \lambda} \Big|_{f_X, \lambda, z} &= \int d\zeta \frac{d^3 N}{d f_X d \lambda d \zeta} \Big|_{f_X, \lambda, \zeta, z} \\ &= \int dM P(f_X, \lambda | M, z) \frac{dN}{dM} \Big|_{M, z}. \end{aligned} \quad (3.9)$$

### 3.1.2 Modeling measurement uncertainties

The intrinsic cluster observables are not directly accessible as only their measured values are known. We thus need to characterise the mapping between intrinsic and measured observables.

For the X-ray flux, we assume that the relative error on the flux  $\hat{\sigma}_X$  is the same as the relative error in the count rate. For each object ( $i$ ) in our catalog we can determine

$$P(\hat{f}_X^{(i)} | f_X) = \frac{1}{\sqrt{2\pi(\hat{\sigma}_X^{(i)})^2}} \frac{1}{\hat{f}_X^{(i)}} \exp \left\{ -\frac{1}{2} \frac{(\ln \hat{f}_X^{(i)} - \ln f_X)^2}{(\hat{\sigma}_X^{(i)})^2} \right\}. \quad (3.10)$$

For an application described below, it is necessary to know the measurement uncertainty on the X-ray flux for arbitrary  $\hat{f}_X$ , also those fluxes for which there is no corresponding entry in the catalog. As described in more detail in Appendix 3.7.1, we extrapolate the relative measurement uncertainty rescaled to the median exposure time in our footprint, creating a function  $\hat{\sigma}_X^2(\hat{f}_X, z, t_{\text{exp}})$ , which in turn allows us to compute

$$P(\hat{f}_X | f_X, z, t_{\text{exp}}) = \frac{1}{\sqrt{2\pi\hat{\sigma}_X^2(\hat{f}_X, z, t_{\text{exp}})}} \frac{1}{\hat{f}_X} \exp \left\{ -\frac{1}{2} \frac{(\ln \hat{f}_X - \ln f_X)^2}{\hat{\sigma}_X^2(\hat{f}_X, z, t_{\text{exp}})} \right\}. \quad (3.11)$$

Following S15, the measurement uncertainty on the optical richness is modeled as Poisson noise in the Gaussian limit, that is

$$P(\hat{\lambda} | \lambda) = \frac{1}{\sqrt{2\pi\lambda}} \exp \left\{ -\frac{1}{2} \frac{(\hat{\lambda} - \lambda)^2}{\lambda} \right\}. \quad (3.12)$$

The measurement uncertainty on the SZE signal to noise follows the prescription of Vanderlinde et al. (2010), who have determined the relation between measured SZE signal to noise  $\xi$  and the intrinsic signal to noise, as a function of the effective field depth  $\gamma_f$ <sup>5</sup>, namely

$$P(\xi | \zeta, \gamma_f) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \xi - \sqrt{\gamma_f^2 \zeta^2 + 3} \right)^2 \right\}. \quad (3.13)$$

### 3.1.3 Modeling selection functions

The selection functions in optical and SZE observables are easy to model as the mapping between measured and intrinsic observables is known and the selection criterion is a sharp cut in measured observable. For the optical case, the removal of random superpositions by imposing  $f_c < 0.05$  in the optical followup leads to a redshift dependent minimal measured richness  $\lambda_{\min}(z)$ , as discussed in K19. This leads to an optical selection function which is a step function in measured richness

$$P(\text{DES} | \hat{\lambda}, z) = \Theta(\hat{\lambda} - \lambda_{\min}(z)), \quad (3.14)$$

where  $\Theta(x)$  is the Heavyside step function with value 0 at  $x < 0$ , and 1 at  $x \geq 0$ . Using the measurement uncertainty on richness (equation 3.12), we construct the optical selection function in terms of intrinsic richness  $\lambda$  as

$$P(\text{DES} | \lambda, z) = P(\hat{\lambda} > \lambda_{\min}(z) | \lambda) = \int_{\lambda_{\min}(z)}^{\infty} d\hat{\lambda} P(\hat{\lambda} | \lambda) \quad (3.15)$$

The SPT-selected catalog we use in this work is selected by imposing that the measured signal to noise is  $\xi > 4.5$ , which, analogously to the optical case, is a step function in  $\xi$  and leads to an SZE selection function on  $\zeta$  given by

$$P(\text{SPT} | \zeta, \gamma_f) = P(\xi > 4.5 | \zeta, \gamma_f) = \int_{4.5}^{\infty} d\xi P(\xi | \zeta, \gamma_f). \quad (3.16)$$

---

<sup>5</sup>In de Haan et al. (2016) these factors are presented as renormalizations of the amplitude of the SZE-signal-mass relation. Our notation here is equivalent, but highlights that they describe a property of the mapping between intrinsic SZE-signal and measured signal, and not between intrinsic signal and mass.

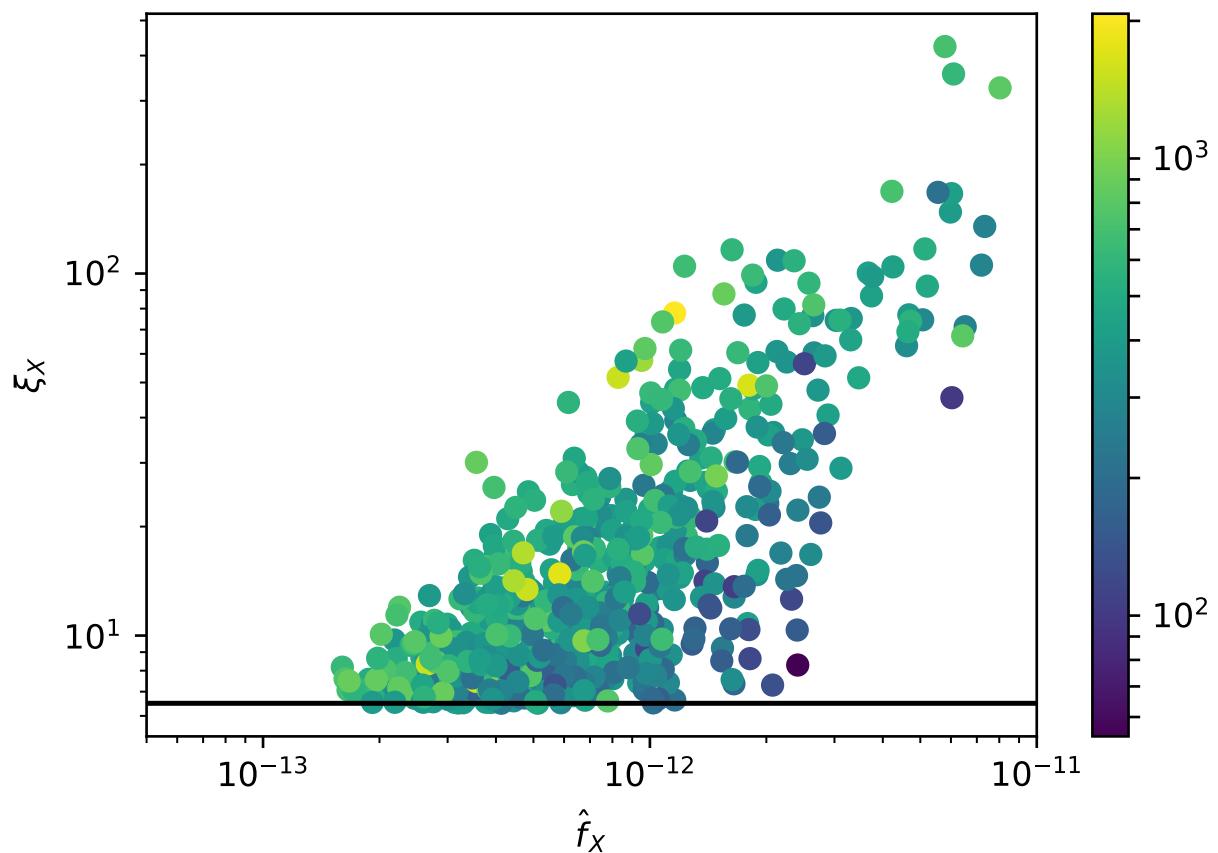


Abbildung 3.1: The measured X-ray flux  $\hat{f}_X$  and the X-ray detection probability  $\xi_X$ , color coded the exposure time. As a black line the X-ray selection  $\xi_X > 6.5$ . While X-ray flux and significance clearly display scaling, the scatter around this scaling correlates with exposure time.

### Constraining the X-ray selection function

The selection of the 2RXS catalog is given by the cut  $\xi_X > 6.5$ , where  $\xi_X$  is the significance of existence of a source, computed by maximizing the likelihood that a given source is not a background fluctuation (Boller et al., 2016). In the space of this observable, the selection function is a simple step function. However, in X-ray studies the selection function in a space of intrinsic X-ray flux is traditionally determined by image simulations (Vikhlinin et al., 1998; Pacaud et al., 2006; Clerc et al., 2018). In such an analysis, the emission from simulated clusters is used to create simulated X-ray images or event files, which are then analyzed with the same source extraction tools that are employed on the actual data. As a function of intrinsic flux, the fraction of recovered clusters is then used to estimate the selection function  $P(\text{X-det}|f_X, \dots)$ . This process captures, to the degree that the adopted X-ray surface brightness model is consistent with that of the observed population, the impact of morphological variation on the selection.

In this work, we take a novel approach, inspired by the treatment of optical and SZe selection functions outlined above. This approach is based on the concept that the traditional selection function can be described as a combination of two distinct statistical processes: the mapping between measured detection significance  $\xi_X$  and measured flux  $\hat{f}_X$ , and the mapping between measured flux  $\hat{f}_X$  and intrinsic flux  $f_X$ , i.e.

$$P(\text{X-det}|f_X, \dots) = \int d\hat{f}_X P(\text{X-det}|\hat{f}_X, \dots) P(\hat{f}_X|f_X, \dots), \quad (3.17)$$

The second part of the integrand is the description of the measurement uncertainty of the X-ray flux. This mapping is required to perform the number counts and any mass calibration. It should therefore be readily accessible. Its construction is described in Appendix 3.7.1. The first term can be easily computed from the mapping between measured flux  $\hat{f}_X$  and X-ray significance  $\xi_X$ ,  $P(\xi_X|\hat{f}_X, \dots)$ . Indeed, it is just the cumulative distribution of that mapping for  $\xi_X > 6.5$ .

The mapping between measured flux  $\hat{f}_X$  and X-ray significance  $\xi_X$  can be seen in Fig. 3.1 for the MARD-Y3 clusters, where we plot the detection significance against the measured fluxes. The relation displays significant scatter, which is partially due to the different exposure times (color-coded). Also clearly visible is the selection at  $\xi_X > 6.5$  (black line). As an empirical model for this relation we make the ansatz

$$\langle \xi_X \rangle = \xi_0(z) e^{\alpha_0} \left( \frac{\hat{f}_X}{f_0(z)} \right)^{\alpha_1} \left( \frac{t_{\text{exp}}}{400\text{s}} \right)^{\alpha_2}, \quad (3.18)$$

where  $\xi_0(z)$  and  $f_0(z)$  are the median significance and measured flux in redshift bins. To reduce measurement noise, we smooth them in redshift. We then assume that the significance of each cluster scatters around the mean significance with a log-normal scatter  $\sigma_\alpha$ . This provides the distribution  $P(\xi_X|\hat{f}_X, z, t_{\text{exp}})$ .

To fit the free parameters of this relation, namely  $(\alpha_0, \alpha_1, \alpha_2, \sigma_\alpha)$ , we determine the likelihood of each cluster  $i$  as

$$L_{\alpha,i} = \frac{P(\xi_X^{(i)}|\hat{f}_X^{(i)}, z^{(i)}, t_{\text{exp}}^{(i)})}{P(\xi_X > 6.5|\hat{f}_X^{(i)}, z^{(i)}, t_{\text{exp}}^{(i)})}, \quad (3.19)$$

where the numerator is given by evaluating  $P(\xi_X|\hat{f}_X, z, t_{\text{exp}})$  for each cluster, while the denominator ensures proper normalization for the actually observable data, i.e. clusters with  $\xi_X > 6.5$ . In properly normalising we account for the Malmquist bias introduced by the X-ray selection. Note also that we do not require the distribution of objects as a function of  $\hat{f}_X$  to perform this fit, as it would multiply both the numerator and the denominator and hence cancel out.

The total log-likelihood of the parameters  $(\alpha_0, \alpha_1, \alpha_2, \sigma_\alpha)$  is given by the sum of the log-likelihoods  $\ln L_\alpha = \sum_i \ln L_{\alpha,i}$ . This likelihood provides stringent constraints on the parameters  $(\alpha_0, \alpha_1, \alpha_2, \sigma_\alpha)$ . We find the best fitting values  $\alpha_0 = -0.113 \pm 0.020$ ,  $\alpha_1 = 1.275 \pm 0.031$ ,  $\alpha_2 = 0.799 \pm 0.038$  and  $\sigma_\alpha = 0.328 \pm 0.012$ . Noticeably, the constraints are very tight, indicating that the sample itself provides precise information about this relation.

Given this relation, the X-ray selection function can be computed as

$$\begin{aligned} P(\text{RASS}|\hat{f}_X, z, t_{\text{exp}}) &= P(\xi_X > 6.5|\hat{f}_X, t_{\text{exp}}, z) \\ &= \int_{6.5}^{\infty} d\xi_X P(\xi_X|\hat{f}_X, t_{\text{exp}}, z) \end{aligned} \quad (3.20)$$

Whenever the X-ray selection function is required, we sample the extra nuisance parameters with the ancillary likelihood (Eq. 3.19), marginalizing over the systematic uncertainties in this element of the X-ray selection function. Further discussion of the parameter posteriors and their use to test for systematics in the selection function can be found in section 3.5.1.

### Testing for additional dependencies

Empirically calibrating the relation governing the X-ray selection function has several benefits: (1) we take full account of the marginal uncertainty in the X-ray selection function. (2) Compared to image simulation, we do not rely on the realism of the clusters put into the simulation. Indeed, we use the data themselves to infer the relation. Together with the aforementioned marginalisation this ensures that we do not artificially bias our selection function. (3) We can empirically explore any further trends of the residuals of the significance–flux relation with respect to other quantities.

The latter is shown in Fig. 3.2, where the residual  $\sigma_\alpha^{-1} \ln(\xi_X^{(i)} / \langle \xi_X \rangle^{(i)})$  is plotted against redshift (upper left panel), Galactic hydrogen column density (upper right), background count rate in an aperture of 5' radius (lower left) and measured extent (lower right). As black dots we show the means if the populations in bins along the x axis. We find a weak trend with hydrogen column density. For simplicity we let this trend contribute to the overall scatter  $\sigma_\alpha$ . We find no correlation with the background brightness. There is a clear trend with measured extent, as can be expected for extended sources like clusters. However, we do not follow up on this trend, as 442 of the 708 cluster that we consider have a measured extent of 0 (due to the large PSF of RASS).

Most troublingly, we find a trend with redshift which is not captured by our model, as can be seen in the upper left panel of Fig. 3.2. At the lowest redshifts, we tend to over predict the significance given flux and exposure time, while at intermediate and high

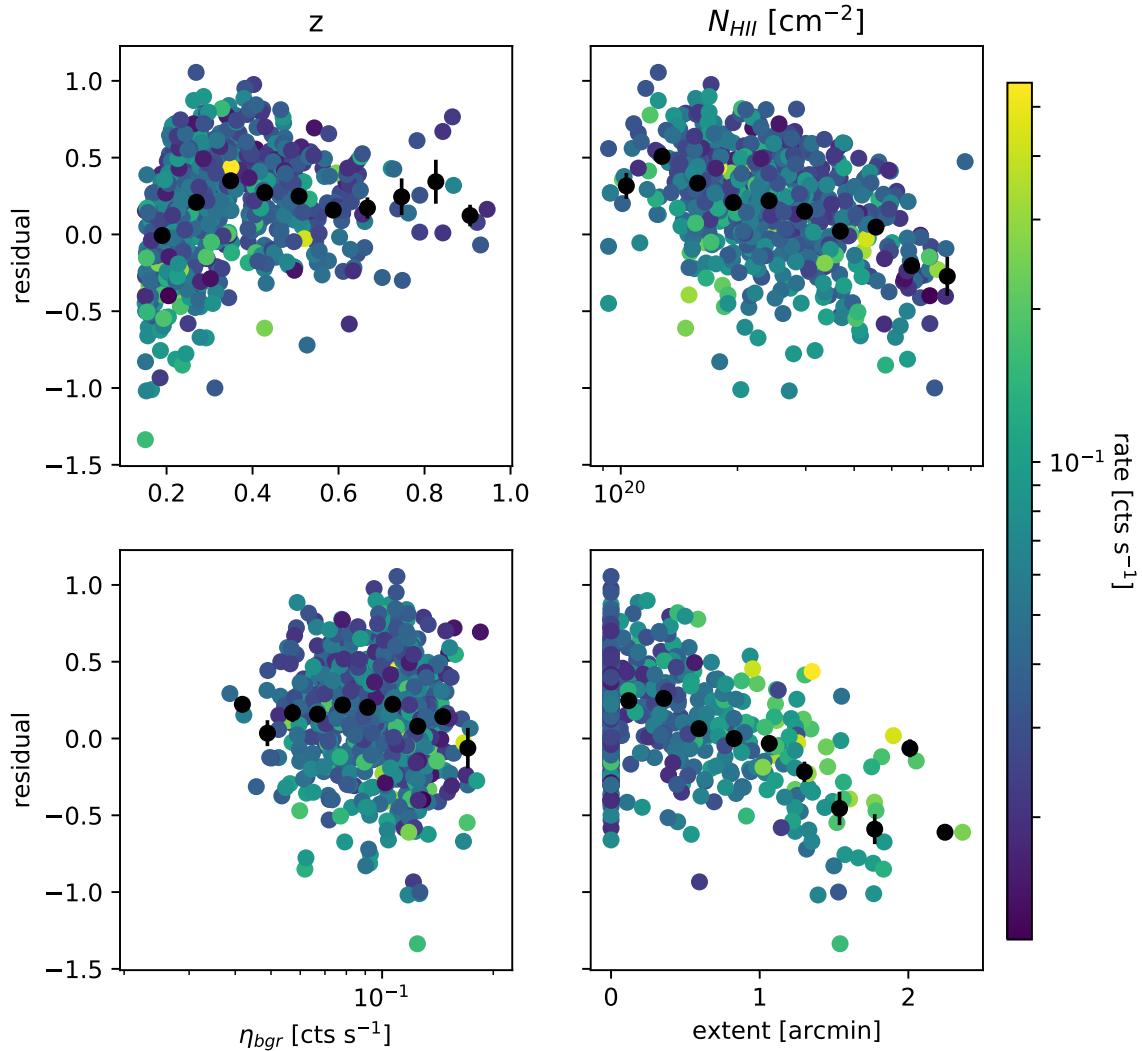


Abbildung 3.2: Residuals of the fitted significance–flux relation against redshift (upper left panel), neutral hydrogen column density (upper right), background counts rate in the aperture (lower left) and measured extent. This plot indicates that the addition of a redshift trend or and extent trend would be natural extension of our model. The current level of systematic and statistical uncertainties however does not require these extensions.

redshifts we tend to underestimate it. This residual systematic manifests itself at different stages in our analysis, and we discuss this as it arises and again in section 3.1.3.

## 3.2 Validation methods

As described above, the selection model for the clusters is specified by the form of the mass–observable relation and the intrinsic and observational scatter of the cluster population around the mean relation. The choice of the form of this relation should be driven primarily by what the data themselves demand, with guidance from the principle of preferred simplicity (Occam’s razor) and informed by predictions from structure formation simulations. This scaling relation should be empirically calibrated using methods such as weak lensing and dynamical masses, whose systematics can be calibrated and corrected using tests on structure formation simulations. Finally, a key step in cosmological analyses of cluster samples is to check consistency of the cluster sample with the best fit model of cosmology and mass–observable relation (e.g., goodness of fit; see Bocquet et al., 2015).

Given that the mass-observable relation can be calibrated using multiple sources of mass information, including direct mass information and from the cluster counts themselves, which is the distribution in observable and redshift of the cluster sample, there is ample opportunity for validation of the scaling relation and the selection function. In future cosmology analyses, blinding of the cosmological and nuisance parameters will be the norm, and cluster cosmology is no exception. The validation of a cluster sample through the requirement that all different reservoirs of information about the scaling relation lead to consistent results can be carried out in a blinded manner and should lead to improved stability and robustness in the final, unblinded cosmological results. We note that given the sensitivity of mass measurements to the distance-redshift relation and the sensitivity of the counts to both distance-redshift and growth of structure, these blinded tests should in general be carried out within each family of cosmological models considered (e.g., flat or curved  $\nu$ - $\Lambda$ CDM, flat or curved  $\nu$ -wCDM, etc).

In this work we seek to perform the following tests to validate the selection function modeling of the MARD-Y3 sample: (1) we investigate whether the X-ray flux–mass and richness–mass relation obtained by cross calibration using SPT-SZ mass information is consistent with the relation derived from the number counts of the MARD-Y3 sample; (2) we compare the scaling relation constraints from different flavours of number counts with each other (e.g., number counts in X-ray flux and redshift, in optical richness and redshift, and in both X-ray flux, optical richness and redshift); (3) finally, we constrain the probability of incompleteness in the SPT-SZ sample or contamination in the MARD-Y3 sample by comparing the clusters with and without counterparts in the other survey to the probabilities of having or not having counterparts as estimated using the selection functions. We take advantage in these validation tests of the fact that these scaling relations have been previously studied, and so we can compare our results not only internally but also externally to the literature. Finally, a key validation test could be carried out with the weak lensing information from DES, but we delay that to a future analysis where we hope

also to present unblinded cosmological results.

Given the stochastic description of the cluster population outlined above, we set-up different likelihood functions for each of these tests. These likelihoods are functions of the parameters determining the mapping between intrinsic observables and mass, the scatter around these relations and the correlation coefficients among the different components of scatter. Consequently, sampling these likelihoods with the data constrains the parameters. In the following sections we present the likelihoods used for each of the three validation methods listed above.

### 3.2.1 SPT-SZ cross calibration

For each object in the matched MARD-Y3 – SPT-SZ sample, we seek to predict the likelihood of observing the measured SZE signal to noise  $\xi^{(i)}$  given the measured X-ray flux  $\hat{f}_X^{(i)}$ , measured richness  $\hat{\lambda}^{(i)}$  and the scaling relation parameters. This likelihood is constructed by first making a prediction of the intrinsic SZE-signal to noise  $\zeta$  that is consistent with the measured X-ray flux  $\hat{f}_X^{(i)}$  and measured richness  $\hat{\lambda}^{(i)}$ , depending on the scaling relation parameters. To this end, the joint distribution of intrinsic properties is evaluated at the intrinsic fluxes and richesses consistent with the measurements

$$P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}) \propto \int d\lambda P(\hat{\lambda}^{(i)} | \lambda) \int df_X P(\hat{f}_X^{(i)} | f_X) \frac{d^3 N}{df_X d\lambda d\zeta} \Big|_{f_X, \lambda, \zeta, z^{(i)}}. \quad (3.21)$$

This expression of expected intrinsic SZE-signal takes account of the Eddington bias induced by the observational and intrinsic scatter in the X-ray and optical observable acting in combination with the fractionally larger number of objects at low mass, encoded in the last term of the expression.

To evaluate the likelihood of the measured SZE signal to noise  $\xi^{(i)}$  given the measured X-ray flux  $\hat{f}_X^{(i)}$  and measured richness  $\hat{\lambda}^{(i)}$ , we need to compare the predicted distribution  $P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)})$  with the likely values of intrinsic SZE-signal derived from the measurement  $\xi^{(i)}$  and the measurement uncertainty. This is written

$$P(\xi^{(i)} | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}) = \frac{\int d\zeta P(\xi^{(i)} | \zeta, \gamma_f^{(i)}) P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)})}{\int d\zeta P(\text{SPT} | \zeta, \gamma_f^{(i)}) P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)})}. \quad (3.22)$$

Notably, the denominator ensures the proper normalisation and also takes into account the Malmquist bias introduced by the SPT-SZ selection. Also note that the normalization cancels the dependence of this likelihood on the amplitude of the number of objects at the redshift  $z^{(i)}$ , measured flux  $\hat{f}_X^{(i)}$  and measured richness  $\hat{\lambda}^{(i)}$ . This strongly weakens its cosmological dependence and makes it independent of the X-ray and the optical selection function (see also Liu et al., 2015). For sake of brevity we omitted that this likelihood

depends on the scaling relation parameters and the cosmological parameters, all needed to compute the distribution of intrinsic properties.

The total log-likelihood of SPT-SZ cross calibration over the matched sample is given by the sum of the individual log likelihoods

$$\ln L_{\text{SPTcc}} = \sum_{i \in \text{matched}} \ln P(\xi^{(i)} | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}), \quad (3.23)$$

which is a function of the scaling relation parameters and cosmology. Sampling it with priors on the SZE scaling relation parameters that come from an external calibration will then transfer that calibration to the X-ray flux and richness scaling relations.

### 3.2.2 Calibration with number counts

The number of clusters as a function of measured observable and redshift is a powerful way to constrain the mapping between observable and mass, because the number of clusters as a function of mass is known for a given cosmology (see self-calibration discussions in Majumdar and Mohr, 2003; Hu, 2003; Majumdar and Mohr, 2004).

#### X-ray number counts

The likelihood of number counts is given by

$$\ln L_{\text{nc X}} = \sum_i \ln N \Big|_{\hat{f}_X^{(i)}, z^{(i)}} - N_{\text{tot}}, \quad (3.24)$$

where the expected number of objects as a function of measured flux  $\hat{f}_X^{(i)}$  and redshift  $z^{(i)}$  is

$$\begin{aligned} N \Big|_{\hat{f}_X^{(i)}, z^{(i)}} &= P(\text{RASS} | \hat{f}_X^{(i)}, z^{(i)}, t_{\text{exp}}) \int df_X P(\hat{f}_X^{(i)} | f_X) \\ &\quad \int d\lambda P(\text{DES} | \lambda, z^{(i)}) \frac{d^2 N}{df_X d\lambda} \Big|_{f_X, \lambda, z^{(i)}} df_X, \end{aligned} \quad (3.25)$$

where the first factor takes into account the X-ray selection, the second factor models the measurement uncertainty on the X-ray flux and the third factor models the optical incompleteness.

The total number of objects is computed as

$$\begin{aligned} N_{\text{tot}} &= \int dt_{\text{exp}} P(t_{\text{exp}}) \int dz \int d\hat{f}_X P(\text{RASS} | \hat{f}_X, z, t_{\text{exp}}) \\ &\quad \int df_X P(\hat{f}_X | f_X, z, t_{\text{exp}}) \int d\lambda P(\text{DES} | \lambda, z) \\ &\quad \frac{d^2 N}{df_X d\lambda} \Big|_{f_X, \lambda, z}, \end{aligned} \quad (3.26)$$

where  $P(t_{\text{exp}})$  is the solid angle weighted exposure time distribution. We highlight here that differently from previous work, we explicitly model not only the selection on the X-ray observable, but also fold in the incompleteness correction due to the MCMF optical cleaning via the term  $P(\text{DES} | \lambda, z)$ .

### Optical number counts

While not customary for a predominantly X-ray selected sample, the number counts of clusters can also be performed in measured richness. In this case, the likelihood reads

$$\ln L_{\text{nc } \lambda} = \sum_i \ln N \Big|_{\hat{\lambda}^{(i)}, z^{(i)}} - N_{\text{tot}}, \quad (3.27)$$

where  $N_{\text{tot}}$  is given by equation (3.26), whereas the expected number of clusters as a function of measured richness  $\hat{\lambda}^{(i)}$  and redshift  $z^{(i)}$  is computed as follows

$$N \Big|_{\hat{\lambda}^{(i)}, z^{(i)}} = \int dt_{\text{exp}} P(t_{\text{exp}}) \int d\hat{f}_X P(\text{RASS} | \hat{f}_X, z, t_{\text{exp}}) \\ \int df_X P(\hat{f}_X | f_X, z, t_{\text{exp}}) \int d\lambda P(\hat{\lambda}^{(i)} | \lambda, z) \frac{d^2 N}{df_X d\lambda} \Big|_{f_X, \lambda, z} d\hat{\lambda}, \quad (3.28)$$

where the first three integrals take account of the X-ray selection, whereas the last integral models the measurement uncertainty on the richness.

### Combined X-ray and optical number counts

Besides performing the number counts in only one observable, one can also perform the number counts in more than one observable (e.g. Mantz et al., 2010), in our case by fitting for the number of objects as a function of both measured flux  $\hat{f}_X^{(i)}$  and richness  $\hat{\lambda}^{(i)}$ . In that case the likelihood reads

$$\ln L_{\text{nc } X, \lambda} = \sum_i \ln N \Big|_{\hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}} - N_{\text{tot}}, \quad (3.29)$$

where the expected number of objects as a function of measured flux  $\hat{f}_X^{(i)}$  and richness  $\hat{\lambda}^{(i)}$  is computed as

$$N \Big|_{\hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}} = P(\text{RASS} | \hat{f}_X^{(i)}, z^{(i)}, t_{\text{exp}}^{(i)}) \int df_X P(\hat{f}_X^{(i)} | f_X) \\ \int d\lambda P(\hat{\lambda}^{(i)} | \lambda) \frac{d^2 N}{df_X d\lambda} \Big|_{f_X, \lambda, z^{(i)}} d\hat{f}_X d\hat{\lambda}, \quad (3.30)$$

computed by folding the intrinsic number density with the measurement uncertainties on flux and richness.

#### 3.2.3 Consistency check using two cluster samples

Given the selection functions for two cluster samples, one can calculate the probability that any member of one sample is present in the other. This provides a powerful consistency check of the two selection functions, and if anomalies are found this approach can be used, for example, to probe for contamination or unexplained incompleteness in the cluster samples.

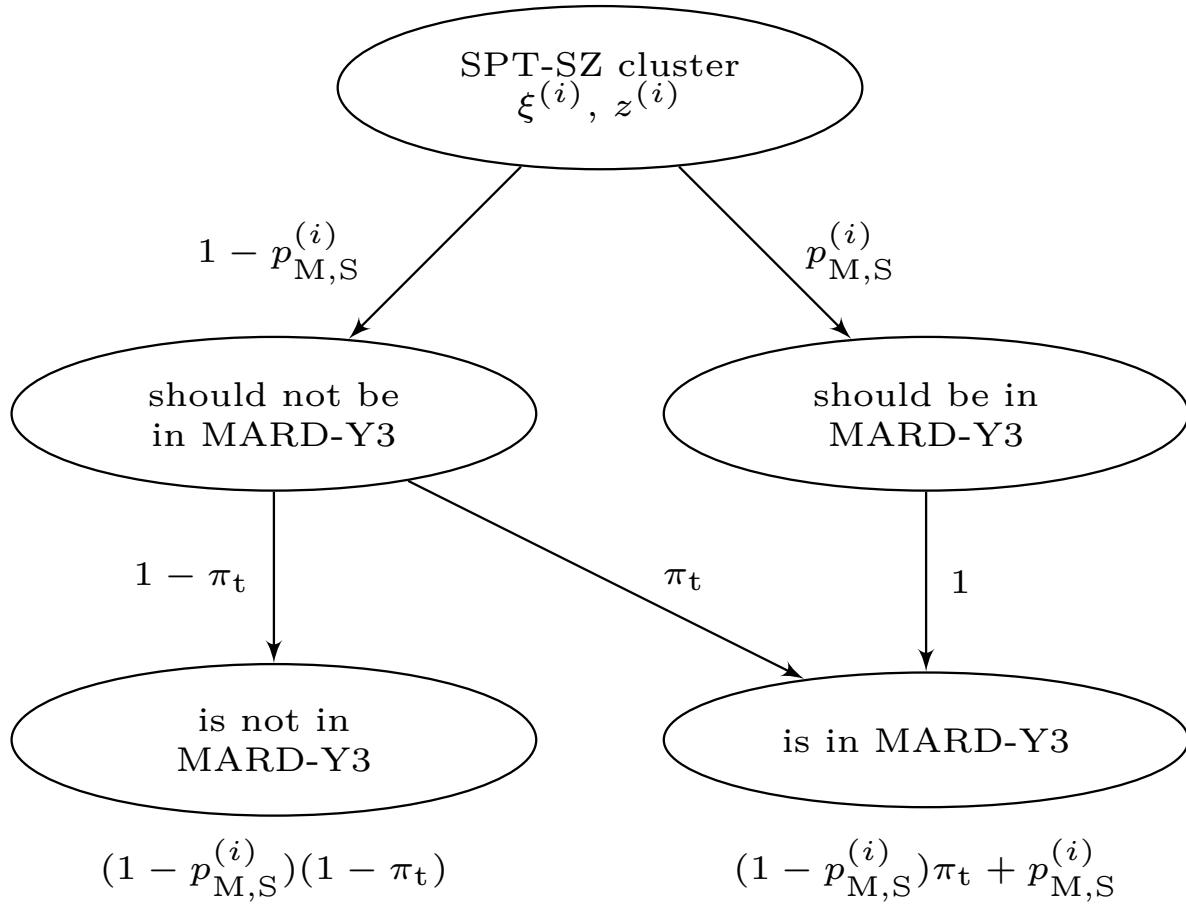


Abbildung 3.3: Probability tree describing the probability of an SPT-SZ cluster being detected in MARD-Y3. Besides the matching probability computed from the scaling between intrinsic observables and mass, the scatter around this relation, the observational uncertainties on the observables and the selection functions  $p_{M,S}^{(i)}$ , we also introduce the chance of either X-ray flux boosting or SZE signal dimming  $\pi_t$ , which would lead to the MARD-Y3 detection of SPT-SZ cluster that should otherwise not have been matched. Summarized at the end of each branch are the probabilities of matching or of not matching.

### MARD-Y3 detection probability for SPT-SZ clusters

For any SPT-SZ cluster with measured SZE signal to noise  $\xi^{(i)}$  and redshift  $z^{(i)}$  in the joint SPT-DES-Y3 footprint we can compute the probability of being detected by MARD-Y3 as follows. We first predict the probability distribution of intrinsic fluxes and richnesses associated with the measured SZE-signal to noise as

$$P(f_X, \lambda | \xi^{(i)}, z^{(i)}) \propto \int d\zeta P(\xi^{(i)} | \zeta, \gamma_f^{(i)}) \frac{d^3 N}{df_X d\lambda d\zeta} \Big|_{f_X, \lambda, \zeta, z^{(i)}}. \quad (3.31)$$

This expression needs to be properly normalised to be a distribution in intrinsic flux and richness. This is achieved by imposing  $\int df_X \int d\lambda P(f_X, \lambda | \xi^{(i)}, z^{(i)}) = 1$ , which sets the proportionality constant of the equation above. Note that this normalization cancels the dependence of this expression on the number of clusters observed. The predicted distribution of intrinsic fluxes and richnesses needs to be folded with the selection functions to compute the detection probability. The optical selection function is simply given by equation (3.15) evaluated at the cluster redshift  $z^{(i)}$ . On the other hand, when computing the X-ray selection function we take the RASS exposure time at the SPT-SZ position into account, while marginalising over all possible measured fluxes. The X-ray selection function thus reads

$$P(\text{RASS} | f_X, t_{\text{exp}}^{(i)}, z^{(i)}) = \int d\hat{f}_X P(\text{RASS} | \hat{f}_X, t_{\text{exp}}^{(i)}, z^{(i)}) P(\hat{f}_X | f_X, t_{\text{exp}}^{(i)}, z^{(i)}), \quad (3.32)$$

where the second factor is taken from equation (3.11), the expression for the X-ray measurement error at arbitrary measured flux  $\hat{f}_X$ .

The probability of detecting in MARD-Y3 a SPT-SZ cluster with measured SZE-signal to noise  $\xi^{(i)}$  and redshift  $z^{(i)}$  can then be computed by folding the predicted distribution of fluxes and richnesses with the selection functions in flux and richness as follows

$$p_{M,S}^{(i)} := P(\text{RASS, DES} | \xi^{(i)}, z^{(i)}) = \int df_X P(\text{RASS} | f_X, t_{\text{exp}}^{(i)}, z^{(i)}) \int d\lambda P(\text{DES} | \lambda, z^{(i)}) P(f_X, \lambda | \xi^{(i)}, z^{(i)}), \quad (3.33)$$

where we omit the dependence on the SPT-SZ field depth  $\gamma_f^{(i)}$  at the position of the SPT-SZ selected cluster.

Given these probabilities, we can define two interesting classes of objects: (1) SPT-SZ clusters that should not have a MARD-Y3 match given their low probability but have been nonetheless matched, and (2) SPT-SZ clusters with a very high chance of being matched that have nonetheless not been matched. For the discussion in this paper we adopt a low probability threshold of  $p_{M,S}^{(i)} < 0.025$  for the first class and call them SPT-SZ false positives, and we adopt a high probability threshold of  $p_{M,S}^{(i)} > 0.975$  for the second class and call them SPT-SZ true negatives.

Anticipating that we find several SPT-SZ false positives and no SPT-SZ true negative, we introduce here the probability  $\pi_t$  that an SPT-SZ cluster has a SZE signature lower than expected by the scaling relation (which includes the effects of intrinsic and observed

scatter), or equivalently an X-ray flux that is larger than expected, effectively leading to an underestimation of  $p_{M,S}^{(i)}$  or a matched detection despite low  $p_{M,S}^{(i)}$ . The likelihood of  $\pi_t$  can be computed by following the probability tree shown in Fig 3.3. The probability of being matched is  $(1 - p_{M,S}^{(i)})\pi_t + p_{M,S}^{(i)}$ , while the probability of not being matched is  $(1 - p_{M,S}^{(i)})(1 - \pi_t)$ . Thus, the log-likelihood is given by

$$\begin{aligned} \ln L(\pi_t) = & \sum_{i \in \text{match}} \ln \left( (1 - p_{M,S}^{(i)})\pi_t + p_{M,S}^{(i)} \right) + \\ & + \sum_{i \in \text{!match}} \ln \left( (1 - p_{M,S}^{(i)})(1 - \pi_t) \right) \end{aligned} \quad (3.34)$$

This likelihood also depends on the scaling relation parameters through the detection probabilities  $p_{M,S}^{(i)}$ . Marginalizing over these scaling relation parameters accounts for the systematic uncertainty on the observable–mass relations.

This approach allows us to go beyond a consistency check and use the cross-matched cluster samples to quantify the inadequacy of the adopted and calibrated selection functions to describe the samples.

### SPT-SZ detection probability for MARD-Y3 clusters

Similarly to the case in the previous section, for each MARD-Y3 cluster with measured X-ray flux  $\hat{f}_X^{(i)}$ , measured richness  $\hat{\lambda}^{(i)}$  and redshift  $z^{(i)}$  in the joint SPT-SZ – DES Y3 footprint, we can compute the probability of it being detected by SPT

$$\begin{aligned} p_{S,M}^{(i)} &:= P(\text{SPT} | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}) \\ &= \int d\zeta P(\text{SPT} | \zeta, \gamma_f^{(i)}) P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)}), \end{aligned} \quad (3.35)$$

where the first factor in the integrand is the SPT-SZ selection function evaluated for the field depth at the MARD-Y3 cluster position, while the second factor is the prediction for the intrinsic SZE-signal to noise consistent with the measured X-ray and optical properties. The latter is taken from equation (3.21) while ensuring that it is properly normalized,  $\int d\zeta P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}) = 1$ .

MARD-Y3 false positives and MARD-Y3 true negatives can be defined as for the SPT-SZ cases discussed in Section 3.2.3. In this case, we introduce the probability of each individual MARD-Y3 cluster being a contaminant  $\pi_c$ , and the probability that a SPT-SZ cluster that should be detected has not been detected, leading to SPT-SZ incompleteness,  $\pi_i$ . From the probability tree shown in Fig. 3.4 we can determine the probability of a MARD-Y3 cluster being matched by SPT-SZ as  $(1 - \pi_c)p_{S,M}^{(i)}(1 - \pi_i)$ , and the probability of a cluster not being matched as  $\pi_c + (1 - \pi_c)(1 - p_{S,M}^{(i)} + \pi_i p_{S,M}^{(i)})$ . Thus, the log-likelihood is given by

$$\begin{aligned} \ln L(\pi_c, \pi_i) = & \sum_{i \in \text{match}} \ln \left( (1 - \pi_c)p_{S,M}^{(i)}(1 - \pi_i) \right) + \\ & + \sum_{i \in \text{!match}} \ln \left( \pi_c + (1 - \pi_c)(1 - p_{S,M}^{(i)} + \pi_i p_{S,M}^{(i)}) \right). \end{aligned} \quad (3.36)$$

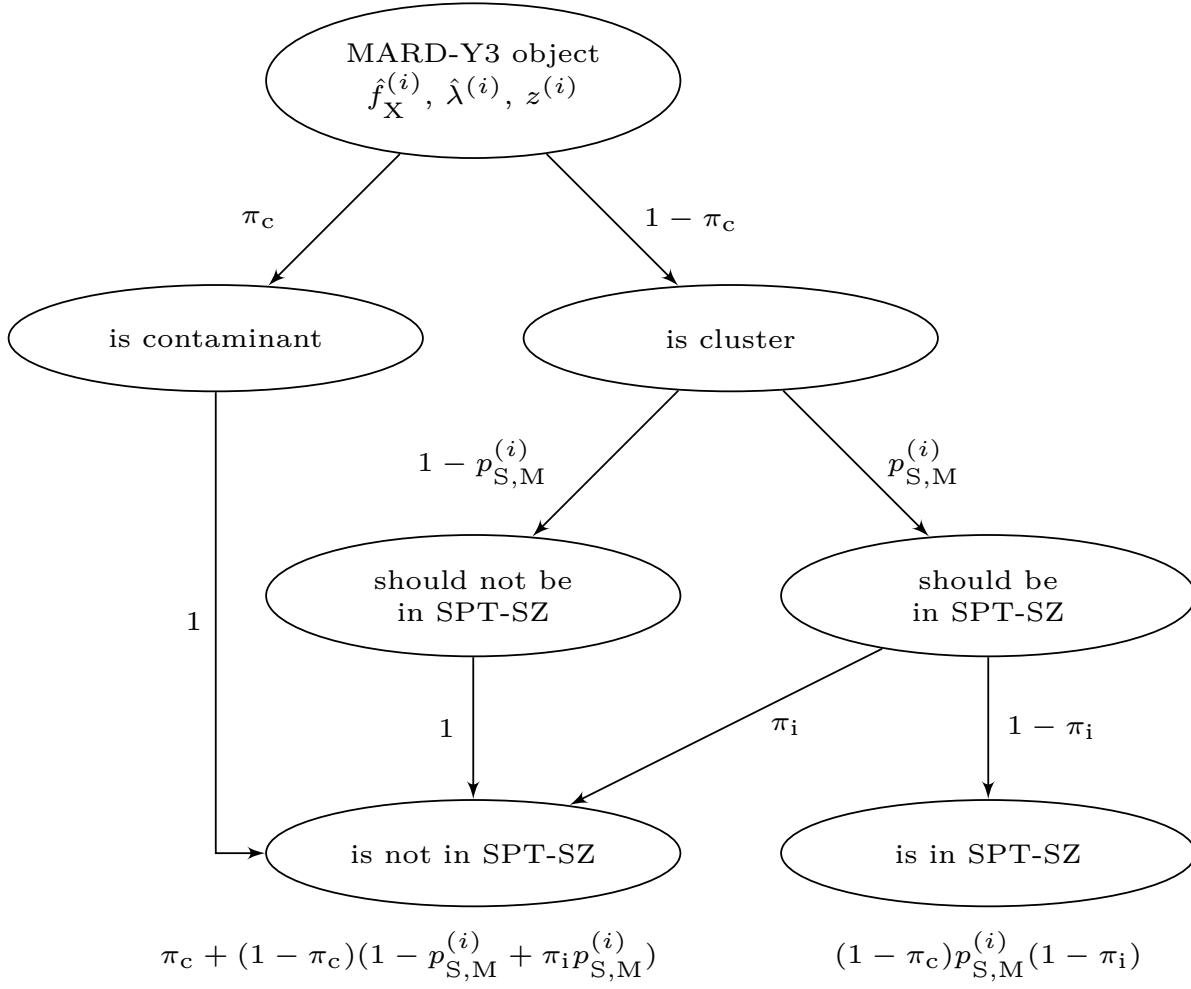


Abbildung 3.4: Probability tree describing the probability of a MARD-Y3 cluster being detected by SPT. Besides the matching probability computed from the scaling between intrinsic observables and mass, the scatter around this relation, the observational uncertainties on the observables and the selection functions  $p_{SM}^{(i)}$ , we also introduce the chance that a MARD-Y3 cluster is a contaminant  $\pi_c$ , and the chance that SPT-SZ misses a cluster it should detect, indicating incompleteness in SPT,  $\pi_i$ . Summarized at the end of each branch are the probabilities of being matched or not matched.

This likelihood also depends on the scaling relation parameter through the detection probabilities  $p_{M,S}^{(i)}$ . Marginalizing over the scaling relation parameters accounts for the systematic uncertainties on the observable–mass relations.

## 3.3 Dataset and Priors

We present here the cluster samples and then the priors used in obtaining the results presented in the following section.

### 3.3.1 Cluster samples

Here we summarize not only the main properties of the prime focus of our validation, the MARD-Y3 cluster sample, but also the SPT-SZ sample that we use for validation and for cross-matching with MARD-Y3.

#### MARD-Y3 X-ray selected clusters

In this work we seek to validate the mass information and the selection function modeling of the MARD-Y3 cluster sample, presented in K19. In that work, optical follow-up with the MCMF algorithm (Klein et al., 2018) of the RASS 2nd faint source catalog (hereafter 2RXS, Boller et al., 2016) is performed by scanning the DES photometric data with a spatial filter based on X-ray candidate position and inferred mass, and a color filter based on the red-sequence model at a putative redshift. This process provides a cluster richness estimate  $\hat{\lambda}$  and photometric redshift  $z$ . Comparison to the richness distribution in lines of sight without X-ray candidates allows one to estimate the probability  $f_c$  that the X-ray source and optical system identified by MCMF are a random superposition (contamination). In cases of multiple richness peaks along a line of sight toward a 2RXS candidate, the redshift with lowest  $f_c$  is identified as the optical counterpart. The redshifts display sub-percent level scatter w.r.t. spectroscopic redshifts, and the richnesses  $\hat{\lambda}$  can be adopted as an additional cluster mass proxy.

In K19, careful comparison of the luminosities from 2RXS to luminosities reported by Piffaretti et al. (2011) for the MCXC catalog allows one to transform the 2RXS luminosity measured in a fixed aperture with radius of  $5'$  into a luminosity within  $R_{500c}$ , the radius enclosing an over-density of 500 w.r.t. to the critical density. This correction is only reliable at  $z > 0.15$ .

In this work we focus on the  $z > 0.15$  sample with  $f_c < 0.05$  with an additional rejection of luminosity–richness outliers with an infra-red signature compatible with an active galactic nucleus (c.f. K19, Section 3.11). Our MARD-Y3 sample is then 708 clusters in a footprint of  $5204 \text{ deg}^2$  with an expected contamination of 2.6% (K19).

In Figure 3.5 we show the redshift–X-ray inferred mass distribution of this sample, color coded to reflect the cluster richnesses. We also show as a black line the mass corresponding to 40 photon counts in the first eROSITA full sky survey (eRASS1), computed using the

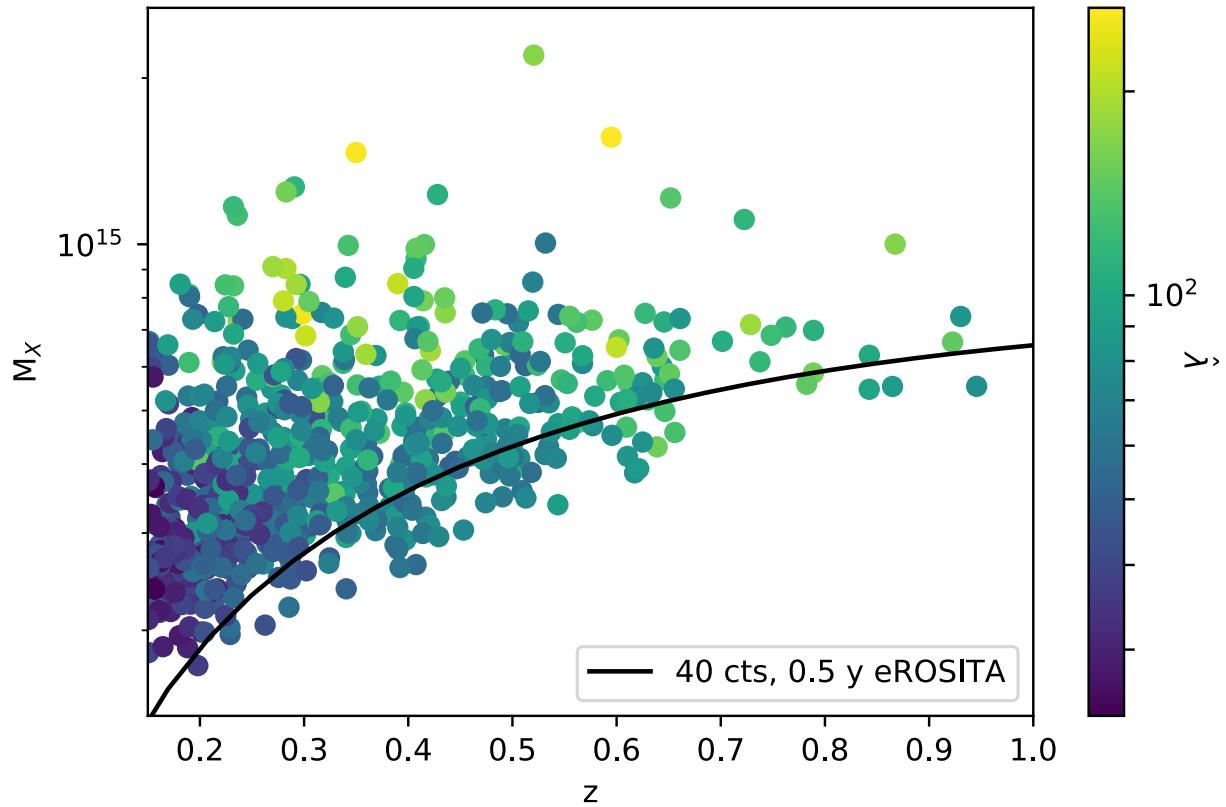


Abbildung 3.5: MARD-Y3 sample of 708 cluster constructed by cleaning the 2nd RASS faint source catalog with DES data. While not used in the rest of the analysis, the X-ray inferred mass  $M_X$  is used here to highlight the mass range of our sample. The color encodes the measured richness of the counterpart in the DES data. The black line indicates the forecast of the mass corresponding to 40 photon counts in the first eROSITA full sky survey after half a year of observing time.

eROSITA count rate–mass relation forecast by Grandis et al. (2019). This indicates that the MARD-Y3 sample we study here is comparable to the one we expect to study in the eRASS1 survey.

Finally, several other X-ray properties, such as the detection significance  $\xi_X$  and the RASS exposure time  $t_{\text{exp}}$  are available from 2RXS. The X-ray flux  $\hat{f}_X$  we employ in the following is computed as  $\hat{f}_X = L_X/(4\pi d_L^2(z))$ , where  $L_X$  is the X-ray luminosity within  $R_{500c}$ , and  $d_L(z)$  is the luminosity distance evaluated at the reference cosmology. This leads to the fact that technically our X-ray flux corresponds to the rest frame [0.5,2] keV. The transformation from the observed [0.1,2.4] keV band to this band is discussed in K19. It is also noteworthy that MCMF allows one to detect the presence of more than one significant optical structure along the line of sight towards an X-ray candidate.

### SPT-SZ SZE selected clusters

We adopt the catalog of clusters selected via their SZE signatures in the SPT-SZ  $2500 \text{ deg}^2$  survey Bleem et al. (2015). Utilising this sample to an SZE signal to noise of 4.5, we confirm the clusters in the DES-Y3 footprint using MCMF (Klein et al., in prep.). The low contamination level of the parent sample allows one to achieve a low level of contamination by imposing the weak cut of  $f_c < 0.2$ . Above a redshift of  $z < 0.2$  this provides us with a sample of 436 clusters. The X-ray properties, as well as the optical properties of these objects have been extensively studied (see for instance McDonald et al., 2014; Saro et al., 2015; Hennig et al., 2017; Chiu et al., 2018; Bulbul et al., 2019; Capasso et al., 2019b, and references therein). Furthermore, successful cosmological studies have been performed with this sample (Bocquet et al., 2015; de Haan et al., 2016; Bocquet et al., 2019b), indicating that the survey selection function is well understood and that the mass information derived from the SZE is reliable. This motivates us to employ this sample as a reference for our validation of the observable mass relations and the selection function of the MARD-Y3 sample.

### Cross-matched sample

To identify clusters selected both by SPT-SZ and by MARD-Y3, we perform a positional matching within the angular scale of 2 Mpc at the MARD-Y3 cluster redshift. We match 120 clusters in the redshift range  $z \in (0.2, 1.1)$ . We identify 3 clusters where the redshift determined by the MCMF run on RASS,  $z_{\text{RASS}}$ , is significantly different from the redshift MCMF assigns for the SPT-SZ candidate,  $z_{\text{SPT}}$ . While for all three cases  $z_{\text{RASS}} < z_{\text{SPT}}$ , in all cases the MCMF run on the SPT-SZ candidate list identifies optical structures at  $z_{\text{RASS}}$  as well. While we consider these clusters as matched, we do not employ them in SPT-SZ cross calibration of the X-ray flux and richness mass information. Both their X-ray fluxes and SZE signals are likely biased w.r.t. to the nominal relation for individual clusters due to the presence of several structures along the line of sight. Disentangling the respective contributions of the different structures along the line of sight is complicated by different scaling of X-ray flux and SZE signal with distance.

Tabelle 3.1: Summary of the priors employed in this work. These priors are implemented as Gaussians. In Section 3.4 we will specify which priors are used for which analysis.

Cosmological Parameters		
$H_0$	$70.6 \pm 2.6$	Rigault et al. (2018)
$\Omega_M$	$0.276 \pm 0.047$	SPT (Bo19)
$\sigma_8$	$0.781 \pm 0.037$	SPT (Bo19)
SZE $\zeta$ -mass Relation		
$A_{\text{SZ}}$	$5.24 \pm 0.85$	SPT (Bo19)
$B_{\text{SZ}}$	$1.53 \pm 0.10$	
$C_{\text{SZ}}$	$0.47 \pm 0.41$	
$\sigma_{\text{SZ}}$	$0.16 \pm 0.08$	
X-ray $L_X$ -mass Relation		
$A_X$	$4.15 \pm 1.01$	SPT (Bu19)
$B_X$	$1.91 \pm 0.17$	
$C_X$	$0.20 \pm 0.42$	
$\sigma_X$	$0.25 \pm 0.10$	
Optical $\lambda$ -mass Relation		
$A_\lambda$	$71.9 \pm 6.1$	SPT-DES (S15)
$B_\lambda$	$1.14 \pm 0.20$	
$C_\lambda$	$0.73 \pm 0.76$	
$\sigma_\lambda$	$0.15 \pm 0.08$	

In only one case, two MARD-Y3 clusters are associated with the same SPT-SZ cluster: ‘SPT-CL J2358-6129’,  $z_{\text{SPT}} = 0.403$ . Visual inspection (c.f. Fig 3.7.2) reveals that one of the MARD-Y3 clusters,  $z_{\text{RASS}} = 0.398$ , is well centered on the SZE signal, and also coincides with a peak in the galaxy density distribution. The second MARD-Y3 cluster in the north-northwest,  $z_{\text{RASS}} = 0.405$ , is offset from the peaks in galaxy density, and does not correspond to any SZE signal. Given the lack of the SZ-counterpart, we do consider this MARD-Y3 cluster not being matched by SPT. Our final matched sample therefore contains 123 clusters. Including the aforementioned clusters in our analysis does not change the posterior results in a noticeable manner.

### 3.3.2 Priors

Our priors reflect the assumptions we make for each analysis. On one side we calibrate the X-ray flux-mass and richness-mass relation assuming SPT-SZ mass information. On the other hand, we investigate which X-ray flux-mass and richness-mass relation reproduce the observed abundance of objects when assuming an SPT-SZ derived cosmology. Contingent upon the validity of both the SPT-SZ cosmology and the mass information, when can therefore establish whether the mass information contained in our X-ray fluxes and richnesses is consistent with the abundances of clusters as a function of these observables. We implement the SPT-SZ masses and cosmological parameters as priors described below

and summarized in Table 3.1.

### Priors on cosmology

Throughout this work, we marginalise over the following cosmological parameters to propagate our uncertainty on these parameters. We adopt the prior on the present day expansion rate, also called the Hubble constant,  $H_0 = 70.6 \pm 2.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$  from cepheid calibrated distance ladder measurements presented by Rigault et al. (2018)<sup>6</sup>. Marginalising over the Hubble constant is necessary, because the X-ray flux–mass scaling relation (equation 3.2) has a distance dependence. Therefore, inference of the amplitude of the X-ray scaling relation is systematically limited by our knowledge of the distance–redshift relation.

When inferring the scaling relation parameters from number counts, we need to account for the fact that the halo mass function is cosmology dependent, especially on the matter density parameter  $\Omega_M$  and the amplitude of matter fluctuations on the scale of 8 Mpc  $h^{-1}$ ,  $\sigma_8$ . Under the assumption of a cosmological constant, flatness and minimal neutrino mass, the other cosmological parameters have a minor impact on the halo mass function. For this reason, in this work we marginalise over the priors  $\Omega_M = 0.276 \pm 0.047$  and  $\sigma_8 = 0.781 \pm 0.037$ , derived by Bocquet et al. (2019b, hereafter Bo19) from the number counts analysis of 343 SZE selected galaxy clusters supplemented with gas mass measurements for 89 clusters and weak lensing shear profile measurement for 32 clusters.

### Priors on SZE $\zeta$ -mass relation

When performing the SPT-SZ cross calibration, we require priors on the SZE scaling relation parameters to infer the X-ray flux–mass and richness–mass scaling relation parameters. These priors are derived from the X-ray and WL calibrated number counts of SPT-SZ selected clusters as described in Bo19. The adopted values are reported in Table 3.1. These priors were derived simultaneously with the cosmological priors discussed above, and both rely on the assumption that the SPT-SZ selection function is well characterised and that the SZE-signal–mass relation is well described by equation (3.4). Note that Bo19 only considered SPT-SZ clusters with SZE-signal to noise  $\xi > 5$  and  $z > 0.25$ , while we adopt their results to characterize a sample with  $\xi > 4.5$  and  $z > 0.2$ . Considering that this is an extrapolation from typical masses of  $\sim 3.6 \cdot 10^{14} \text{ M}_\odot$  for  $\xi = 5$  to  $\sim 3.3 \cdot 10^{14} \text{ M}_\odot$  for  $\xi = 4.5$  we view this as a minor extrapolation.

### Priors on X-ray $L_X$ -mass relation

Our priors on the X-ray flux–mass relation (c.f. Table 3.1) have been determined by Bu19, who studied the X-ray luminosities of 59 SPT-SZ selected clusters observed with XMM-Newton. The authors then use priors on the SZE -signal–mass relation to infer the

---

<sup>6</sup>Given the still unresolved controversy on the exact value of the Hubble constant, the value adopted here has the benefit of not being in significant tension with any other published result.

Tabelle 3.2: Mean and standard deviation estimated from the one dimensional marginal posterior plots for the parameters of the X-ray scaling relation and the richness scaling relation. Besides the constraints of the mass trend of the X-ray–mass relation and the corresponding intrinsic scatter, we find good agreement among our different analysis methods and with the literature values. This provides strong evidence that our selection function modeling is adequate.

	$A_X$	$B_X$	$C_X$	$\sigma_X$	$A_\lambda$	$B_\lambda$	$C_\lambda$	$\sigma_\lambda$
liter.	$4.15 \pm 1.01$	$1.91 \pm 0.17$	$0.20 \pm 0.42$	$0.25 \pm 0.10$	$71.9 \pm 6.1$	$1.14 \pm 0.20$	$0.73 \pm 0.76$	$0.15 \pm 0.08$
SPT calibr.	$5.14 \pm 2.29$	$1.31 \pm 0.42$	–	$0.42 \pm 0.21$	$80.2 \pm 18.9$	$1.01 \pm 0.25$	$0.41 \pm 1.58$	$0.26 \pm 0.11$
X NC	$3.81 \pm 0.81$	$1.85 \pm 0.12$	$-0.29 \pm 0.32$	$0.22 \pm 0.15$		$74.8 \pm 14.0$	$1.03 \pm 0.17$	$0.73 \pm 0.40$
opt NC								$0.28 \pm 0.15$
2d NC	$3.00 \pm 1.37$	$1.19 \pm 0.11$	$-0.07 \pm 0.41$	$0.44 \pm 0.18$	$87.2 \pm 20.0$	$0.80 \pm 0.07$	$1.06 \pm 0.39$	$0.18 \pm 0.12$

luminosity–mass relation parameters. Given that we are ourselves interested in determining the X-ray scaling relation parameters, we adopt priors on these parameters only when carrying out the number counts experiment in optical richness, where we need to account for the X-ray selection of our sample without being able to directly constrain it. Furthermore, we employ these priors when determining the systematic uncertainty on the outlier probability  $\pi_t$ , the MARD-Y3 contamination  $\pi_c$  and the SPT-SZ incompleteness  $\pi_i$  (see definitions of these in Section 3.2.3).

### Priors on optical $\lambda$ -mass scaling relation

When analysing the number counts as a function of flux and redshift we adopt priors on the richness–mass relation derived by S15 from a sample of 25 SPT-SZ selected cluster, matched with DES redmapper selected clusters. In that work the SZE-signal–mass relation parameters were determined by fitting the SPT-SZ selected cluster number counts at fixed cosmology. The resulting constraints on the richness–mass relation are reported in Table 3.1. These priors are also utilised when estimating the systematic uncertainty on the outlier probability  $\pi_t$ , the MARD-Y3 contamination  $\pi_c$  and the SPT-SZ incompleteness  $\pi_i$  (Section 3.2.3).

## 3.4 Application to MARD-Y3 and SPT-SZ

In this section we present the results of validation tests on the MARD-Y3 sample by way of examining the consistency of the X-ray–mass and the richness–mass scaling relations derived using different methods. First, we present the cross calibration of the fluxes and richnesses using the externally calibrated SPT-SZ sample (Section 3.4.1). Then in Section 3.4.2, we derive the parameters of the X-ray–mass scaling relation from the X-ray number counts, the parameters of the richness–mass scaling relation from the optical number counts, and then explore the constraints on both scaling relations from a joint 2-dimensional X-ray and optical number counts analysis. We explore the implied cluster masses in Section 3.4.3, and in Section 3.4.4 we validate our selection functions by computing the probabilities of each

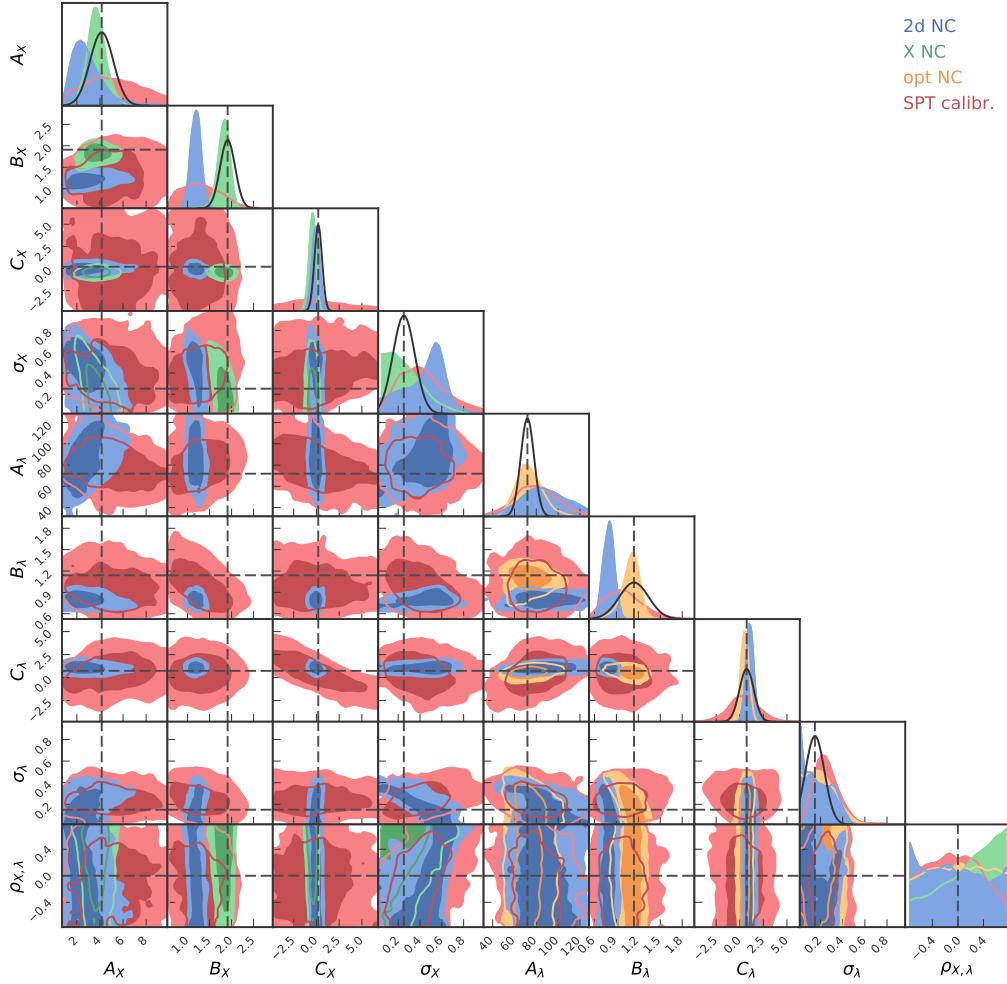


Abbildung 3.6: Marginal posterior contours of the free parameters in the SPT-SZ cross calibration (SPT calibr., red), the number counts in X-ray flux and redshift (X NC, green), the number counts in richness and redshift (opt NC, orange), and the number counts in X-ray flux, richness and redshift (2d NC, blue). In black the literature values from Bu19 and S15. The SPT-SZ cross calibration shows good agreement with the different number counts constraints and the literature. With the exception of the mass slope of the X-ray flux–mass relation inferred from 2d number counts, also the constraints from the different number count experiments show good agreement with the literature values. This provides strong evidence that our selection function modeling is adequate.

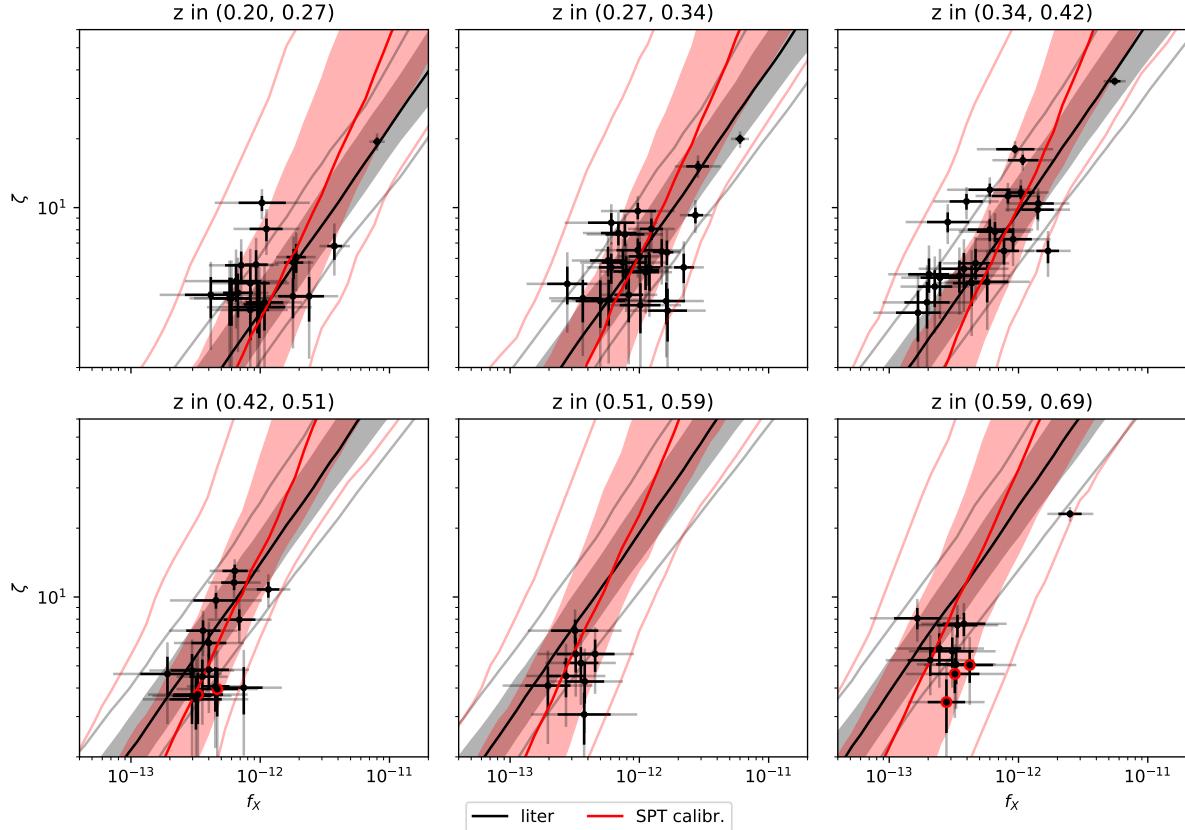


Abbildung 3.7: Black points mark the intrinsic X-ray flux  $f_X$  and SZE-signal to noise  $\zeta$  inferred from the respective error models for our cross-matched sample. The X-ray flux-SZE-signal relation is shown either marginalized over the literature priors (black and grey) or over the posterior of our cross calibration to SPT-SZ (red). While both sets of scaling relation parameters are consistent, we find a tendency for a weaker mass trend in the X-ray observable than reported in the literature. Marked as red points are MARD-Y3 false positives (an SPT-SZ cluster with  $< 0.025$  MARD-Y3 detection probability that is nonetheless is in MARD-Y3).

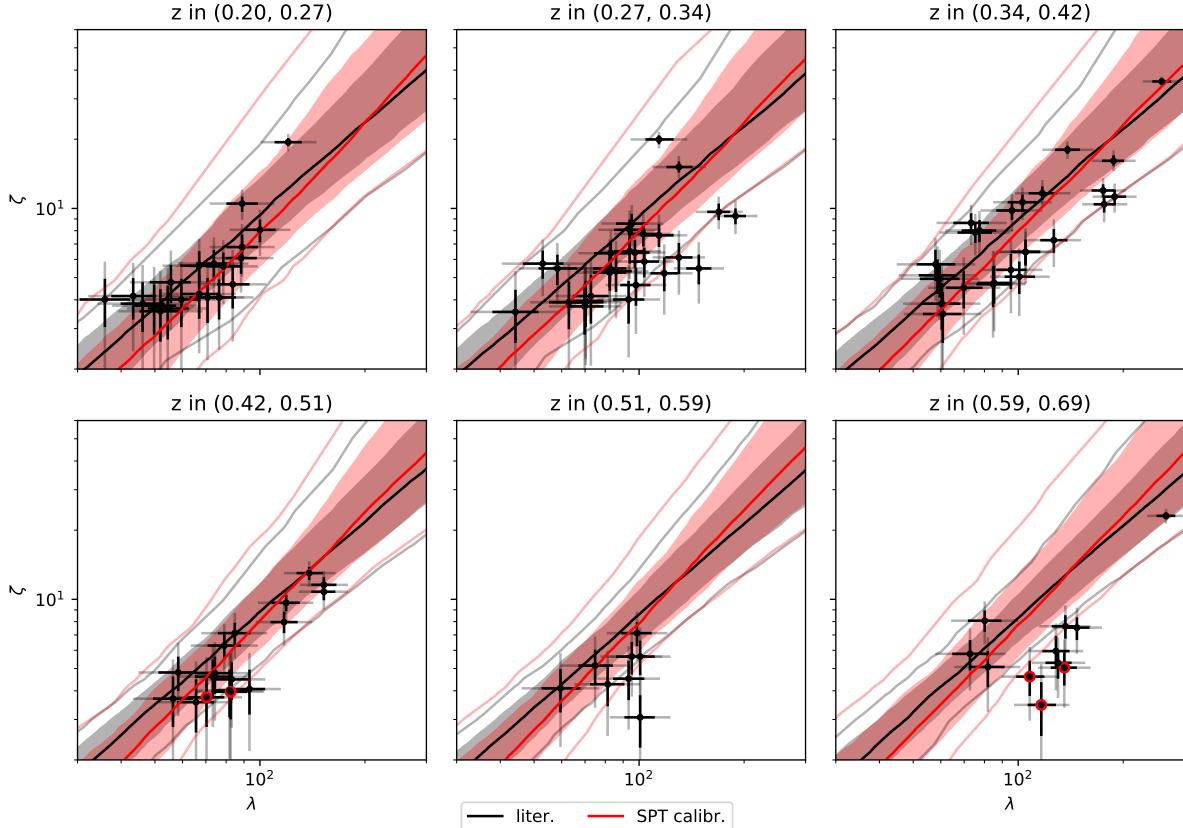


Abbildung 3.8: Black points mark the the intrinsic richness  $\lambda$  and SZE-signal to noise  $\zeta$  inferred from the respective error models for our cross-matched sample. The richness–SZE-signal relation is shown either marginalized over the literature priors (black) or over the posterior of our cross calibration to SPT-SZ (red). Both sets of scaling relation parameters are consistent, however, at intermediate redshift they fail to describe a part of the population with low SZE-signal and high richness. Marked as red points are MARD-Y3 false positives (SPT-SZ clusters with  $< 0.025$  MARD-Y3 detection probability that are nonetheless in MARD-Y3).

cluster in one sample having a counterpart in the other and comparing these probabilities to the actual set of matched pairs and unmatched single clusters in each sample. This last exercise allows us to study outliers in observable beyond the measured scaling relation and observational scatter and has implications for the incompleteness in the SPT-SZ sample and the contamination in the MARD-Y3 sample.

### 3.4.1 Validation using SPT-SZ cross calibration

As implied in the methods discussion in Section 3.2.1, the results of the SPT-SZ cross calibration of the MARD-Y3 mass indicators X-ray flux and richness are extracted by sampling the likelihood in equation (3.23). The free parameters of this fit are the parameters of the X-ray scaling relation ( $A_X, B_X, C_X, \sigma_X$ ), of the richness scaling relation ( $A_\lambda, B_\lambda, C_\lambda, \sigma_\lambda$ ) and the correlation coefficients between the intrinsic scatters ( $\rho_{X,\lambda}, \rho_{X,SZ}, \rho_{\lambda,SZ}$ ). We put priors on the parameters of the SZE-signal–mass relation ( $A_{SZ}, B_{SZ}, C_{SZ}, \sigma_{SZ}$ ) and on the cosmological parameters ( $H_0, \Omega_M, \sigma_8$ ), as described in Section 4.2.6.

The resulting marginal posterior contours on the parameters without priors are shown in red (SPT calibr.) in Fig. 3.6 and in Table 4.1. The same figure also shows as a black line the literature values for these parameters, where we use Bu19 for the X-ray parameters, and S15 for the optical parameters. Our constraints are in agreement with these works, but display comparable or larger uncertainties despite the larger number of objects. This is due to different effects.

The difference between these constraints on the richness–mass relation from S15 are mainly due to the tighter priors on the SZE-signal–mass relation parameters utilized in that work. For instance, the prior on the amplitude of the SZE-signal–mass relation is four times smaller than the one used in this work. On the other side, we analyze a 4 times larger sample, which warrants at best an improvement of the constraints by a factor of 2. Our larger uncertainties on the richness–mass relation parameters are thus reflecting our more conservative treatment of systematic uncertainties on the SZE inferred masses.

This issue does however not explain why our constraints on the luminosity–mass relation are weaker than those reported by Bu19, as that work used priors on the SZE-signal–mass relation comparable to ours. Two different effects play a role in this case. Firstly, the measurement uncertainty on the luminosities extracted from pointed XMM observations is much smaller than on RASS based luminosities, especially at high redshift. Secondly, we marginalize over the systematic uncertainty in the luminosity distance resulting from imperfect knowledge of the matter density  $\Omega_M$  and the Hubble parameter  $H_0$ . This source of uncertainty is not considered in Bu19. Both these effects are especially important at high redshift, likely resulting in our much weaker constraint on the redshift evolution parameter  $C_X$ .

In Fig. 3.7 we show the results for the cross calibration of the X-ray fluxes. In different redshift bins we plot the intrinsic X-ray flux inferred from the X-ray flux error model (equation 3.10) against the intrinsic SZE signal to noise inferred from the SZE error model (equation 3.13), as black points with 1 and 2 sigma uncertainties. We also plot the X-ray flux–SZE-signal relation obtained by combining the respective scaling relations. We show

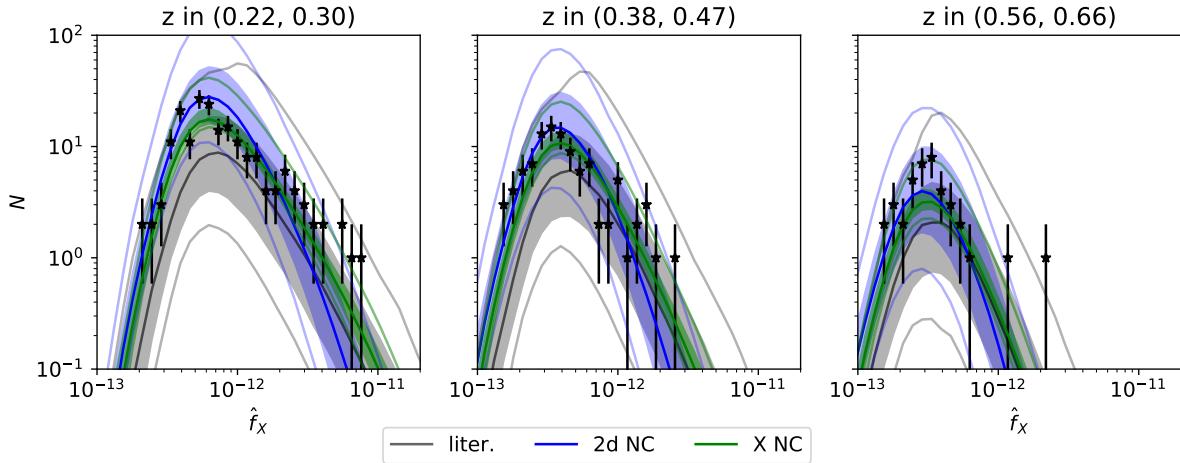


Abbildung 3.9: Measured number of clusters in bins of measured X-ray flux for different redshift bins as black points with Poissonian error bars. We over-plot the prediction for the number of objects with the uncertainties derived from the literature values (black and gray), from our 1 d fit (green), and from our 2d fit (blue). The latter captures adequately both the increasing rarity of high redshift objects, as well as the effect of X-ray incompleteness at low flux. The measurement is also consistent with the literature values, although as discussed in Section 3.4.2, our assumption that cosmological and scaling relation parameters are uncorrelated leads to an over-estimation of the uncertainty.

(black and grey) their marginalization over the Bo19 cosmological parameter and SZE-scaling parameter priors, the Bu19 X-ray-scaling parameter priors, and over the posterior of the SPT-SZ cross calibration (red). As already noted from the contour plots of the marginal posteriors, our inferred scaling relation parameters are statistically consistent with the literature. However, our calibration prefers a steeper relation, resulting from a lower inferred value on the X-ray mass trend  $B_X$ .

The results for the SPT-SZ cross calibration of the richness–mass relation are shown in Fig. 3.8. In different redshift bins we plot as black point the intrinsic richness  $\lambda$  and the intrinsic SZE-signal  $\zeta$  inferred from the respective error models (equations 3.12 and 3.13). We also plot the richness–SZE scaling derived from combining the richness–mass and the SZE-signal–mass relation. The resulting relation is shown with the uncertainties derived from the literature priors and the cross calibration posteriors. The two constraints are in very good agreement. Yet, at high redshift  $z > 0.5$ , we note the presence of a high richness, low SZE-signal population, not well described by the both the literature relation and our cross calibrated relation. These objects will be discussed in more detail in Section 3.4.4.

### 3.4.2 Validation using number counts

As described in the method discussion in Section 3.2.2, we perform three different number counts experiments in this work: (1) we infer the X-ray flux–mass relation by fitting for the

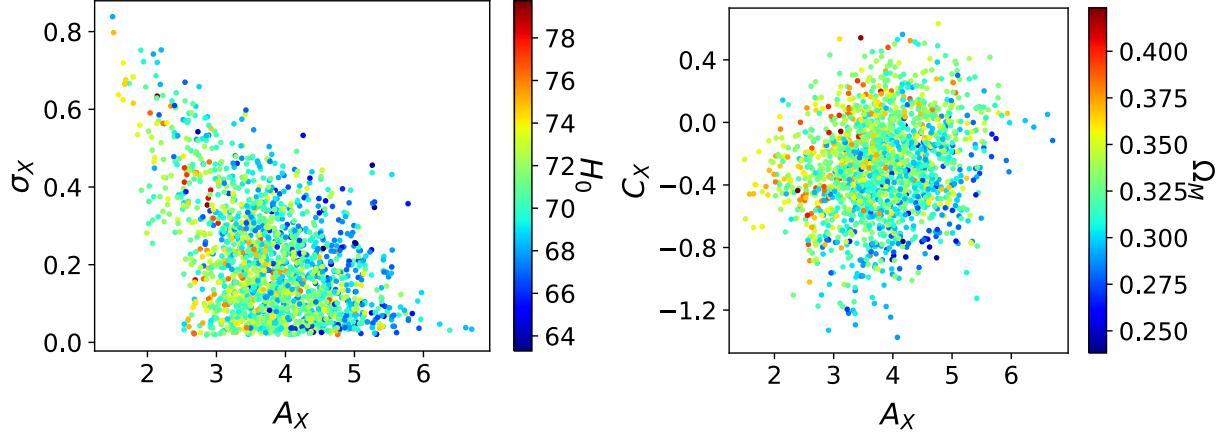


Abbildung 3.10: Examples of two 3 dimensional degeneracies between pairs of scaling relation parameters and cosmological parameters. For instance (lower panel), the lower right half of the  $(A_X, C_X)$  marginal plot is populated preferably by low  $\Omega_M$  points, while the contrary is true for the upper left half. The posterior thus lives on a relatively thin volume extending from the lower left for low  $\Omega_M$  to the upper right for large  $\Omega_M$ , leaving a significant fraction of the volume of the parameter space unpopulated. Similar behaviour can be seen in the upper panel. In summary, the posterior distribution occupies less parameter space volume than the expected from the marginal contours.

number counts of cluster as a function of measured flux and redshift, (2) we constrain the richness–mass relation by fitting for the number counts as a function of measured richness and redshift, and (3) we determine both relations by fitting the number of objects as a function measured flux, measured richness and redshift.

### X-ray number counts

While sampling the likelihood of the number counts in X-ray flux (equation 3.24), we let the parameters of the X-ray flux–mass relation ( $A_X, B_X, C_X, \sigma_X$ ) float within wide, flat priors. We adopt priors on the relevant cosmological parameters ( $H_0, \Omega_M, \sigma_8$ ) as described in Table 3.1. We also put priors on the richness–mass relation parameter ( $A_\lambda, B_\lambda, C_\lambda, \sigma_\lambda$ ). Furthermore, we empirically constrain the relation between X-ray detection significance  $\xi_X$ , measured flux  $\hat{f}_X$  and exposure time  $t_{\text{exp}}$  from the sample. As described in more detail in Sections 3.1.3 and 3.5.1, this results in four tightly constrained nuisance parameters that impact the X-ray selection function. The resulting posteriors on the X-ray scaling relation parameters are shown in green in Fig. 3.6. We find tight agreement with the literature values, at comparable accuracy on the marginal uncertainties.

In Fig. 3.9 we plot the number counts in measured X-ray flux bins in three different redshift bins with the respective Poissonian errors. We also plot the prediction for the number of objects in the same bins, once marginalized over the literature values (black and grey), and over our 1d fit (green). While for high fluxes the number of objects decreases due to the increasing rarity of such high mass objects, at low fluxes the decrease is due to

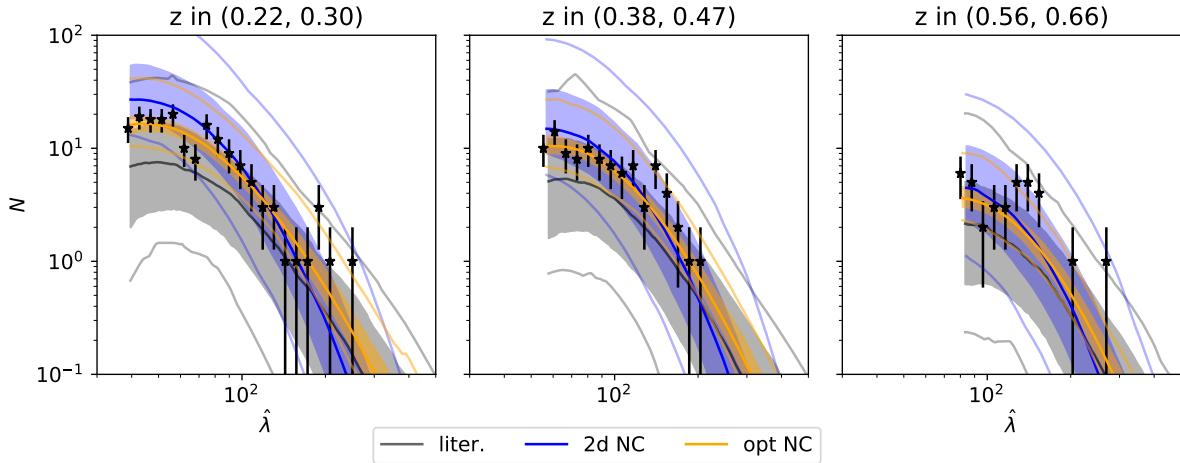


Abbildung 3.11: Number of objects in bins of measured richness  $\hat{\lambda}$  in different redshift bins as black points with Poissonian error bars. Over-plotted is the expected number of objects as a function of measured richness for the same bins marginalized over the uncertainties from the literature (black and grey) and from our fit (orange). The shape of the abundance at low richness would in principle be closer to a power-law, but at low richness the X-ray selection of the samples leads to a decrease in the number of objects, which is well fit by our selection model.

the incompleteness introduced by the X-ray selection. In both regimes, our model provides an adequate description of the abundance of clusters.

Similarly also the prediction from the literature provides a statistically consistent description of the data, albeit systematically one sigma low. Furthermore, this systematic uncertainty is larger than in the case of our own fit. This might seem to contradict the fact that the marginal posterior constraints are of comparable width as the constraints reported by the literature. However, note that when computing the prediction from the literature, we assume that the constraints on the cosmological parameters are independent from the constraints on the scaling relation parameters. While this holds true for any single pair of cosmological and scaling relation parameters, the number counts likelihood leads to hidden high dimensional degeneracies among scaling relation parameters and cosmological parameters, as can be seen in Fig. 3.10. In the upper panel, we show the well known amplitude-scatter degeneracy of number counts. Noticeably, the position along the degeneracy correlates with the value of  $H_0$ , as can be seen by the colour coding. Thus, in 3-d, the posterior occupies a relatively thin surface for each value of  $H_0$ , which however appears as an uncorrelated structure when projected in 2-d. This can also be noted in the lower panel of Fig. 3.10, which demonstrates that while  $A_X$  and  $C_X$  might be uncorrelated when marginalized over  $\Omega_M$ , parameter values with similar  $\Omega_M$  show stronger correlation. Consequently, the posterior from the number counts occupies a much smaller part of the parameter space than the prior from the literature, which assumes that all parameters are uncorrelated. These effects manifest themselves most drastically in the space of predicted

number of objects, which through its dependence on the different parameters dictates the high dimensional geometrical form of the hidden degeneracies.

### Optical number counts

Just as the number counts as a function of measured flux can be used to infer the X-ray scaling relation parameters, the number counts in richness can be used to infer the richness–mass relation parameters. To this end, we sample the likelihood of number counts in richness bins (equation 3.27). We let the parameters of the richness–mass relation ( $A_\lambda$ ,  $B_\lambda$ ,  $C_\lambda$ ,  $\sigma_\lambda$ ) free, while we adopt priors from the literature on the cosmological parameters ( $H_0$ ,  $\Omega_M$ ,  $\sigma_8$ ). Importantly, modeling the X-ray incompleteness in the space on measured richness requires a way to transform from measured richness to X-ray flux. Thus, while the transformation from richness to mass is fit, we need to assume a transformation from mass to X-ray flux. This is done by putting priors on the X-ray scaling relation parameters ( $A_X$ ,  $B_X$ ,  $C_X$ ,  $\sigma_X$ ). As for the X-ray number counts, we empirically constrain the relation between X-ray detection significance  $\xi_X$ , measured flux  $\hat{f}_X$  and exposure time  $t_{\text{exp}}$  from the sample and predict the X-ray selection function on the fly.

The resulting marginal posterior contours are shown in Fig. 3.6 in orange. We find good agreement with the literature values and with the SPT-SZ cross calibration. The marginal uncertainties are comparable to the literature values, despite being marginalized over cosmological parameters. We also find that the constraints from the number counts are more stringent than those derived from the SPT-SZ cross calibration.

One can visually assess the quality of the resulting fit in Fig. 3.11, where we plot the number of objects in measured richness for different redshift bins as black points with Poissonian error bars. We also plot the predicted number of objects with the uncertainties derived from the literature priors (black and grey) and from our fit (orange). As noted earlier with the X-ray number counts, the prediction from our fit has smaller uncertainties than those derived from the literature despite the comparable size of the marginal uncertainties on the underlying parameters. Given the very similar structure of the employed likelihood when compared to the X-ray number counts, this effect is most likely also explained by degeneracies between cosmology and scaling relation parameters that are not captured by assuming uncorrelated priors from the literature. Also in this case we note that the literature prediction is systematically 1 sigma low.

### Combined X-ray and optical number counts

We also fit for the abundance of clusters as a function of measured X-ray flux  $\hat{f}_X$ , measured richness  $\hat{\lambda}$  and redshift, which we will refer to a ‘2d number counts’, by sampling the likelihood in equation (3.30). We allow the parameters of both the X-ray scaling relation ( $A_X$ ,  $B_X$ ,  $C_X$ ,  $\sigma_X$ ) and the richness scaling relation ( $A_\lambda$ ,  $B_\lambda$ ,  $C_\lambda$ ,  $\sigma_\lambda$ ) float within wide, flat priors. We adopt priors on the cosmological parameters from Table 3.1. Furthermore, we empirically constrain the relation between X-ray detection significance  $\xi_X$ , measured flux  $\hat{f}_X$  and exposure time  $t_{\text{exp}}$  from the MARD-Y3 sample and predict the X-ray selection

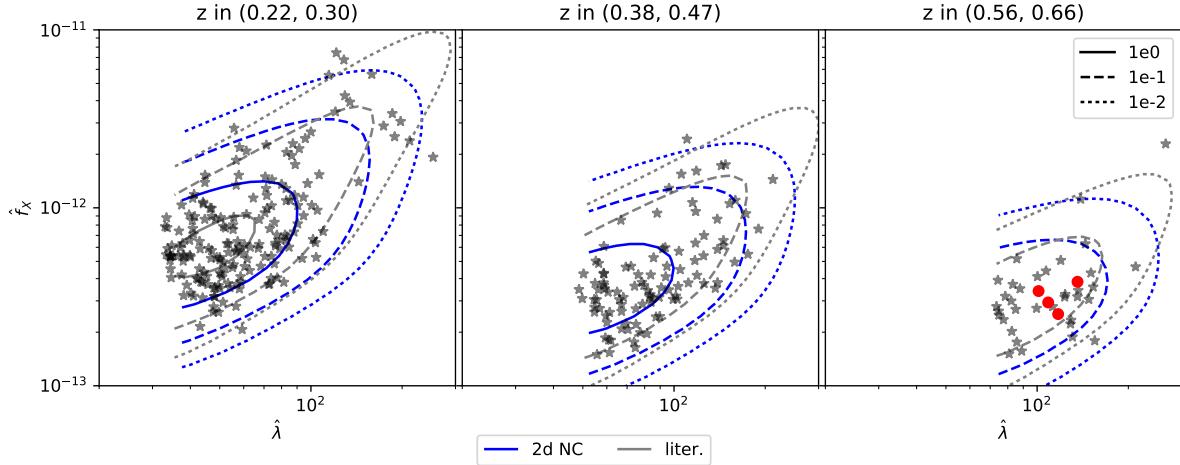


Abbildung 3.12: Number counts in both measured richness  $\hat{\lambda}$  and measured X-ray flux  $\hat{f}_x$  visualized by presenting in different redshift bins the distribution of our sample (black stars), the contours of the 2-dimensional predicted number of objects for the literature values (grey, liter.) and for the mean of the posterior of our fit to the data (blue, 2d NC). Our 2d fit prefers larger scatter and provides a better description of the data than does the literature prediction. Marked as red points are SPT-SZ false positives (SPT-SZ clusters with a MARD-Y3 detection probability  $< 0.025$  that are nonetheless in MARD-Y3).

function on the fly (c.f. Sections 3.1.3 and 3.5.1).

In Fig. 3.6 we show the marginal posterior contours on the scaling relation parameters in blue. We find good agreement with the results from the SPT-SZ cross calibration on all parameters. When comparing the constraints from 2d number counts (blue) on the X-ray scaling relation parameters to the constraints from the number counts in X-ray flux (green), we find good agreement on the values of the amplitude and redshift evolution. However, we find a shallower X-ray observable mass trend than from the X-ray number counts, and we see a similar shift in the optical mass trend parameter, although in this case the statistical significance is small. Given the agreement of the X-ray number counts result is with Bu19, the results from the 2d number counts are in some tension with both. As shown in Section 3.4.3 below, these constraints however do not result in statistically inconsistent mass estimates. Nevertheless, possible systematic effects impacting our validation tests are discussed in section 3.5.1 and 3.5.2.

Of interest is also the constraint the 2d number counts put on the two intrinsic scatters in X-ray flux and richness. Inspecting their joint marginal posterior in Fig. 3.6 reveals a distinct degeneracy in the form of an arc. This is the natural result of the fact that the 2d number counts can only constrain the total scatter between the two observables, but not the two individual scatters between each observable and mass. The total scatter between observables, being the squared sum of the individual scatter, sets the radius of the arc. Noticeably, this arc-like degeneracy excludes the possibility that both the X-ray and the richness scatter are small.

For visual inspection of the 2d number counts fit in Fig. 3.12 we present the distribution in measured X-ray flux and measured richness of our sample in different redshift bins as black stars. We also plot the contours of the predicted number of objects in equally spaced logarithmic bins: in blue the prediction for the best fit value of the 2d number counts, while in grey the prediction from the literature. The selection in richness due to the  $f_c < 0.05$  cut is at every redshift a sharp cut in measured richness, as can be seen up to the intra bin scatter due to the large bins used for plotting. The effect of the X-ray selection function is harder to see, but can be appreciated in the shape of the contours at low flux: they show a bend, predicting very small numbers of objects at the lowest fluxes. Notably, the distribution of the data displays a large dispersion, which is better captured by our fit (blue) than by the prediction from the literature (grey). This confirms that the measurement of a larger X-ray scatter is indeed a feature of the data visible in the 2 dimensional cluster abundance. Despite the larger intrinsic scatter, 2d number counts posterior provide also an prediction of the X-ray and optical 1d number counts that is consistent with the data within the systematic and statistical uncertainties, as can be see by the blue predictions in Fig 3.9 and 3.11.

### 3.4.3 Validation using cluster masses

In this section we investigate the prediction of the individual halo masses derived from the different constraints on the scaling relation parameters extracted above.

To estimate the masses for each cluster given its measured X-ray flux  $\hat{f}_X^{(i)}$  (or analogously the measured richness  $\hat{\lambda}^{(i)}$ ), we compute the distribution of probable masses

$$P(M|\hat{f}_X^{(i)}, z^{(i)}, \vec{p}) \propto \int df_X P(\hat{f}_X^{(i)}|f_X) P(f_X|M, z^{(i)}, \vec{p}) \frac{dN}{dM} \Big|_{M, z^{(i)}, \vec{p}}, \quad (3.37)$$

where  $P(f_X|M, z^{(i)})$  is the mapping between intrinsic flux and mass obtained by only considering the first component of equation (3.5). Note also that the above equation needs to be normalized in such a way that  $\int dM P(M|\hat{f}_X^{(i)}, z^{(i)}, \vec{p}) = 1$ , which sets the proportionality constant.

The X-ray mass  $M_X$  (and analogously the optical mass  $M_\lambda$ ) can then be estimated as

$$\ln M_X^{(i)}|_{\vec{p}} = \int dM P(M|\hat{f}_X^{(i)}, z^{(i)}, \vec{p}) \ln M. \quad (3.38)$$

Note that these masses naturally take account of the Eddington bias, which is fully described by equation (3.37).

The X-ray and optical masses are affected by systematic uncertainties in the scaling relation and cosmological parameters. We capture this uncertainty in each case by marginalising the mass posterior over the appropriate posterior distribution of the parameters that we determined above. We marginalize the mass over different scaling relation parameter posteriors, including those from the literature (liter.), those from the SPT-SZ cross

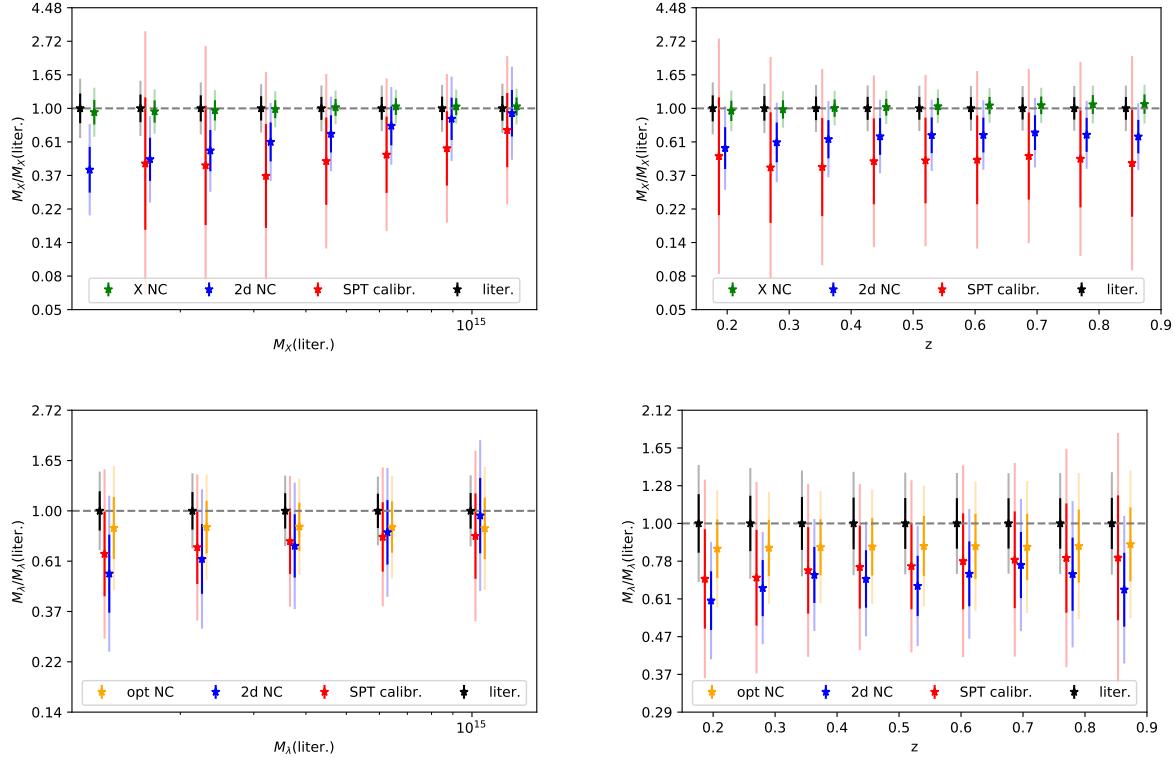


Abbildung 3.13: Ratio of the masses derived by our analysis methods and the masses derived from the literature values, for masses inferred from their measured X-ray flux ( $M_X$ , upper row) and their measured richness ( $M_\lambda$ , lower row), as functions of mass (left column) and of redshift (right column), together with the 1 and 2-sigma systematic uncertainties on the individual masses due to the incomplete knowledge of the scaling relation parameters. The masses we recover from SPT-SZ cross calibration (SPT calibr, red) and different flavours of number counts, while being in most cases systematically low, are statistically consistent with the masses inferred by adopting the literature values. Tension beyond 1 sigma, but still smaller than 2 sigma appear at the low mass end of the inferred X-ray masses.

calibration (SPT calibr.), and those from the combined X-ray and optical number counts (2d NC), the X-ray number counts (X NC) and the optical number counts (opt NC). The mass posteriors are derived for all clusters in the MARD-Y3 sample.

In the upper row of Fig. 3.13 we present the ratio between the X-ray masses derived from our posteriors to the X-ray masses derived from the literature (Bu19) as a function of inferred literature mass (left panel) and of redshift (right panel). We find that the mass inferred from the number counts in X-ray flux is consistent with the literature values, while the masses inferred from the 2d number counts and the SPT-SZ calibration are lower than the literature masses. In the case of the SPT masses the difference never exceeds one sigma at all redshifts and masses we considered. For the 2d number count masses, we find that they are 1 sigma low at all redshifts, and up to 2 sigma low at masses of  $1-2 \cdot 10^{14} M_{\odot}$ . At masses of around  $10^{15} M_{\odot}$  they are in perfect agreement with the other mass estimates. This is due to the different values of inferred mass trend. As a function of redshift, the masses inferred from 2d number counts and the SPT-SZ calibration are also lower, reflecting on one side the prevalence of low mass systems. Furthermore, this shift is also due to the larger intrinsic scatter recovered from the 2d number counts and the SPT-SZ calibration, that together with the shallower mass slope leads to a larger intrinsic mass scatter. This results in larger Eddington bias corrections and ultimately lower inferred masses. At the current level of statistical and systematic uncertainty we conclude that different methods predict mutually consistent individual masses from the X-ray flux at less than 2 sigma. Yet, the magnitude of the intrinsic scatter of the X-ray luminosity at fixed mass and redshift, together with its mass trend, are indications of possible internal tensions and unresolved systematics (c.f. Section 3.5.2).

In the lower row of Fig. 3.13 we also show the ratio between the optical mass inferred from our fits to the one inferred from the literature value (S15). Here we find that all our methods provide a lower, yet statistically consistent mass estimate. The difference is likely due to an analysis choice in the literature values. Namely, S15 utilizes priors for the SZE-scaling relation parameters derived from fitting the SZE number counts at fixed cosmology. In that work, however, the CMB derived cosmology from Planck Collaboration et al. (2014) was used, which results in  $A_{\text{SZ}, \text{S15}} = 4.02 \pm 0.16$ , and therefore is an overestimation of masses by  $\sim 18\%$  compared to our work. This shift accounts for most of the shifts seen in  $M_{\lambda}$  here. Even without this correction, at the current level of systematical uncertainties, the individual optical masses inferred from our different analysis methods are mutually consistent. This is expected because our  $A_{\text{SZ}}$  prior is consistent with the value used by S15. Furthermore, while 2d number counts predict a shallower mass trend than all other methods, in the mass range we consider this does not lead to significant tension with the other analysis methods.

This consistency check of mass estimates underscores the importance of weak lensing mass calibration as a component of the validation of cluster samples. If the cosmology marginalized constraints on cluster masses from weak lensing are not consistent with those from cluster counts, then this would be clear evidence of an inadequacy in the selection model or an unaccounted for bias in the weak lensing calibration analysis. As noted previously, we will examine the validation with the weak lensing constraints in a forthcoming

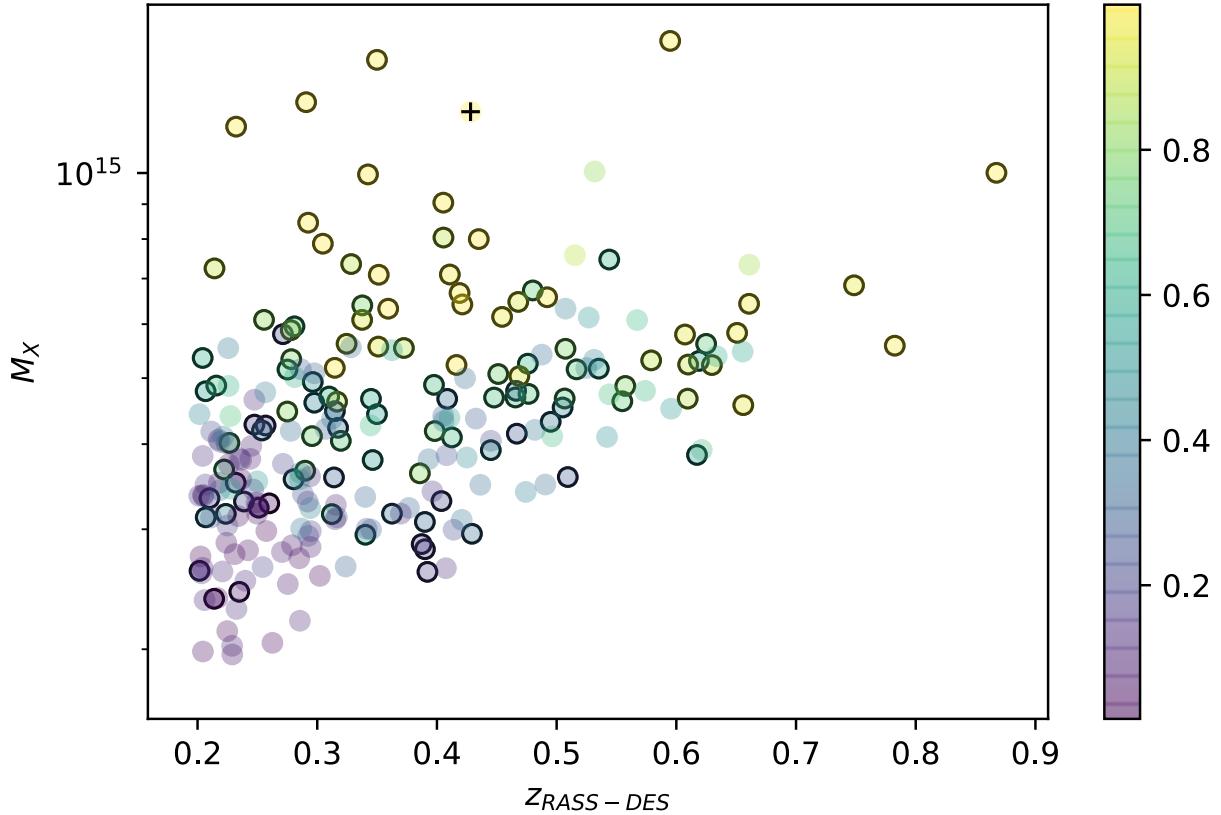


Abbildung 3.14: MARD-Y3 sample in the joint SPT-DES Y3 footprint at redshift  $z > 0.2$ . Color encodes the probability of an SPT-SZ detection for each object, showing the characteristic mass selection of the SPT-SZ catalog. Black circles indicate matched clusters, while the cross marks MARD-Y3 true negatives (clusters with high SPT-SZ detection probability and yet no match).

analysis.

#### 3.4.4 Validation using independent cluster samples

Having established in the section above that our selection function modeling allows us to infer the masses of the MARD-Y3 clusters consistently to within the systematic uncertainties, we now move to a further test of the selection functions of the two samples.

As described in the methods Section 3.2.3, we investigate the SPT-SZ and MARD-Y3 selection functions by comparing the probability of each MARD-Y3 object being detected by SPT-SZ to the actual occurrence of such a detection. As established earlier, there are 123 clusters in the cross-matched sample, but the validation we do here also uses information from unmatched clusters. This will provide an estimate of SPT-SZ incompleteness as well as MARD-Y3 contamination. We then consider the SPT-SZ sample and compute the probability that an SPT-SZ cluster is detected in MARD-Y3. In this case, we also constrain

Tabelle 3.3: List of MARDY-Y3 true negatives (those with high SPT-SZ detection probability but no detection). Probabilities are given as the 16th percentile.

2RXS name	RA [°]	DEC [°]	$z_{\text{RASS}}$	$\hat{f}_X$ [erg s <sup>-1</sup> cm <sup>-2</sup> ]	$\hat{\lambda}$	$L_X$ [erg s <sup>-1</sup> ]	$p_{\text{S.M.}}^{\text{liter.}}$	$p_{\text{S.M.}}^{\text{SPT c.}}$	$p_{\text{S.M.}}^{\text{2d NC}}$	Fig.
2RXS J042056.4-524651	52.69	-52.48	0.428	2.249e-12	110.3	$1.609 \cdot 10^{45}$	>0.976	>0.994	>0.99997	3.7.2

Tabelle 3.4: List of SPT-SZ false positives (SPT-SZ clusters with low MARF-Y3 detection probability that were nevertheless detected). Probabilities are given as the 68th percentile. *Comments:* 1) classified only when marginalising over the SPT-SZ cross calibration posterior. 2) classified only with the literature priors or the 2d number counts posterior. 3) different redshifts, discussed in Section 3.3.1.

SPT-SZ name	RA [°]	DEC [°]	$z_{\text{SPT}}$	$\xi$	$p_{\text{M.S.}}^{\text{liter.}}$	$p_{\text{M.S.}}^{\text{SPT c.}}$	$p_{\text{M.S.}}^{\text{2d NC}}$	Fig.	visual inspection
SPT-CL J0218-4233	34.55	-42.55	0.638	4.517	<0.007	<0.015	<0.018	3.7.2	
SPT-CL J0202-5401 <sup>3)</sup>	30.57	-54.02	0.711	5.653	<0.001	<0.003	<0.003	3.7.2	2 structures in LoS, c.f. K19, Fig.18
SPT-CL J0111-5518	17.84	-55.31	0.495	4.544	<0.031	>0.045	<0.026	3.7.2	radio source nearby
SPT-CL J0048-4548	12.25	-45.80	0.476	4.831	<0.024	<0.041	<0.017	3.7.2	radio source nearby
SPT-CL J0041-5107	10.29	-51.13	0.510	4.598	<0.028	<0.043	<0.023	3.7.2	radio source nearby
SPT-CL J2332-5053 <sup>3)</sup>	353.02	-50.89	0.567	4.581	<0.003	<0.008	<0.007	3.7.2	mismatch
SPT-CL J2034-5936 <sup>3)</sup>	308.54	-59.60	0.916	8.529	<0.022	<0.052	<0.012	3.7.2	mismatch
SPT-CL J0324-6236	51.05	-62.60	0.755	8.755	<0.013	<0.024	<0.018	3.7.2	X-ray point sources (Chandra)
SPT-CL J0430-6251	67.71	-62.85	0.782	5.292	<0.001	<0.002	<0.006	3.7.2	mismatch
SPT-CL J0010-5112 <sup>1)</sup>	2.74	-51.21	0.210	4.515	<0.112	<0.06	<0.042	3.7.2	
SPT-CL J2146-5736 <sup>2)</sup>	326.69	-57.61	0.608	6.190	<0.020	<0.055	<0.030	3.7.2	complex X-ray structure
SPT-CL J0233-5819 <sup>2)</sup>	38.25	-58.33	0.657	6.546	<0.031	<0.068	<0.047	3.7.2	radio source nearby

the outlier fraction beyond the log-normal scatter, more precisely the fraction of objects with an abnormally high X-ray flux or optical richness, or a surprisingly low SZE-signal.

### SPT-SZ detection of MARD-Y3 clusters

In Fig. 3.14 we show the MARD-Y3 cluster sample in the joint SPT-DES Y1 footprint, plotted as a function of the X-ray derived mass and the redshift presented by K19. Note that the mass used in this plot is used solely for presentation purposes, and does not go into any further calculation. We color-code the MARD-Y3 clusters based on their SPT-SZ detection probability  $p_{\text{S.M.}}^{(i)}$ , computed following equation (3.35). This prediction reflects the mass information contained in each cluster's measured flux  $\hat{f}_X^{(i)}$  and measured richness  $\hat{\lambda}^{(i)}$ . It also nicely visualizes the approximate mass selection at  $M \gtrsim 3 \times 10^{14} M_\odot$  of the SPT-SZ sample.

We place black circles around the matched clusters. When determining the detection probabilities using the literature values for the scaling relation parameters, we identify six clusters that have high detection probability, but are not matched, so called *true-negative*. However, when determining the detection probabilities either from the posterior of our SPT-SZ cross-calibration or the 2d number counts, only one of these systems is confirmed as a MARD-Y3 true negative: 2RXS J042056.4-524651.

Several characteristic quantities for this object are presented in Table 3.3. We visually inspect its DES image, its RASS count rate map, its SPT-SZ signal-to-noise map and its MCMF galaxy density map in Fig 3.7.2. The line of sight is occupied by one part of the

merging cluster system A 3126 at  $z \sim 0.05$  (white ellipticals all over the picture) and by a massive cluster with a strong lensing arc at  $z \sim 0.43$  (yellow, diffuse galaxy in the upper left quadrant with a blue arc to the south east). Both application of MCMF to RASS and SPT-SZ find these optical structures with  $f_c = 0$ . Exactly for this reason, however, the automated redshift assignment fails: MCMF on RASS selects  $z_{\text{RASS}} = 0.428$ , while MCMF on SPT-SZ on the other hand selects  $z_{\text{SPT}} = 0.056$ . Given  $z_{\text{SPT}} = 0.056$ , the SPT-SZ object does not fall into our redshift selection  $z_{\text{SPT}} \in (0.2, 1)$ .

More generally, in the case of practically equal  $f_c$ s random rounding errors determine which redshift is selected with smaller  $f_c$ . This automated redshift assignment is likely inaccurate in this specific case, as the X-ray flux is probably dominated by the  $z \sim 0.05$  cluster, while the SPT-SZ signal most likely originates from the  $z \sim 0.43$  cluster, as also reported by Bleem et al. (2015). Such alignments are rare due to the inherent sparsity of clusters on the sky. To keep the pipeline automated and avoid human decision making, and the resulting biases, we do not apply any special treatment to this object. It is just an outlier that does not bias our statistical analysis, as shown in the following<sup>7</sup>.

We also aim to constrain the occurrence of contamination in the MARD-Y3 sample by introducing the probability  $\pi_c$  that a MARD-Y3 object is not a cluster, and should therefore not be detected by SPT. Simultaneously, we also introduce the probability of SPT-SZ incompleteness  $\pi_i$  that any MARD-Y3 cluster that should be detected by SPT, is not detected (c.f. Fig. 3.4). This allows us to use the actual list of detections and non-detections together with the raw probabilities of detection, to constrain these extra probabilities, as discussed in equation (3.36). We find that  $\pi_c$  and  $\pi_i$  are degenerate parameters, with the difference between the two being constrained by our data. Under the assumption of a MARD-Y3 contamination of  $\pi_c = 0.025$ , as derived by K19 for the  $f_c < 0.05$  sample used here, we find  $\pi_i = 0.278 \pm 0.043(\text{stat.})^{+0.108}_{-0.146}(\text{sys.})$ , when marginalising over the literature priors. When marginalizing over the SPT-SZ calibration posterior we find  $\pi_i < 0.027$  (stat.) and  $\pi_i < 0.030$  (sys.) at 68% confidence, while we find  $\pi_i < 0.030$  (stat.) and  $\pi_i < 0.242$  (sys.) at 68% confidence when marginalising over the 2d number counts constraints together with the priors from Bo19 on the SZE-signal scaling relation.

The difference in inferred central value for the SPT-SZ incompleteness is due to the different mass predictions when using the literature priors as compared to our fits. As discussed in Section 3.4.3, our SPT-SZ cross calibration and our 2d number counts analysis imply lower X-ray and optically derived masses than the literature priors. This systematically lowers the SPT-SZ detection probability of MARD-Y3 clusters, resulting in different incompleteness probabilities when comparing to the actual number of matched objects. We interpret this as another piece of evidence that the SPT-SZ cross calibration and the 2d number counts provide a more accurate picture of the observable–mass relation than

<sup>7</sup>For completeness it is worth mentioning that an XMM observation by Werner et al. (2007) identified the high redshift structure via a redshifted Fe-line with  $z = 0.44$  in the X-ray spectrum, confirmed by optical spectroscopy of the central galaxy with Magellan. Furthermore, the central galaxy hosts a radio loud AGN, and is also detected in the Sydney University Molonglo Sky Survey (SUMSS, Mauch et al., 2003). To the authors knowledge, the impact of this radio source on the SZE-signal in SPT-SZ has not been investigated.

the literature priors. In fact, they reveal that the scatter around our luminosity–mass relation is larger than the scatter found by Bu19. Yet, within the statistical and systematic uncertainties the results are still in agreement.

Another interesting aspect is the magnitude of the statistical and systematic uncertainty on the SPT-SZ contamination. Note that the statistical uncertainties when marginalizing over the different posteriors are comparable. This reflects the fact that they are derived from a sample of a given size. The minor differences can be appreciated by noting that in equation (3.36) the individual clusters likelihood of  $\pi_i$  are weighted by the detection probabilities  $p_{S,M}^{(i)}$ , which are different depending on which posterior is used to compute them. On the other hand, the magnitude of the systematic uncertainty introduced by the marginalization over the different posteriors is quite different. Marginalizing over the SPT-SZ cross calibration posterior provides the smallest systematic uncertainty. This is expected when considering that the SPT-SZ cross calibration constrains  $P(\zeta | \hat{f}_X^{(i)}, \hat{\lambda}^{(i)}, z^{(i)})$  (c.f. equations 3.21-3.23), which is the major source of systematic uncertainty when computing the SPT-SZ detection probabilities of MARD-Y3 cluster (c.f. equation 3.35). We also expect that this introduces high dimensional degeneracies among the sampled parameters. The fact that both the literature priors and the 2d number counts do not follow these degeneracies leads to larger systematic uncertainties on the SPT-SZ incompleteness.

### MARD-Y3 detection of SPT-SZ clusters

We also test the MARD-Y3 selection function by computing the probability of detecting each of the SPT-SZ clusters in the DES-Y3 footprint. In Fig. 3.15 we show the SPT-SZ sample as a function of redshift and SZE derived mass. Note that the SZE derived mass shown in this figure is only used for presentation purposes. Color encodes the MARD-Y3 detection probability, computed following equation (3.33). The color coding reflects the approximate flux selection of the MARD-Y3 sample. As black circles we highlight the matched clusters.

Out of the 123 clusters in the cross-matched sample, we identify 11 false-positives, which are SPT-SZ clusters that show up in MARD-Y3 even though they have a  $< 0.025$  MARD-Y3 detection probability, calculated by marginalising over the literature values or the 2d number counts posterior. These clusters are marked in Fig. 3.15 with crosses. When marginalising over the SPT-SZ cross calibration posterior, we find 10 false positives, 9 of which are in common with the aforementioned. In Table 3.4 we present several properties of these objects, while we visually inspect (see Fig. 3.7.2-3.7.2).

Among the 12 false positives ( $10 + 2$  unique from the original 11), 1 object (Fig. 3.7.2) is in a line of sight with two optical structures and is already discussed in K19 (see figure 18). Both X-ray flux and SPT-signal are likely due to contributions from both objects, and thus unreliable. This object is not included in the SPT-SZ cross calibration, as discussed in Section 3.3.1. Inspection of the galaxy density maps at the two redshifts reveals that these two structure align almost perfectly. Three objects furthermore are mismatches due to too generous positional matching. Two of these objects were excluded from the SPT-SZ cross calibration due to the significant difference between the SPT-SZ and MARD-Y3 redshift

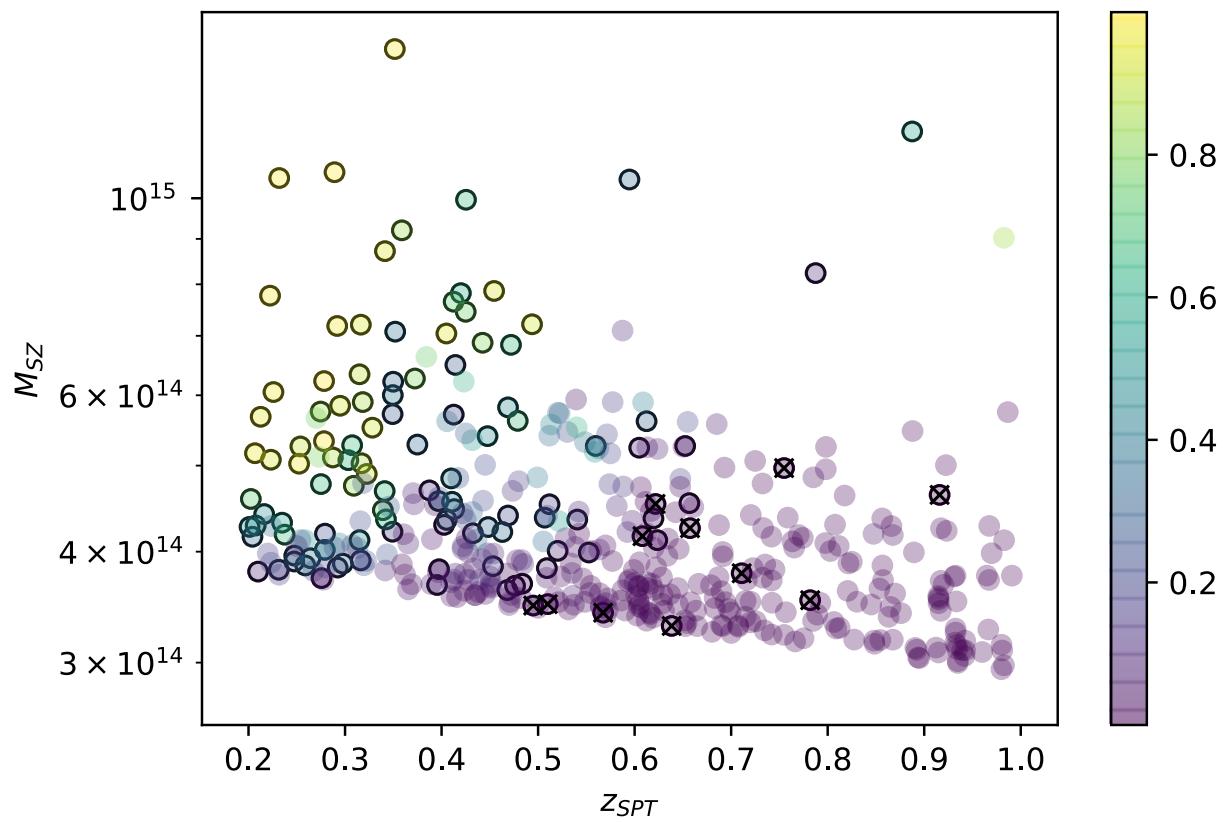


Abbildung 3.15: SPT-SZ selected sample in the joint SPT-DES Y3 footprint. Color encodes the probability of MARD-Y3 detection, showing the characteristic flux selection of an X-ray survey. Black circles indicate matched clusters, while crosses mark false negative (SPT-SZ clusters with  $< 0.025$  MARD-Y3 detection probability that were matched nonetheless).

(c.f. 3.3.1).

The remaining 8 objects are shown as red circles in Fig 3.7 and Fig 3.8, from which we conclude that they are predominantly members of the already discussed class of low SZE-signal / high X-ray or richness objects. We find that these objects do not display any special properties in the X-ray flux–lambda plot, as can be seen by the red point marking them in Fig. 3.12. The clusters outside of the redshift bins chosen for visualisation similarly are not outliers in the X-ray flux – richness population. This gives an indication that these objects likely have an unexpectedly low SZE-signal, rather than a too high X-ray flux. If the latter were the case, we would expect them to populate the low richness – high X-ray flux part of the population.

Furthermore, in the visual inspection these objects look like genuine clusters in X-ray, SZE and optical. In the optical they show a diffuse and bright central galaxy surrounded by similarly coloured smaller galaxies. Their SZE-signal also displays a clear round and extended structure. Their X-ray signatures are more irregular, at a lower contrast but still clearly visible.

Noticeably, however, 4 of them (SPT-CL J0111-5518, SPT-CL J0048-4548, SPT-CL J0041-5107, SPT-CL J0233-5819) are in the vicinity of radio sources identified in the Sydney University Molonglo Sky Survey (SUMSS, Mauch et al., 2003). This may lead to a bias in their SZE-signals, depending on the flux of these radio galaxies at the SPT frequencies of 90 and 150 GHz. One object (SPT-CL J0324-6236) reveals the presence of X-ray point sources in the Chandra image. This could bias the X-ray flux high. For the other objects, no high resolution X-ray imaging is available.

Given that most of the false positives are not accounted for, but display typical contamination signature beyond the log normal scatter around the scaling relations, we further investigate the probability that an SPT-SZ object that should not be matched by MARD-Y3 is matched anyway. It cannot *a priori* be excluded that, intrinsically, the distribution of X-ray luminosities or SZE-signals at fixed mass displays tails beyond the assumed log-normal scatter model. Such tails would lead to unexpected detections. The probability of a cluster living in such a tail, i.e. being an outlier, is given by the parameter  $\pi_t$ . Taking account of the detection probabilities and the actual occurrence of detections, we use the likelihood presented in equation (3.34). We find  $\pi_t = 0.082^{+0.018}_{-0.017}(\text{stat.}) \pm 0.025(\text{sys.})$ , when marginalising over the literature priors. When marginalizing over the SPT-SZ cross calibration posterior we find  $\pi_t = 0.090^{+0.018}_{-0.017}(\text{stat.}) \pm 0.038(\text{sys.})$ , while we find  $\pi_t = 0.101^{+0.019}_{-0.018}(\text{stat.}) \pm 0.026(\text{sys.})$  when marginalising over the 2d number counts constraints together with the priors from Bo19 on the SZE-signal scaling relation.

Note that the outlier probability is stable with respect to the posterior we marginalise over. While this might be surprising at first, it is actually due to the fact that we use the SZE-signal to predict the expected fluxes and richnesses (c.f. equations 3.31- 3.33). Using the same prior on the SZE–mass scaling relation parameters in all three cases, we predict comparable fluxes and richnesses, although with larger scatter in the case of the SPT-SZ cross calibration and the 2d number counts. We thus determine with some confidence that  $\sim 10\%$  of the SPT-SZ clusters display either abnormally high X-ray fluxes or abnormally low SZE-signals that are not adequately explained by the selection models we explore here.

In their section 4.2.2, K19 study the subset of MARD-Y3 clusters containing WISE sources (thus, clusters that are candidates for containing AGN) showing evidence of AGN boosting of the X-ray flux, given their optical richnesses, in only  $\sim 1\%$  of clusters. Indeed, all these SPT-SZ false positives show X-ray properties consistent with those expected given their richnesses. On the other hand, a recent study of the high frequency radio galaxy luminosity function in low redshift X-ray selected clusters indicates that  $\sim 5\%$  of the SZE selected clusters could be removed from the SPT-SZ sample due to radio AGN flux reducing their SZE signatures (Gupta et al., 2017). Contextual evidence and visual inspection suggest that a bias in the SZE-signal is the more likely explanation for these false positives. In our opinion this population of objects merits further study, because it might systematically bias the X-ray flux–mass or the SZE-signal–mass relation.

## 3.5 Discussion

Here we first summarize the findings from the previous section and then discuss implications. We focus on different aspects, including: (1) internal indications for unresolved systematics in the selection function modelling, (2) the outcome of our validation, (3) the impact of optical incompleteness and the resulting benefits from its modeling, and finally (4) the implications of this work for cosmological studies.

### 3.5.1 X-ray selection function systematics

In section 3.1.3 we discussed potential unresolved redshift trends of the selection function fit. This systematic manifests itself in different places, as discussed in the following.

When sampling the X-ray number counts (c.f. Section 3.4.2) we sample the parameters of the richness–mass scaling relation with priors from the literature to estimate the impact of optical incompleteness of the sample. While the prior on the redshift evolution is  $C_\lambda = 0.73 \pm 0.76$ , the posterior is  $C_\lambda = 0.42 \pm 0.47$ , indicating that the X-ray number counts likelihoods slightly prefers a weaker redshift trend of the richness, effectively making the optical incompleteness stronger at low redshift and weaker at high redshift. This preference may be compensation for the fact that our model seems to predict a too large X-ray selection function at low redshift and a too small X-ray selection function at high redshift. Similarly, when sampling the optical number counts, we rely on priors on the X-ray flux–mass scaling relation to propagate the X-ray selection function to the space of measured richness. Also in this case the prior  $C_X = 0.20 \pm 0.42$  is altered to a posterior  $C_X = -0.21 \pm 0.31$ . Consequently, a weaker redshift trend is preferred by the number counts, possibly as in an attempt to compensate the same residual systematic effect. Lastly, we find that both the X-ray, as well as the optical number counts, pull the prior we placed on  $\Omega_M = 0.276 \pm 0.047$  to a posterior  $\Omega_M = 0.325 \pm 0.026$  from X-ray number count, and  $\Omega_M = 0.318 \pm 0.028$  from optical number counts, respectively. If these shifts result in biases of the cosmological results once direct mass information from weak lensing is available, they will be further investigated.

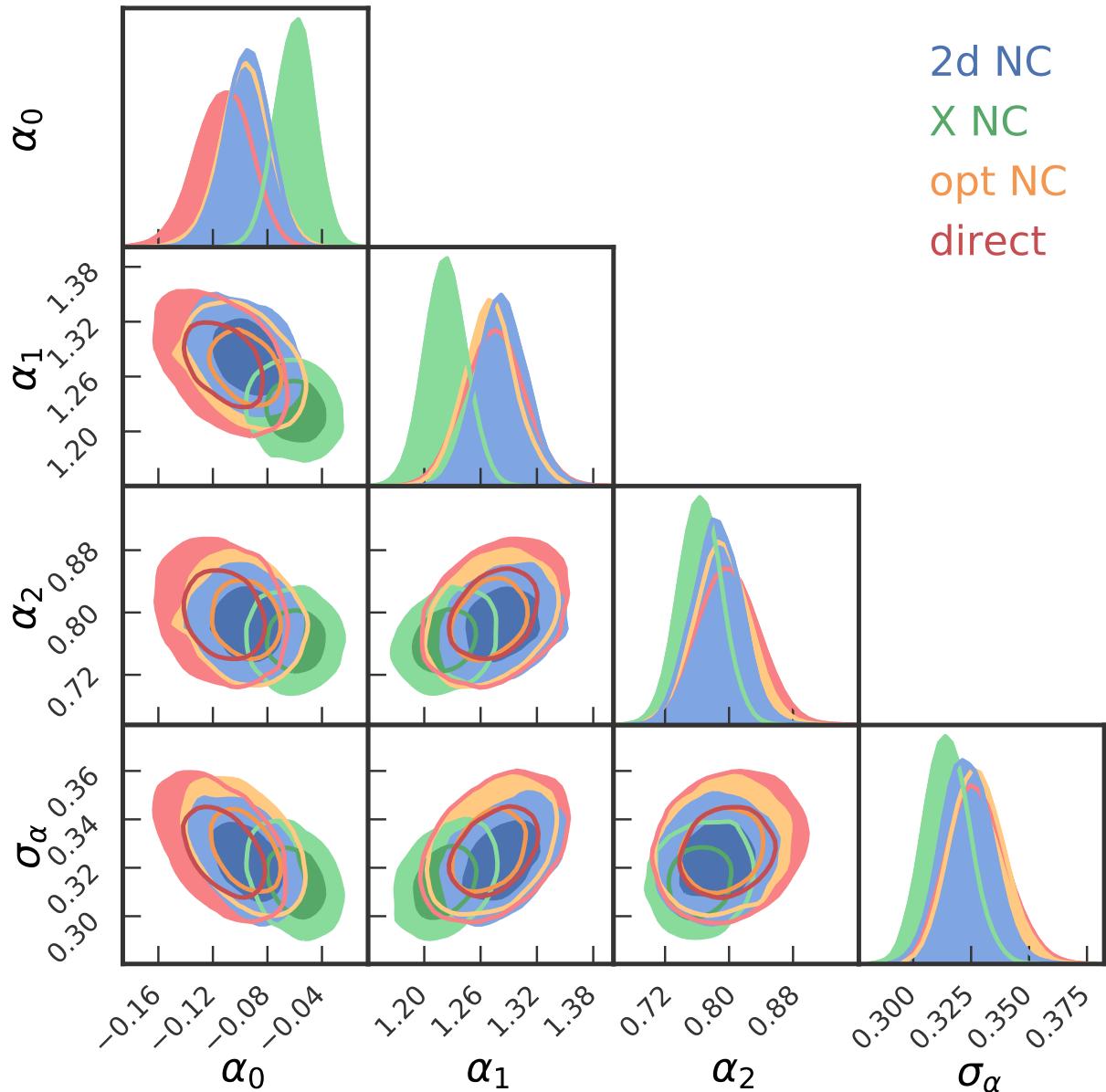


Abbildung 3.16: Marginal posterior contours on the extra nuisance parameters controlling the mapping between X-ray flux and detection significance, and hence the X-ray selection function from the direct fit to the data (red), the sampling of that fit with the 2d number counts (blue), with the X-ray number counts (green), and the optical number counts (orange). Shifts of the contours with respect to the constraints from the data alone are indicative of residual systematics.

As described above in the case of a putative redshift residual, the empirical calibration of the selection function provides an opportunity to uncover unresolved systematics. From this perspective it offers advantages in comparison to selection functions determined from image simulations. For instance, consider in Fig. 3.16 the posterior constraints on the significance–flux scaling parameters resulting from fitting directly to the relevant catalog data by sampling equation 3.19 (red), and with different number counts likelihoods (2d in blue, X-ray in green, and optical in orange). In principle, we expect no extra information from the number counts on the scaling governing the X-ray selection function. Yet, especially the posterior of the X-ray number counts displays shifts with respect to the direct fit. This might hint at unresolved systematic effects in the X-ray number counts. Indeed, we find that the X-ray number counts predict a smaller intrinsic scatter  $\sigma_X$  than both the SPT-SZ cross calibration and the 2d number counts. While at the current stage these putative systematics are smaller than the statistical uncertainties, the empirical methods here already prove to be potent tools for validating the number counts. We plan to include such tests as unblinding conditions for the forthcoming cosmological analysis of this catalog.

### 3.5.2 Outcome of the validation

As outlined in Section 3.4.3, different methods with different sensitivities to the selection function provide statistically consistent masses. This provides strong evidence for the adequacy of the selection functions we constructed in this work. Interestingly, however, non-significant tensions appear on different parameters, mainly in the scaling relation parameters derived from 1d X-ray number counts and 2d number counts. We identify two main scenarios: low intrinsic scatter and steep luminosity–mass trend, preferred by X-ray number counts, and large intrinsic scatter and shallow slope, preferred by 2d number counts. In the following, we will discuss evidence for these two scenarios.

Comparison to the literature does not provide clear guidance on which scenario is more plausible, as can be seen in Fig. 3.17. The low scatter scenario is in very good agreement with the results from Bu19 on XMM luminosities of SPT-SZ selected clusters. On the other side, weak lensing calibrated measurements of the luminosity–mass relation on RASS selected clusters by Mantz et al. (2015) and Mulroy et al. (2019) find shallower mass trends and larger intrinsic scatter in good agreement with our large scatter scenario. In analysing number counts of RASS selected clusters with X-ray mass information, Vikhlinin et al. (2009a) found a mass trend and scatter value consistent with both scenarios.

Further evidence for the amount of intrinsic scatter can be obtained by comparing different measurements of the luminosities. K19 show that there is significant scatter among the luminosities measured by Boller et al. (2016) and those reported by Piffaretti et al. (2011). Namely, a log- normal scatter of  $0.48 \pm 0.05$  for  $0.15 < z < 0.3$  and  $0.40 \pm 0.10$  for  $0.3 < z$ . This is unsettling, considering that the luminosities reported by Piffaretti et al. (2011) are measured on the same ROSAT data as the ones by Boller et al. (2016). Further investigation of the systematics in flux measurement methods is clearly required.

The hypothesis of larger scatter in the X-ray mass scaling is further supported by the constraints on the SPT-SZ incompleteness derived from the different posteriors (see

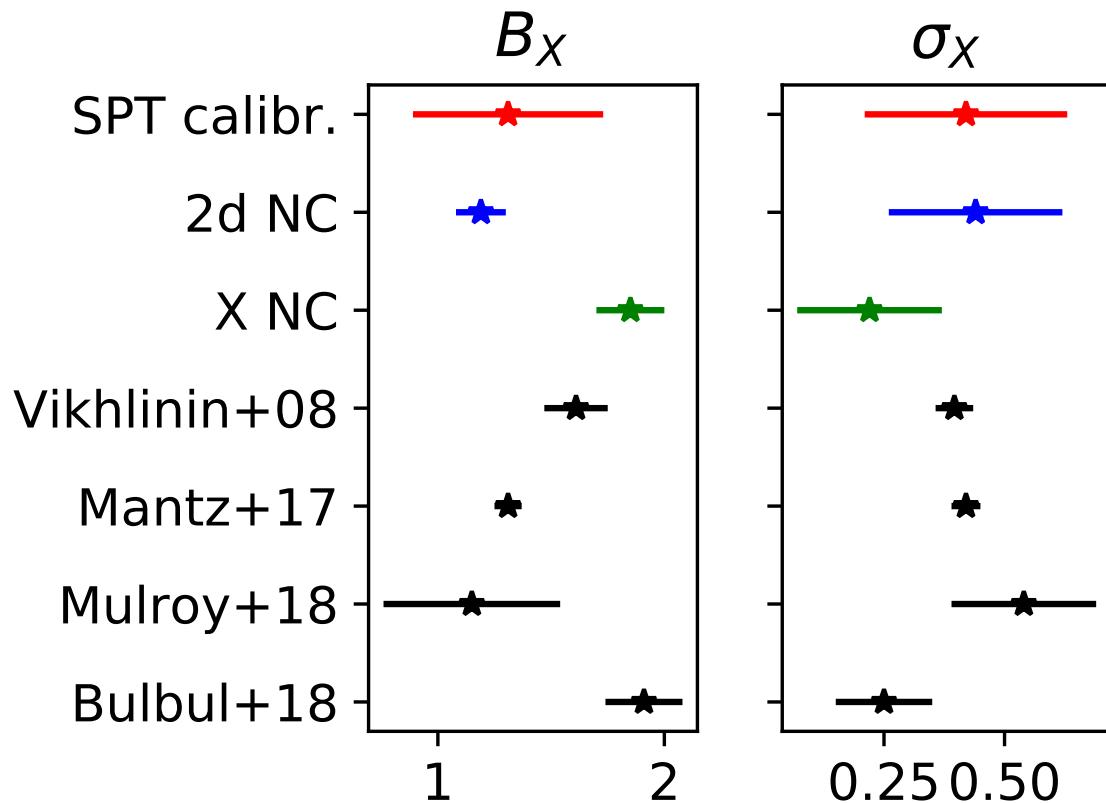


Abbildung 3.17: Compilation of results on the mass trend  $B_X$  and the intrinsic scatter  $\sigma_X$  of the luminosity–mass relation, compared to our results. While all our results lay within the dispersion of the literature results, this dispersion among the results is larger than the reported error bars, indicating that no consensus has yet been reached.

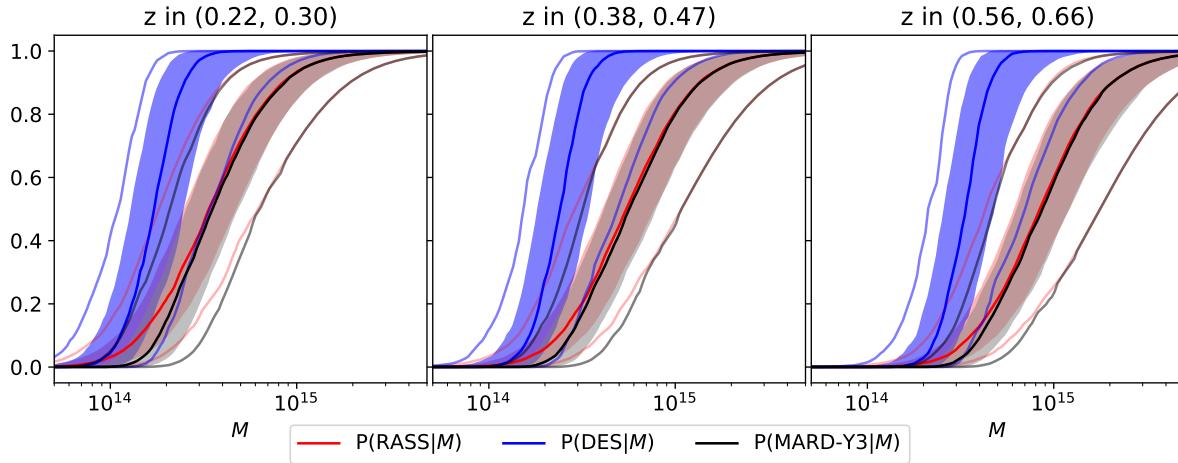


Abbildung 3.18: X-ray (red), optical (blue) and combined (black and grey) selection functions as functions of mass for different redshift bins, plotted with the systematic uncertainties derived from the 2d number counts posterior on the scaling relation parameters. While the combined selection of the sample is clearly dominated by the X-ray selection function at most masses, the optical cleaning introduces some extra incompleteness at low masses, especially at low redshift.

Section 3.4.4). Compared to the literature priors, which prefer small scatter but predict high incompleteness, both the SPT-SZ cross calibration and the 2d number counts predict incompletenesses consistent with zero, mainly due to the larger X-ray intrinsic scatter. On the other side, the mass calibrations of the SZE-mass scaling determined using different, independent methods (Capasso et al., 2019b; Stern et al., 2019; Dietrich et al., 2019; Chiu et al., 2018) match with the masses emerging from a fully self-consistent cosmological analysis of the SPT-SZ cluster sample (Bocquet et al., 2015; de Haan et al., 2016; Bocquet et al., 2019b). In the presence of high incompleteness, this agreement would be coincidental. Larger X-ray scatter is thus made even more plausible, because it predicts low SPT-SZ incompleteness.

In summary, the large scatter scenario is supported by evidence from comparing different luminosity measurements, different literature results and the implications of these scenarios on the inferred SPT-SZ incompleteness. Furthermore, we find that the 2d number count fits introduce less internal tension on the parameters of the significance-flux scaling governing the X-ray selection function.

### 3.5.3 Impact of the optical incompleteness

As shown throughout this work, we model the selection of the MARD-Y3 sample in a two staged approach, which mirrors the operational creation of the catalog: (1) we determine an X-ray selection function based on the fact that the candidate catalog is selected with a X-ray detection significance threshold, and (2) we model the optical cleaning, which is

operationally equivalent to a redshift dependent minimum value for the measured richness. The two result in selection functions in the space of X-ray flux and richness, respectively (c.f. Section 3.1.3).

For ease of representation, we utilise the observable–mass scaling relation to transform these observable selection functions into mass selection functions. This introduces systematic uncertainty through the widths of the posteriors on the scaling relation parameters. The mass selection functions in three redshift bins are shown in Fig. 3.18. As stated already above, the X-ray selection is dominant at most masses. Yet, the optical cleaning introduces an excess incompleteness at the lowest masses, leading to a suppression of the selection probability at those masses.

The fact that the optical selection can not be completely ignored can be appreciated also from Fig. 3.11 and Fig. 3.12. Given that in these plots we show the number of clusters also as a function of measured richness, we can appreciate that the MARD-Y3 sample displays a sharp, redshift dependent cut in measured richness. This is the result of the optical cleaning process. It sets in before the X-ray selection probability is close to zero.

The fact that we can consistently infer the masses when marginalizing over a fiducial cosmology indicates that the two stage selection function modeling is adequately describing the sample. This in turn means that optical cleaning with MCMF can provide clean cluster samples also from highly contaminated candidate samples. At the cost of tracking an extra scaling relation, the richness–mass relation, this has the potential to significantly lowering the limiting mass of ongoing and future surveys with SPT, eROSITA or similar ICM observable based surveys while maintaining a similar contamination level. Given that all selected clusters in such samples would have not only an X-ray or SZE observable, but also a richness, in the context of direct mass calibration the richness–mass relation would be calibrated along side the X-ray or SZE observable, as we demonstrate with our SPT-SZ cross calibration. Furthermore, the possibility to perform number counts not only in the X-ray or SZE observable, but also in richness alone, or even in the combination of multiple observables, provides additional consistency checks that could be used to reveal unappreciated systematics.

### 3.5.4 Implications for cosmological studies

In this work we explore several techniques that allow us to validate the selection function of a cluster survey. However, we would like to caution that in this work we never directly determined the masses of our clusters. This would require either the measurement of the weak lensing signal around our clusters, or the study of the projected phase space distribution of spectroscopically observed cluster members. From a formal perspective, such studies can be treated analogously to our SPT-SZ cross calibration. They will allow us to determine the parameters of the scaling relations to high accuracy, enabling the use of the number counts to study cosmology.

In contrast, our current work assumes the cosmology derived by Bo19 to determine the scaling relation parameters from the number counts of the MARD-Y3 sample. Also the indirect mass information we use in form of the priors on the SZE–mass relation were

derived by Bo19 in the same analysis. So they, too, are contingent upon that analysis. The consistency of their result with our modeling is supported by the fact that we do not find a significant level of SPT-SZ incompleteness.

Our work then demonstrates several techniques that we anticipate will be important for controlling systematics in future X-ray selected cluster samples, especially the sample detected by eROSITA (Predehl et al., 2010; Merloni et al., 2012). First, we have shown that the X-ray selection function can be determined empirically from the selected sample. As such, the simplistic assumptions made in forecast works (e.g. Grandis et al., 2019) can be easily replaced by a more accurate description without introducing much numerical complexity. The empirical determination of the selection function also allows one to check for unresolved systematic effects, as demonstrated in Section 3.1.3. As an addition to the set of systematics tests, such techniques are likely to improve the systematics control within eROSITA cluster cosmological studies.

Our work also highlights the use of secondary mass proxies to inform the number counts experiment. We demonstrate that performing the number counts in optical richness despite the X-ray selection provides a valuable source of mass information. In the presence of a direct mass calibration, that mass information would be provided externally, and optical number counts would provide independent cosmological constraints. This in turn allows one to set up another important consistency check, ensuring a higher level of systematics control. On the same note, we also clearly demonstrate the value of additional mass proxies to put direct constraints on the scatter. Indeed, the analysis of the number counts in X-ray flux and richness space was central to revealing the larger scatter in X-ray observable. Given the planned application of MCMF to eROSITA such multi-observable number counts experiments can be undertaken also in that context.

Furthermore, we present here an expansion of earlier work by S15 on detection probabilities of clusters selected by one survey in another survey. Our formalism allows us to test the selection functions of different surveys and thereby gain precious empirical constraints on the selection function. This method depends on the shape of the mass function for the Eddington bias correction, and on the redshift–distance relation for the X-ray scaling relation. Importantly, however, it is independent of the distribution of clusters in observable and redshift. In turn, these are the major sources of cosmological information in the number counts experiment. Consequently, in the presence of direct mass information to constrain the scaling relation parameters, this technique provides a selection function test that is insensitive to the predicted number of clusters and its redshift evolution. As such this test is ideally suited to validated cluster number counts experiments.

This approach would not only benefit the systematics control in future X-ray and SZE surveys, but also future optical surveys. The selection function in optical surveys remains a source of systematic uncertainty that has been mainly studied through simulations (Co- stanzi et al., 2019). Applying techniques like ours to empirically validate an optical survey cluster selection function offers important advantages and will become more relevant with the upcoming next generation surveys from Euclid and LSST.

## 3.6 Conclusions

We perform a multi-wavelength analysis of the MARD-Y3 sample (K19). This sample was selected by performing an optical follow up of the X-ray selected 2nd ROSAT faint source catalog (Boller et al., 2016) using DES-Y3 data. The optical followup was carried out using MCMF (Klein et al., 2018), which is a tool that includes spatial and colour filters designed to identify optical counterparts of ICM selected cluster candidates and to exclude random superpositions of X-ray and optical systems. The multi-wavelength dataset allows for an extensive set of cross-checks and systematics probes of the MARD-Y3 sample, its selection function and the associated observable–mass scaling relations.

We model the selection function (see Section 3.1.3) of the MARD-Y3 sample as the combination of the X-ray selection function of the candidate sample together with a model of the incompleteness introduced by the optical cleaning of that sample. We then proceed to calibrate the X-ray luminosity–mass and optical richness–mass relation using different sources of mass information to test whether there is tension in the dataset or a flaw in the selection function.

First, we cross-match the MARD-Y3 and the SZE selected SPT-SZ cluster samples, and calibrate the MARD-Y3 scaling relations using the published calibration of the SZE signal-to-noise–mass relation (see Section 3.4.1). Second, assuming priors on the cosmological parameters from the most recent SPT-SZ cluster cosmology analysis (Bocquet et al., 2019b), we calibrate the observable mass scaling relations from the number counts of MARD-Y3 clusters (see Section 3.4.2). In addition to the traditional number counts as a function of X-ray flux and redshift, we also use the number counts as a function of richness and redshift and the number counts as a function of X-ray flux, richness and redshift.

We find that the different flavours of number counts provide scaling relation constraints that are statistically consistent with the constraints from the SPT-SZ calibration performed on the cross-matched sample. This validates the MARD-Y3 selection function, because the SPT-SZ calibration is independent of the MARD-Y3 selection function, while the number count experiments are highly sensitive to it. This leads us to the main conclusion of this work: *optical cleaning with MCMF allows one to create a clean cluster sample with a controllable selection function*. Once direct mass information is available, we will be able to study cosmology using the MARD-Y3 number counts. The fact that the incompleteness (primarily at low masses) introduced by optical cleaning can be modeled using the richness–mass relation implies that much larger, reliable cluster samples extending to higher redshift and lower masses can be constructed from ICM based surveys if appropriately deep optical and NIR data are available.

In these tests we identify some moderate tension between constraints on the luminosity–mass relation from X-ray number counts and 2d (optical+X-ray) number counts: while the former prefers small intrinsic X-ray scatter and a steep mass trend, the latter prefers a shallower mass trend and larger intrinsic scatter. This hints at some unresolved systematic on the X-ray side. As discussed in Section 3.5.2, the high scatter scenario is supported by the scatter among different measurements of luminosity on the same X-ray raw data highlighted in K19, a further indicator of systematics in the flux measurement. Neverthe-

less, the individual masses derived from the different scenarios are consistent within the uncertainties. Because there is no consensus in the literature, this question merits further investigation once direct mass information is available.

In Section 3.4.3 we present the implications for MARD-Y3 masses from different scaling relations that emerge from the tests described above. There is a tendency for these masses to lie below those calculated using externally calibrated relations from the literature (Saro et al., 2015; Bulbul et al., 2019), and the largest tensions occur at low masses.

We also study the MARD-Y3 selection function by comparing the matched and unmatched MARD-Y3 clusters in the SPT-SZ sample and vice-versa. If the selection functions (also that for SPT-SZ in this case) are well understood then the number of matched and unmatched clusters should be fully consistent with the statistical expectations. Simply stated, this test allows us to constrain MARD-Y3 contamination or SPT-SZ incompleteness (the two effects are degenerate in this test). As discussed in Section 3.4.4, in the large scatter luminosity–mass scenario, we find no evidence for either effect, while in the low scatter scenario we find evidence at the 2 sigma level for either contamination or incompleteness. Given that the MARD-Y3 sample contamination is estimated to be 2.5% (K19) and given that the SPT-SZ sample has been used to produce cosmological constraints in good agreement with independent probes (de Haan et al., 2016; Bocquet et al., 2019b), we take this as further evidence supporting the large scatter scenario.

Looking at the probability of a MARD-Y3 confirmation of an SPT-SZ selected cluster we find a subsample of clusters whose SZE properties suggest they should not have been detected in MARD-Y3, but they are. As discussed in Section 3.4.4, if we model this as an outlier fraction in the distribution of scatter about the mass–observable relations (either abnormally high X-ray flux or low SZE signature), we find a preference for an outlier fraction of  $\sim 10\%$  at a significance between 2 and 3 sigma. The locations of these most extreme clusters in the richness–luminosity–SZE signal scatter plots indicate that this sample is most likely caused by low SZE-signal rather than high X-ray flux, but further study is warranted.

From a methodological perspective we demonstrate several new techniques:

1. Optical follow-up allows for three different flavours of number counts. While we demonstrate the potential of multi-observable number counts, the real novelty is that one can perform number counts as a function of optical richness for a predominantly X-ray selected sample in a consistent manner. In a blinded WL calibrated cosmological analysis we would demand that the blinded cosmology from these three likelihoods be consistent.
2. We improve the technique of studying matched and unmatched clusters in two independent samples by including binomial statistics and marginalizing over the systematic uncertainties associated with lack of knowledge of the observable–mass relation parameters. With the use of probability trees, extra probabilities, such as contamination, incompleteness or outlier fractions, can all be constrained in a statistically sound way. This technique does not depend on the amplitude and redshift evolution of the number of objects, reducing its cosmological sensitivity.

3. We present a flexible empirical method to determine the X-ray selection function from the data itself. It does not require any assumptions about cluster morphology. The empirical nature of the constraint also marginalizes over the inherent uncertainty of the selection function by sampling extra nuisance parameters. Shifts in these nuisance parameters when, for example, calibrating the observable–mass relation using different sources of information can serve as a further test of systematic.

The techniques highlighted here have the potential to enable better control of systematic effects in cosmological studies of current and upcoming cluster surveys. They also demonstrate the potential of multi-wavelength analysis of cluster samples not only to inform the selection function modeling of individual surveys, but also to identify interesting cluster populations. This will help exploit the wealth of information provided by deep and wide surveys in X-ray, optical, NIR and millimeter wavelengths.

## 3.7 Appendix

### 3.7.1 X-ray flux error model

As outlined in Section 3.1.2 in some application it is not sufficient to know the measurement uncertainty only for the objects in the catalog, but the measurement uncertainty is also needed for arbitrary values of measured flux  $\hat{f}_X$  and redshift  $z$ . We therefore seek to predict  $\hat{\sigma}_X^2(\hat{f}_X, z, t_{\text{exp}})$  from the measured entries  $\hat{\sigma}_X^{(i)}$ . First we note that the measurement uncertainties in the catalog scale with the exposure time approximately like  $\hat{\sigma}_X^{(i)} \sim t_{\text{exp}}^{-0.5}$ . We thus bin the quantity  $(\hat{\sigma}_X^{(i)})^2 t_{\text{exp}}/400\text{s}$  in fine redshift and measured flux bins, as shown in the upper panel of Fig. 3.19. This is then extrapolated and smoothed to provide a prediction of the measurement uncertainty  $\sigma_{\text{pred}}^2(\hat{f}_X, z)$  at each measured flux  $\hat{f}_X$  and redshift  $z$ , if the exposure time was  $t_{\text{exp}} = 400\text{s}$ , shown in the lower panel of Fig. 3.19. This prediction can then be scaled to the desired exposure time assuming the scaling above, i.e.

$$\hat{\sigma}_X^2(\hat{f}_X, z, t_{\text{exp}}) = \sigma_{\text{pred}}^2(\hat{f}_X, z) \frac{400\text{s}}{t_{\text{exp}}}. \quad (3.39)$$

Applying this prediction the cluster in our catalog and comparing the resulting uncertainties to the actual measurement uncertainty leads to a mean relative error of 5.6%. Furthermore, these residuals display no strong trends with background brightness, neutral hydrogen column density or measured extent. Given the small magnitude, we choose to ignore this source of systematic uncertainty, which could be included at the cost of sampling extra nuisance parameters.

### 3.7.2 Gallery of Multi-wavelength Cluster Images

For each of the clusters we show four different maps: the DES rgb-pseudo color image in the upper left, the RASS count rate in the upper right, the SPT-SZ signal to noise in the

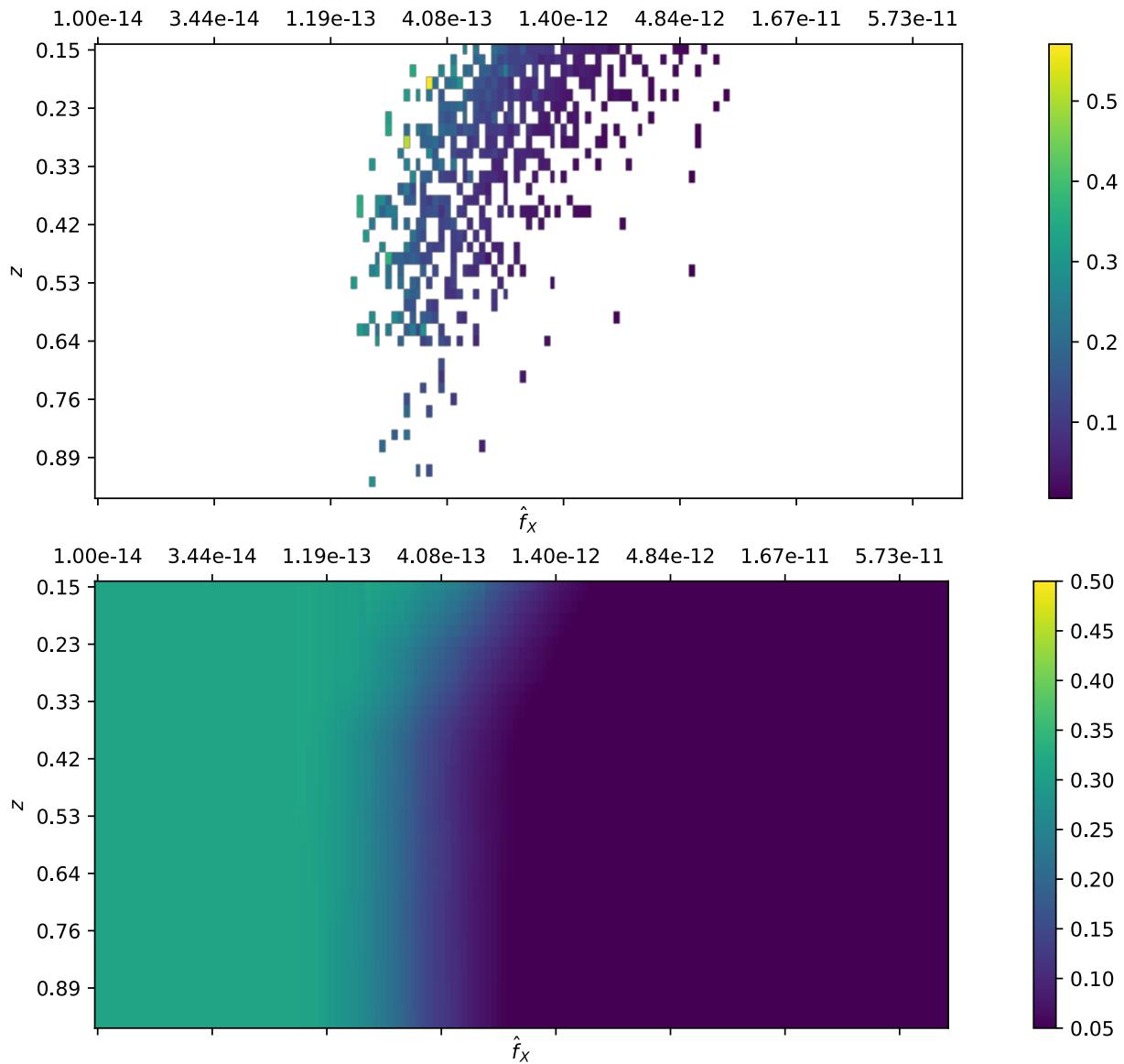
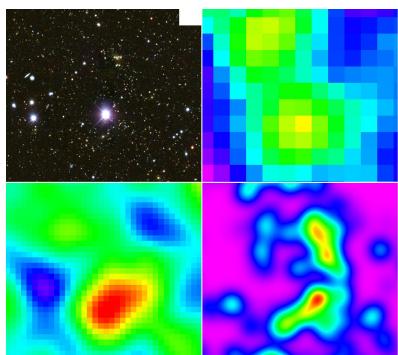
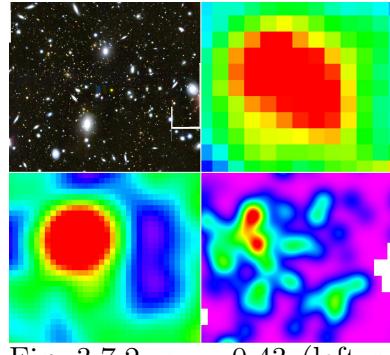
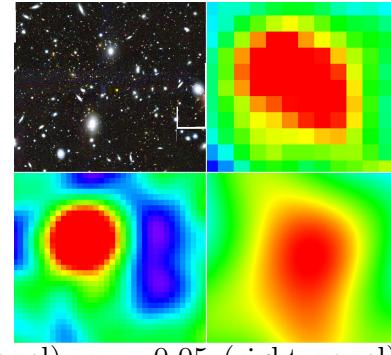
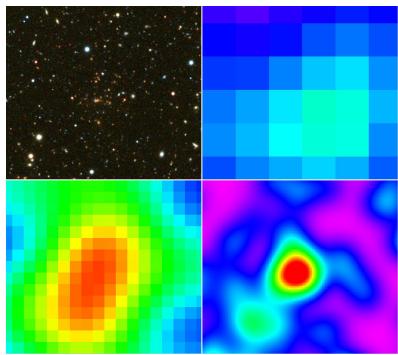
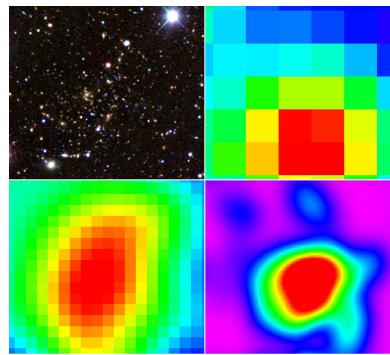
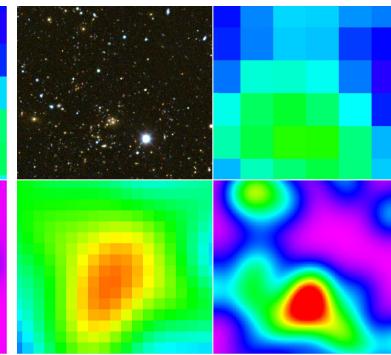
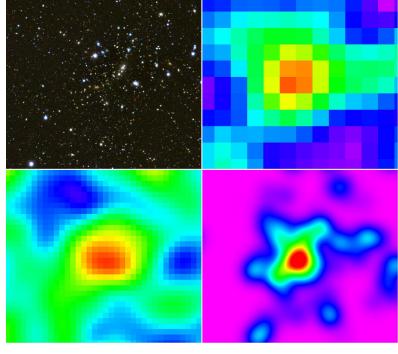
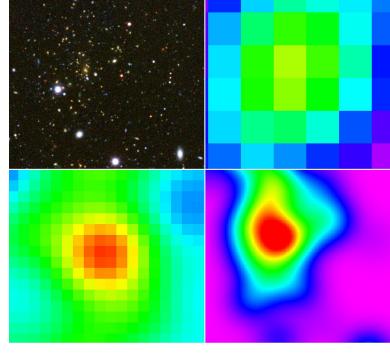
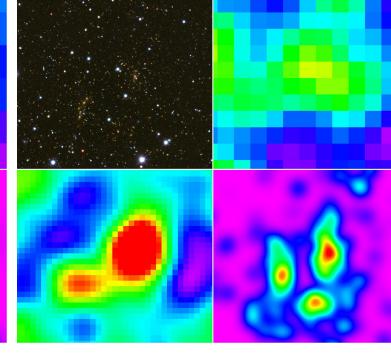
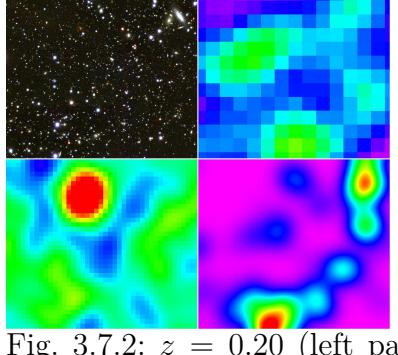
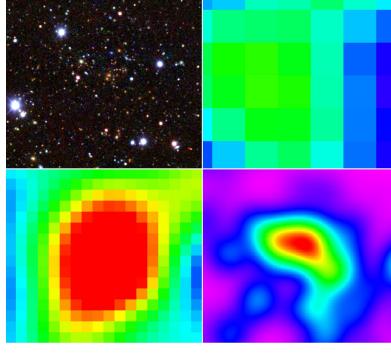
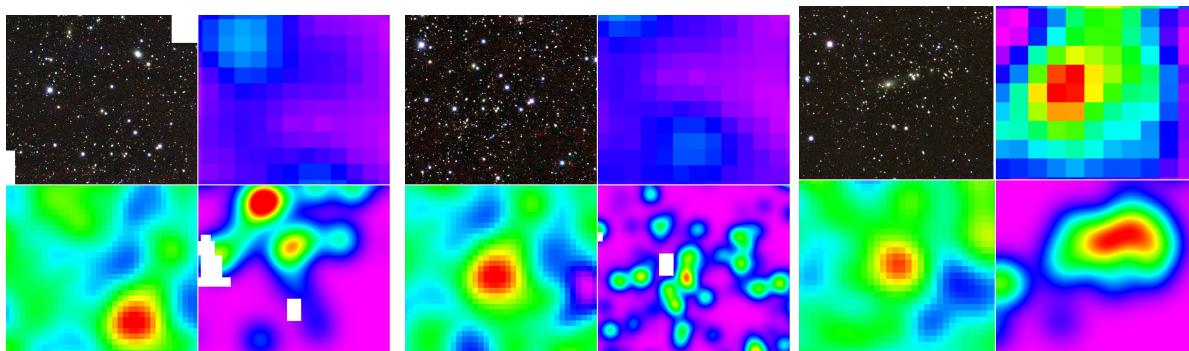
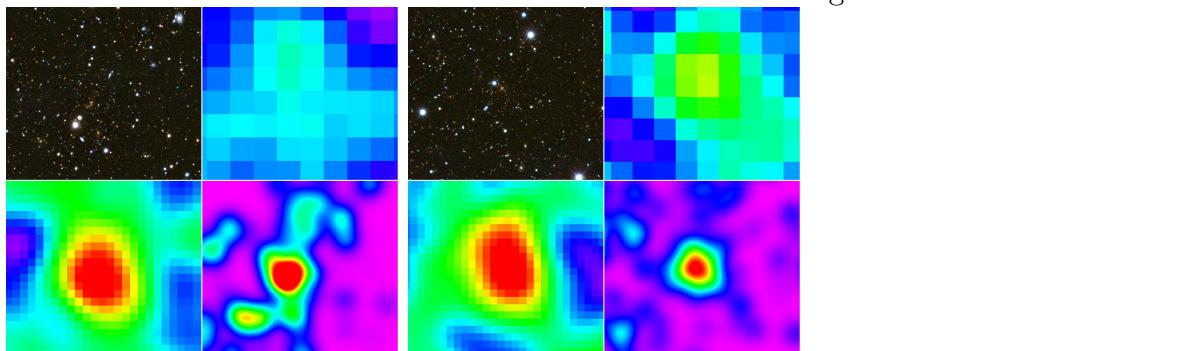


Abbildung 3.19: Construction of the expected measurement uncertainty at the median redshift (lower panel) as a function of redshift (y axis) and measured flux (x axis), from the normalized measurement uncertainties reported in the catalog (upper panel). In the range where we have data, the predicted observational measurement uncertainty nicely extrapolates the trends in the catalog.

smallest filter scale in the lower left and the galaxy density map at the redshift noted in the respective caption in the lower right. We discourage inspection of the DES color image in printed version and recommend to inspect it by zooming into at PDF format of the paper. Some clusters are show twice with galaxy density maps at different redshifts.

Fig. 3.7.2:  $z = 0.403$ Fig. 3.7.2:  $z = 0.43$  (left panel),  $z = 0.05$  (right panel)Fig. 3.7.2:  $z = 0.403$ Fig. 3.7.2:  $z = 0.64$ Fig. 3.7.2:  $z = 0.71$ Fig. 3.7.2:  $z = 0.50$ Fig. 3.7.2:  $z = 0.48$ Fig. 3.7.2:  $z = 0.510$ Fig. 3.7.2:  $z = 0.57$ Fig. 3.7.2:  $z = 0.20$  (left panel),  $z = 0.92$  (right panel)Fig. 3.7.2:  $z = 0.76$

Fig. 3.7.2:  $z = 0.20$  (left),  $z = 0.78$  (right)Fig. 3.7.2:  $z = 0.76$ Fig. 3.7.2:  $z = 0.61$ Fig. 3.7.2:  $z = 0.66$

# Kapitel 4

## Empirical cross calibration of the DES-Y1 Cluster Sample with SPT-SZ selected clusters.

### 4.1 Cluster samples

In this work we investigate the scatter model, the purity and the outlier fraction w.r.t. the mean scaling relation of the optically selected sample based on the DES-Y1 observations. These measurements are performed by cross matching and cross calibration of the optical sample with the selected cluster from the SPT-SZ survey. This limits this analysis to the joint footprint of SPT-SZ and DES-Y1, which is shown in Fig. 4.1 with the relative SPT field depth color coded. It comprises an area of 1463 deg<sup>2</sup>. In the following, we will touch on the main aspects of the two samples relevant to our analysis.

#### 4.1.1 Optically selected samples

We employ the optically selected cluster sample extracted from DES-Y1 data (Drlica-Wagner et al., 2018) with the RM algorithm (Rykoff et al., 2014; Rykoff et al., 2016; Dark Energy Survey Collaboration, in prep.). This sample provides a measured cluster richness  $\hat{\lambda}^i$  with associated measurement uncertainty  $\delta\lambda^i$ , and photometric redshift  $z^i$  for each cluster  $i$ . The photometric redshifts display percent scatter around spectroscopic redshifts. In the joint DES-Y1–SPT-SZ footprint, this sample consists of 1005 objects above  $\hat{\lambda} > 40$ , shown are gray points in Fig. 4.2. For all of these objects we extract the relative SPT field depth  $\gamma_f^i$ .

After our cross matching studies, we extrapolate our results to the cosmological sample. It is defined over the entire DES-Y1 footprint, is selected by  $\hat{\lambda} > 20$ , and comprises 7066 objects. Both samples span the redshift range  $z \in (0.2, 0.7)$ .

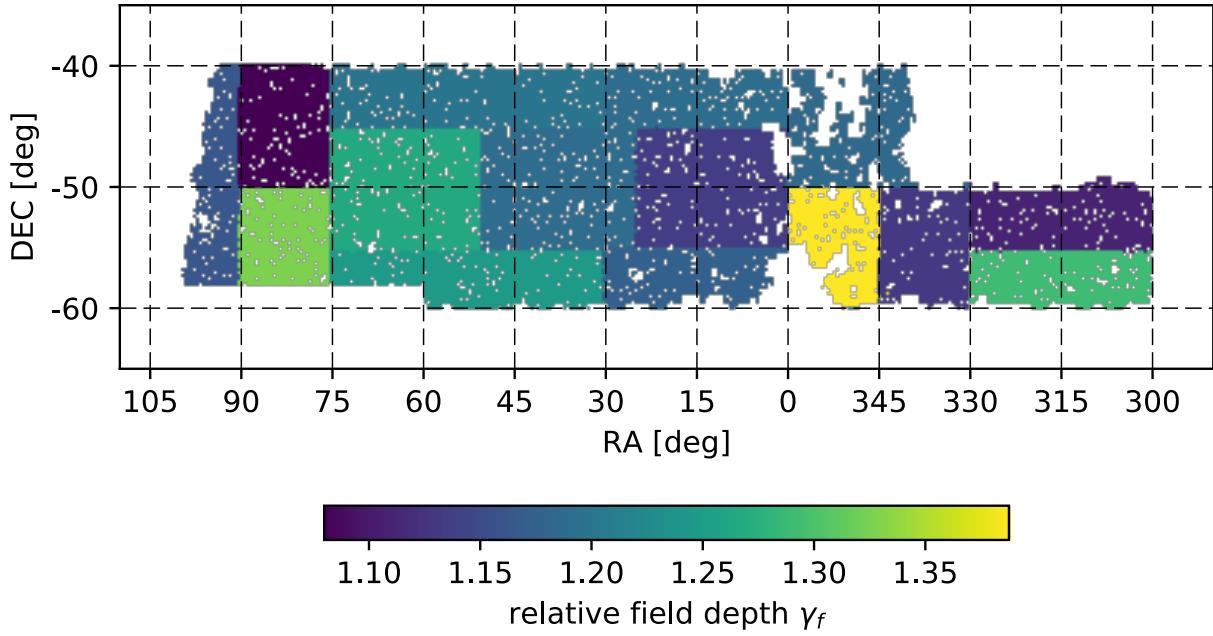


Abbildung 4.1: Joint SPT-SZ DES-Y1 footprint with color coded the relative field depth as reported in de Haan et al. (2016).

#### 4.1.2 SPT matched sample

We match the  $\hat{\lambda} > 40$  RM sample with the SPT-SZ selected sample selected above SZe signal to noise  $\xi > 4$  (Bleem et al., 2015). To reduce the contamination by noise fluctuations, we employ the SPT-SZ catalog that was cleaned by the automated cluster confirmation and redshift measurement tool by MCMF by Klein et al. (in prep.) (see also Klein et al., 2018; Klein et al., 2019, for recent applications), using DES-Y3 photometric data. MCMF allows one to select based of the change of random superposition  $f_{\text{cont}} < 0.1$ , which enables us to produce an uncontaminated SZe selected sample down to signal to noise  $\xi > 4$ . However, at low signal to noise the optical confirmation introduces optical incompleteness. We show in Appendix 4.6.1 that for objects with  $\hat{\lambda} > 40$  this incompleteness is always  $< 2.5\%$ . For the RM- $\hat{\lambda} > 40$  sample the SPT-SZe selection is thus solely given by the SZ signal to noise.

We match the two sample by positional matching within a radius of 1.0 Mpc. We match 207 objects, shown as color circles in Fig. 4.2, where the color represents their SZe signal to noise. 2 SPT clusters are matched to 2 RM  $\hat{\lambda} > 40$  objects each. We confirm that in these cases, both RM objects correspond to a significant detection by MCMF of an optical structure along the line of sight of the SPT cluster. The redshift and the MCMF-richness of these objects match the RM redshifts and richnesses. The SZe signals from these objects might likely have contribution from both cluster along the line of sight. Consequently, it would be biased high for every single object. Given their rarity, we do not treat these objects in a special way.

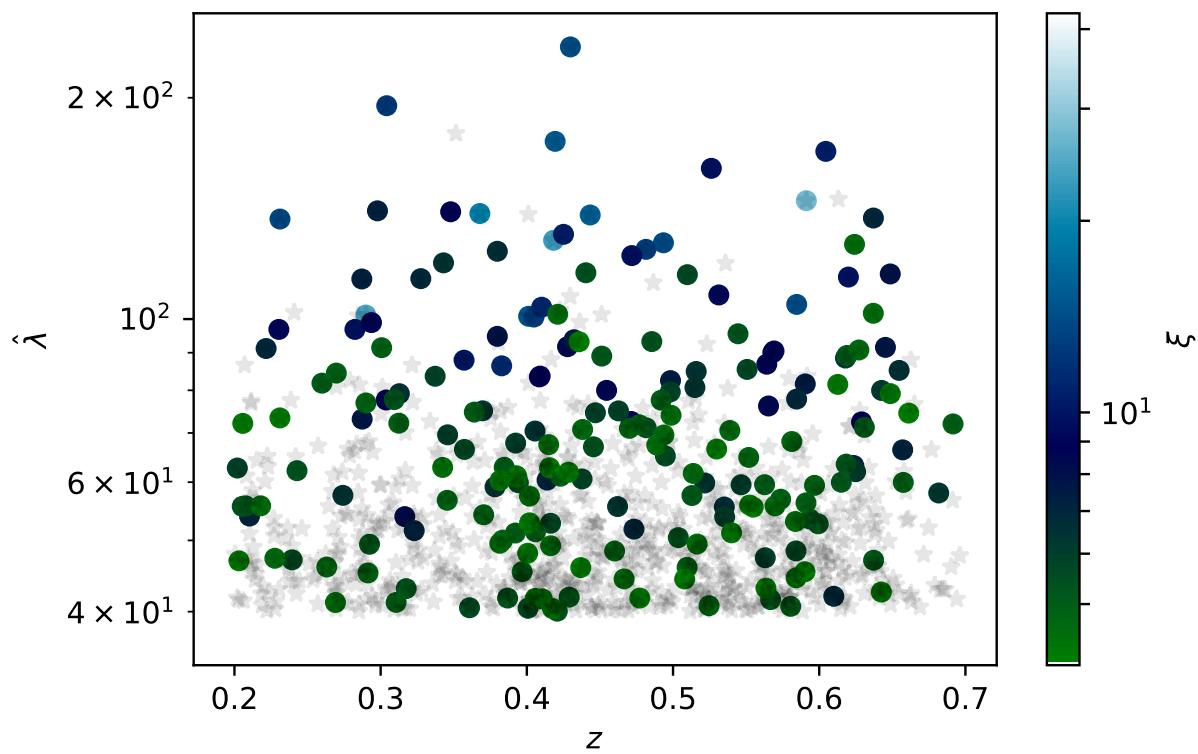


Abbildung 4.2: Distribution in measured richness and redshift of the optically selected RM sample with measured richness  $\hat{\lambda} > 40$  in the joint SPT-SZ-DES-Y1 footprint as gray point. In colored circles the objects matched by SPT-SZ, with their SZ detection significance color coded.

## 4.2 Methods

In this section we will outline our modelling framework. We will describe how we utilize the SPT mass information and its systematic uncertainty to constrain the richness–mass relation and the scatter around that relation. This calibration is done for two distinct error model, which are shortly summarized. We then proceed to present our statistical formalism to constrain the RM purity and the outlier fractions of the scatter model. Furthermore, we discuss how our inferred richness–mass relation, in combination with cosmological priors, allows us to predict the number counts of RM clusters and their stacked weak lensing measurements. Finally, we outline how these predictions, together with measurements of these quantities can be transformed into estimated of the RM purity beyond the sensitivity of the SPT confirmation.

### 4.2.1 General cluster population model

The cluster population model adapted in this work follows closely the model present in (Grandis et al., in prep.[a]), which conceptually build on work by Bocquet et al. (2015). Given a differential number of cluster  $\frac{dN}{dM}\big|_{M,z}$  as a function of mass  $M$  at a redshift  $z$  for a specific survey area, we determine the differential number of cluster in the joint space of intrinsic SZe signal  $\zeta$  and richness  $\lambda$  as

$$\frac{d^2N}{d\zeta d\lambda}\big|_z = \int dM P(\zeta, \lambda|M, z) \frac{dN}{dM}\big|_{M,z} \quad (4.1)$$

where  $P(\zeta, \lambda|M, z)$  describes the scatter around the mean SZe signal and mean richness. A graphic representation of the differential number of objects as a function of the two intrinsic observables is shown in the central panel of Fig. 4.3. For these mean relations we adopt the same powerlaw behaviour as outlined in S15, reading

$$\langle \zeta \rangle = A_{\text{SZ}} \left( \frac{M}{M_0} \right)^{B_{\text{SZ}}} \left( \frac{E(z)}{E(z_0)} \right)^{C_{\text{SZ}}}, \quad (4.2)$$

with  $M_0 = 3.0 \cdot 10^{14} M_\odot h^{-1}$  and  $z_0 = 0.6$ .

We define an analogous relation for the mean intrinsic richness,

$$\langle \lambda \rangle = A_\lambda \left( \frac{M}{M_0} \right)^{B_\lambda} \left( \frac{E(z)}{E(z_0)} \right)^{C_\lambda}. \quad (4.3)$$

Following again S15 the scatter in SZe signal is modeled a log-normal distribution with dispersion  $\sigma_{\text{SZ}}$ , while the scatter in richness also has a Poissonian contribution. We thus have four free parameters for each relation: an amplitude  $A_{\text{SZ}/\lambda}$ , a mass slope  $B_{\text{SZ}/\lambda}$ , redshift evolution  $C_{\text{SZ}/\lambda}$  and an intrinsic scatter  $\sigma_{\text{SZ}/\lambda}$ . We also introduce the correlation coefficient  $\rho$  between the scatter in SZe signal and in richness as a free parameter of our analysis.

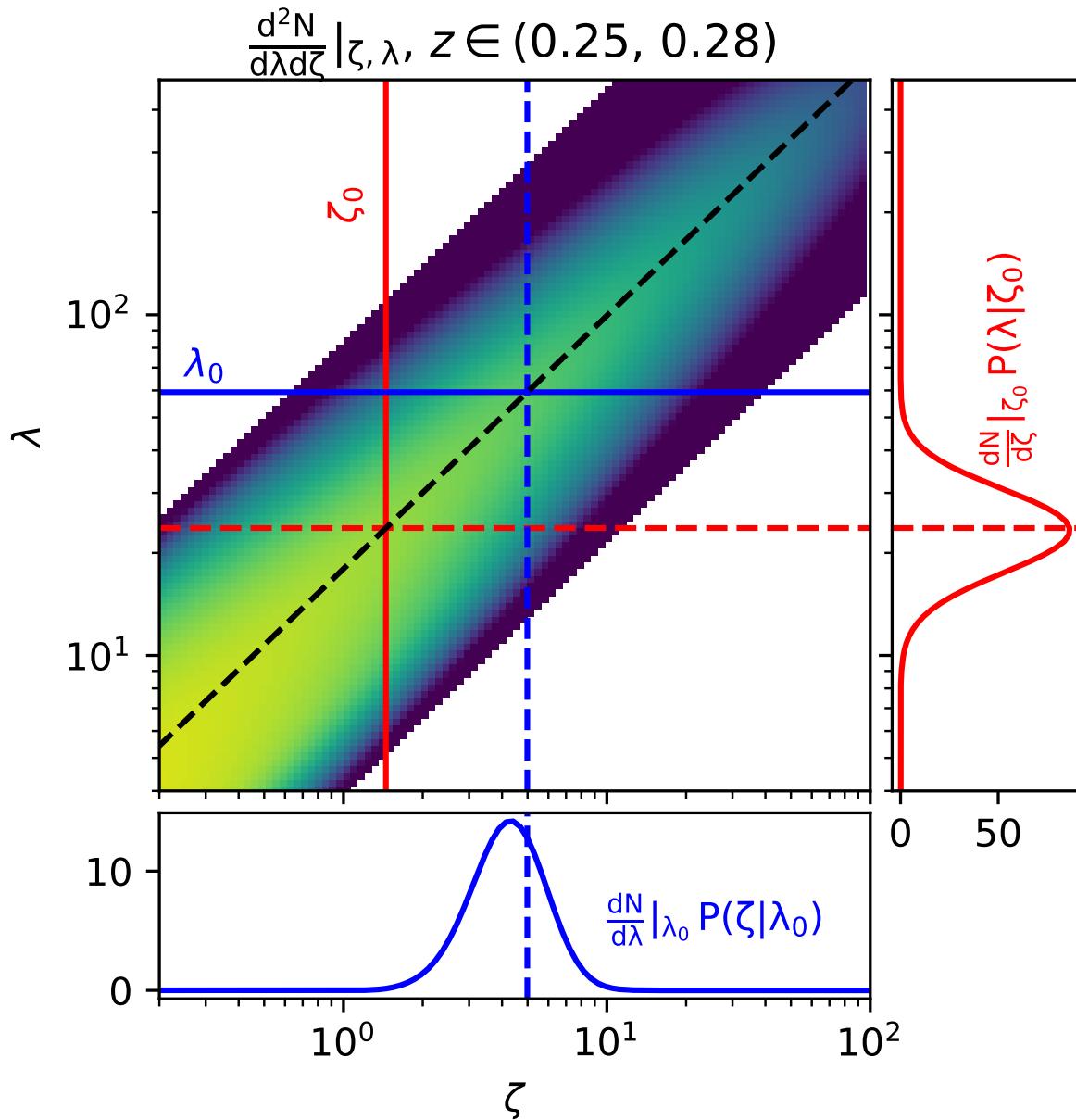


Abbildung 4.3: Graphic representation of the differential number density in space of intrinsic observables SZe signal  $\zeta$  and richness  $\lambda$  (central panel). In red the result of conditioning that distribution on a given value of SZe observable  $\zeta_0$ , with the resulting conditional probability in richness (right insert). The shape of the distribution are the richnesses consistent with  $\zeta_0$ , while the amplitude is the differential number of objects at  $\zeta_0$ . The same argument is repeated when conditioning on a richness  $\lambda_0$  in blue and in the lower insert.

### 4.2.2 Observation error models

Having established our population model as a function of intrinsic observable, we need to account for the observational uncertainty affecting our actual measurements of the SZe signal and the richness.

In the case of the SZe signal, the measured signal  $\xi$  follows the distribution established by Vanderlinde et al. (2010), which reads

$$P(\xi|\zeta, \gamma_f) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \xi - \sqrt{\gamma_f^2 \zeta^2 + 3} \right)^2 \right\}, \quad (4.4)$$

where  $\gamma_f$  is the normalized field depth at the cluster position.

For the observational error on the richness, we follow two prescriptions. The first follows the method used in S15. Together with the measured richness  $\hat{\lambda}^i$ , the RM cluster catalog provides an estimate on the error of the richness  $\delta\lambda^i$  for each entry  $i$ , which is interpreted as a Gaussian standard deviation, yielding

$$P_{\text{S15}}(\hat{\lambda}^i|\lambda, \delta\lambda^i) = \frac{1}{\sqrt{2\pi(\delta\lambda^i)^2}} \exp \left\{ -\frac{1}{2} \left( \frac{\hat{\lambda}^i - \lambda}{\delta\lambda^i} \right)^2 \right\}. \quad (4.5)$$

For applications where the average measurement uncertainty as a function of arbitrary measured richness  $\hat{\lambda}$  is required, we adopt the extrapolation scheme presented in Grandis et al. (in prep.[a]), appendix A to estimate  $\delta\lambda(\hat{\lambda}, z)$  directly from the catalog.

A detailed study of projection effects on simulations by Costanzi et al. (2019, hereafter C19), expanded the prescription above to provide an accurate description of the impact of correlated structures on mapping between intrinsic and measured richness. This effect is summarized by the fitted probability density function which we call  $P_{\text{C19}}(\hat{\lambda}|\lambda, z)$ . For the exact definition of this function see Eq. 15 in C19.

All analysis steps that follow are performed for both models in an attempt to grasp the impact projection effects might have on our inference.

### 4.2.3 SPT cross calibration

All objects of the matched RM-SPT sample have, besides a redshift  $z^i$ , both a measured SZe signal  $\xi^i$  and a measured richness  $\hat{\lambda}^i$ . For each set of scaling relation parameters, we then use the measured SZe signal  $\xi^i$  to predict the distribution of intrinsic richnesses  $\lambda$ , by convolving the joint distribution of intrinsic observables with the measurement uncertainty of the SZe signal, i.e.

$$P(\lambda|\xi^i, \gamma_f^i, z^i) \propto \int d\zeta P(\xi^i|\zeta, \gamma_f^i) \frac{d^2 N}{d\zeta d\lambda} \Big|_{z^i}. \quad (4.6)$$

This equation depends on the scaling relation parameters through the last term, the differential number of objects. As can be seen in Fig. 4.3, for each intrinsic SZe signal to noise

this expression selects a range intrinsic richness, which correspond to the distribution of richesses for clusters with that SZe signal.

The equation above can be propagated into the space of observed richness by folding with the optical error model, i.e.

$$P(\hat{\lambda}|\xi^i, \gamma_f^i, z^i) \propto \int d\lambda P(\hat{\lambda}|\lambda, z)P(\lambda|\xi^i, \gamma_f^i, z^i). \quad (4.7)$$

The proportionality constant is determined by ensuring that the equation above is properly normalized for all possible measured richesses, that is for  $\hat{\lambda} > 40$ . This is achieved by imposing  $\int_{40}^{\infty} d\hat{\lambda} P(\hat{\lambda}|\xi^i, \gamma_f^i, z^i) = 1$ . Note that this normalization cancels the dependence on the absolute number of objects, strongly reducing the cosmological dependence in this analysis.

Evaluation of the properly normalized Eq. 4.7 at the measured richness  $\hat{\lambda}^i$  gives the single cluster likelihood

$$\ln \mathcal{L}^i = \ln P(\hat{\lambda}^i|\xi^i, \gamma_f^i, z^i) \quad (4.8)$$

The total log-likelihood then results from summing the log-likelihoods of the individual clusters. Sampling it with priors on the SZe scaling relation parameters provides constraints on the richness–mass scaling relation parameters.

#### 4.2.4 Constraining outlier fractions and purity

The analysis outlined in the previous section can only be performed on a matched SPT-RM sample, as for any cluster both a measurement of the richness and the SZe signal are required. Considering instead the entire RM sample above  $\hat{\lambda} > 40$ , we can view being matched or not being matched by SPT as a measurement, just in this case of boolean nature and not a number. We can also seek to predict the outcome of this event for each single RM cluster  $i$  based on the observed richness  $\hat{\lambda}^i$ , the redshift  $z^i$ , and the values of the scaling relation parameters. Indeed, given a  $\hat{\lambda}^i$ , we can predict the probable intrinsic SZe signal associated to it by computing

$$P(\zeta|\hat{\lambda}^i, z^i) \propto \int d\lambda P(\hat{\lambda}^i|\lambda, z^i) \frac{d^2 N}{d\zeta d\lambda} \Big|_{z^i}, \quad (4.9)$$

where in this case the proportionality constant is set by  $\int d\zeta P(\zeta|\hat{\lambda}^i, z^i) = 1$ , as no selection of SZe properties was performed. For a graphical representation of this equation, see Fig. 4.3, where we highlight how conditioning on a given richness selects a range of intrinsic SZe signal to noises, based on the joint number density as a function of both observables. Note that in this case too the normalization cancels the dependence on the absolute number of objects, strongly reducing the cosmological dependence in this analysis.

Given the relative field depth of SPT  $\gamma_f^i$  at the RM cluster position, we can compute the probability  $p^i$  that the cluster will have a measured SZe signal  $\xi > 4$  as

$$p^i = \int d\zeta P(\xi > 4|\zeta, \gamma_f^i) P(\zeta|\hat{\lambda}^i, z^i) \quad (4.10)$$

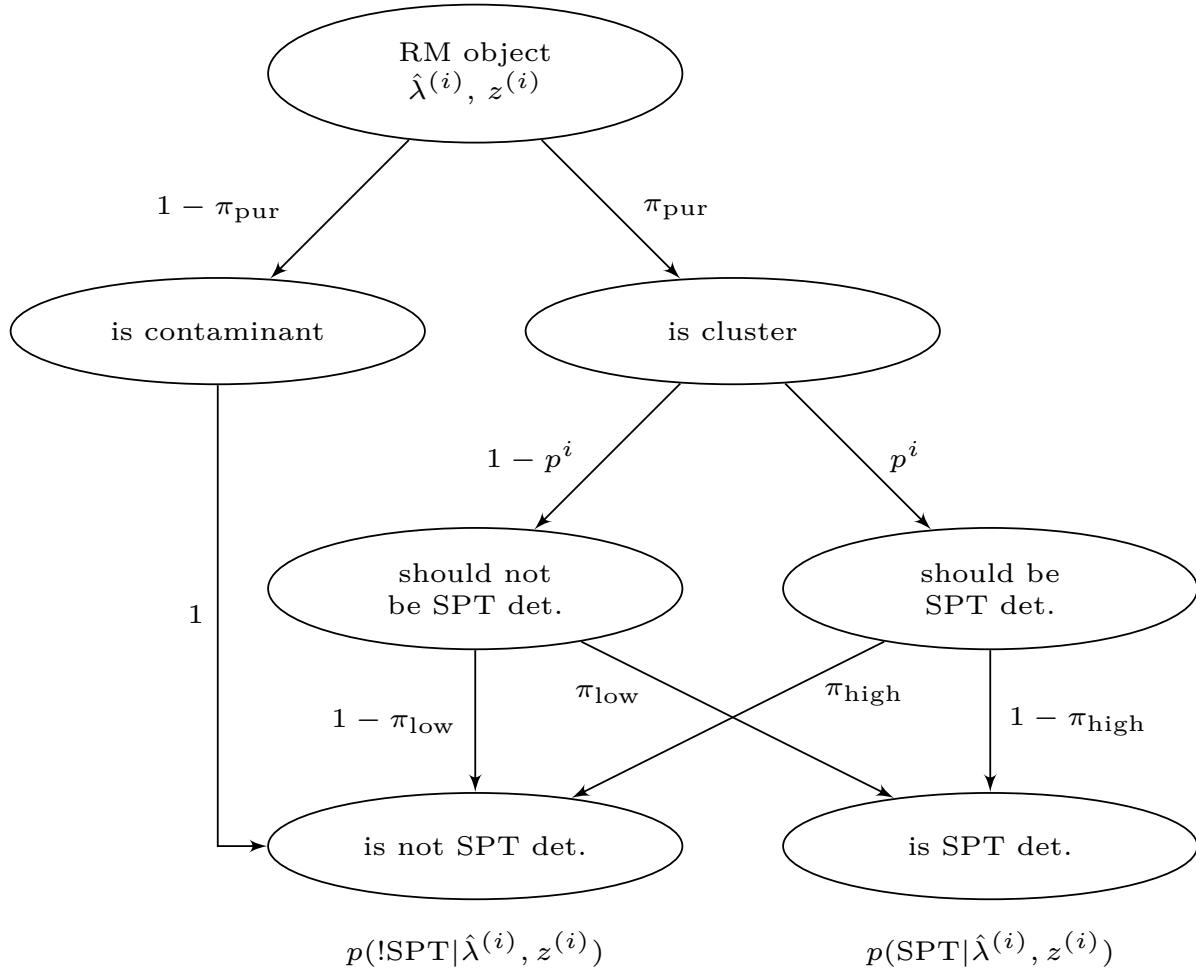


Abbildung 4.4: Probability tree describing the different possibilities for a RM object with measured richness  $\hat{\lambda}^{(i)}$  and redshift  $z^{(i)}$  to be detected by SPT  $p(\text{SPT}|\hat{\lambda}^{(i)}, z^{(i)})$  and not being detected  $p(\text{!SPT}|\hat{\lambda}^{(i)}, z^{(i)})$ . These probabilities depend not only on the raw detection probabilities  $p^i$  as obtained from the observable–mass relations and the selection function, but also on the purity  $\pi_{\text{pur}}$  and the outlier fractions  $\pi_{\text{low/high}}$ .

In this simple example, the likelihood of matching a system is given by the probability of being detected  $p^i$ , as we remind the reader than in a Bayesian context the likelihood is defined as the probability of the data given the model. Similarly, the likelihood of a not matched system is given by the probability of not being detected, that is  $1 - p^i$ .

As in Grandis et al. (in prep.[a]), we expand this formalism in order to investigate different properties of the selection function. To do so we introduce the probability tree shown in Fig. 4.4. For every RM object, we first entertain the possibility that the object is a contaminant. This is modelled by the purity  $\pi_{\text{pur}}$ . Every RM object has only a chance  $\pi_{\text{pur}}$  to be an actual cluster, and, conversely, a chance  $1 - \pi_{\text{pur}}$  of being a contaminant. Anticipating the constraining power of this experiment, we do not investigate possible trends of the purity with richness or redshift at this stage. If the object is a cluster, the detection probability  $p^i$  modulates if it should be detected by SPT or not. We then introduce two further probabilities: first  $\pi_{\text{low}}$  for cluster with a surprisingly low richness leading to an underestimation of the detection probability and thus a detection in SPT despite the cluster should not have been detected. The same physical effect on detection could be achieved by a boosting of the SZe signal, which however is highly unlikely due to its distinct spectral feature. Second, we also allow clusters that we thought should have been detected by SPT to be missed with probability  $\pi_{\text{high}}$ . This can be due to surprisingly high richness, and consequently an overestimated detection probability, or due to a surprisingly low SZe signal due to phenomena such as radio or dust emission in the cluster. Following the probability tree and adding up the weight of all the branches leading to a none detection (represented by the notation ‘!SPT’), we find

$$p(\text{!SPT}|\hat{\lambda}^{(i)}, z^{(i)}) = 1 - \pi_{\text{pur}} + \pi_{\text{pur}}((1 - p^i)(1 - \pi_{\text{low}}) + p^i \pi_{\text{high}}). \quad (4.11)$$

For matched objects, i.e. SPT detected objects, the likelihood is

$$p(\text{SPT}|\hat{\lambda}^{(i)}, z^{(i)}) = \pi_{\text{pur}}((1 - p^i)\pi_{\text{high}} + (1 - p^i)(1 - \pi_{\text{low}})). \quad (4.12)$$

The total likelihood for the RM sample can then be obtained by summing the log-likelihood of the individual clusters based on whether they were detected or not, reading

$$\ln \mathcal{L} = \sum_{i \in \text{matched}} \ln p(\text{SPT}|\hat{\lambda}^{(i)}, z^{(i)}) + \sum_{i \in \text{!matched}} \ln p(\text{!SPT}|\hat{\lambda}^{(i)}, z^{(i)}). \quad (4.13)$$

This likelihood depends on the scaling relation parameters, on the purity, and the outlier fractions.

### 4.2.5 Predictions

After having determined the richness–mass scaling relation parameters, we employ the different posteriors on these parameters to predict several quantities which we compare with data: 1) the fraction of SPT detected RM clusters as a function of redshift and measured richness, 2) the number density of RM cluster in redshift and measured richness bins, and 3) the mean mass of clusters in redshift and richness bins. We also outline how

predictions and measurements of both the number counts of RM clusters and the mean mass can be transformed in estimates of the purity.

### SPT confirmation fraction

When working in a bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$ , we can determine the range of intrinsic richnesses  $\lambda$  that fall into that bin due to observational error as

$$\begin{aligned} P(j|\lambda, z) &= \chi_{z_-^j, z_+^j}(z) dz \int_{\hat{\lambda}_-^j}^{\hat{\lambda}_+^j} d\hat{\lambda} P(\hat{\lambda}|\lambda, z) = \\ &= \chi_{z_-^j, z_+^j}(z) \left[ P(\hat{\lambda} > \hat{\lambda}_-^j | \lambda, z) - P(\hat{\lambda} > \hat{\lambda}_+^j | \lambda, z) \right]. \end{aligned} \quad (4.14)$$

Note that in the limit of infinitely small bins around a measured richness  $\hat{\lambda}^i$  and redshift  $z^i$ , that is  $\hat{\lambda}_{+/-}^j \rightarrow \hat{\lambda}^i$  and  $z_{+/-}^j \rightarrow z^i$ , this equation tends towards the error model for a single clusters given in Eq. 4.5. Also,  $\chi_{z_-^j, z_+^j}(z) = 1$  for  $z_-^j < z < z_+^j$ , and = 0 else.

Predicting the distribution of SZe signals  $\zeta$  in the bin then follows closely the expression for a single cluster given in Eq. 4.9, reading

$$P(\zeta|j) \propto \int d\lambda \int dz P(j|\lambda, z) \frac{d^2 N}{d\zeta d\lambda} \Big|_z, \quad (4.15)$$

where the normalization is given by the condition  $\int d\zeta P(\zeta|j, z) = 1$ . As above, this normalization makes this expression loose much of its dependence on cosmology.

The confirmation fraction  $f(\text{SPT}|j)$  is then obtained by folding the predicted distribution of intrinsic SZe signal to noises with the SPT selection function given by the condition  $\xi > 4$ , that is

$$f(\text{SPT}|j) = \frac{1}{\sum_f A_f} \sum_f A_f P(\xi > 4 | \zeta, \gamma_f) P(\zeta|j), \quad (4.16)$$

where  $A_f$  is the area of the field  $f$  in the joint footprint. The weighted sum over the fields properly accounts for the spatially varying SPT observing conditions.

### Mean mass

The prediction for the mean mass in a bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  can be computed from the predicted distribution of masses

$$P(M|j) \propto \int dz \int d\lambda P(j|\lambda, z) P(\lambda|M, z) \frac{dN}{dM} \Big|_{M,z}, \quad (4.17)$$

where the normalization is given by  $\int dM P(M|j) = 1$ . The mean mass  $\bar{M}(j)$  is then simply

$$\bar{M}(j) = \int dM M P(M|j). \quad (4.18)$$

### Number counts

The number of objects  $N(j)$  in a bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  is given by

$$N(j) = \int dz \int d\lambda P(j|\lambda, z) \int dM P(\lambda|M, z) \frac{dN}{dM} \Big|_{M, z} A(z), \quad (4.19)$$

where  $A(z)$  is the fraction of the survey area in which an object with redshift  $z$  can be detected (Dark Energy Survey Collaboration, in prep.). We compared our prediction to other code available within DES and found that the numerical difference are clearly smaller than the systematic uncertainty on this quantity.

### Purity

We can utilize the prediction for the mean mass in a bin  $j$  defined by  $\hat{\lambda}_-^j < \hat{\lambda} < \hat{\lambda}_+^j$  and  $z_-^j < z < z_+^j$  in combination with the stacked WL measurement of the mean mass  $\bar{M}_{WL}(j)$  in that bin by McClintock et al. (2019a, hereafter McC19) to estimate the purity of the RM  $\hat{\lambda} > 20$ . We assume that a contaminant would contribute  $M_{cont} = 0$ , at least on average. It then follows that

$$\pi_{pur, WL}(j) \bar{M}(j) = M_{WL}(j), \quad (4.20)$$

as the predicted mean mass  $\bar{M}(j)$  is extrapolated from the SPT matched sample. In light of their SZe signal these objects can safely be considered clusters. This estimate of the low mass contamination is conservative, as in reality the WL signal associated to a low mass contaminant is larger or equal to that of a typical large elliptical galaxy. These objects populate the tails of the richness distribution beyond the modelled scatter, which we call low mass contaminants. Due to their mass they contribute less to the dilution of the WL signal than a putative  $M_{cont} = 0$  object. As a consequence, the fraction of low mass contaminants required to bring extrapolated and measured WL mass in agreement is larger than the fraction inferred for putative  $M_{cont} = 0$  objects.

Practically, to account for systematic and statistical uncertainties on the inferred purity, we draw 500 random point from the posterior of the richness–mass scaling relation parameters derived using derived using Eq. 4.20 to predict a sample of  $\bar{M}(j)$ . This propagates the systematic uncertainty of the mass calibration onto our prediction. We then resample the WL measurement  $M_{WL}(j)$  employing the measurement uncertainties reported in McC19 500 times to take account of the correct statistical uncertainties. Finally for each element of the two samples, we estimate of  $\pi_{pur, WL}(j)$ , obtaining a sample thereof. The dispersion in this sample characterizes the systematic and statistical uncertainties on  $\pi_{pur, WL}(j)$ .

An alternative method is to consider the number counts. While we expect  $N(j)$  cluster in each bin, contamination might increase this number (and excess incompleteness reduce it). We therefore assume

$$N(j) = \pi_{pur, NC}(j) N_{RM}(j), \quad (4.21)$$

where  $N_{RM}(j)$  is the expected number of RM objects in the bin  $j$ . We only have access to a Poisson realisation  $\hat{N}_{RM}(j)$ . We create a sample of  $N_{RM}(j)$  by sampling from the

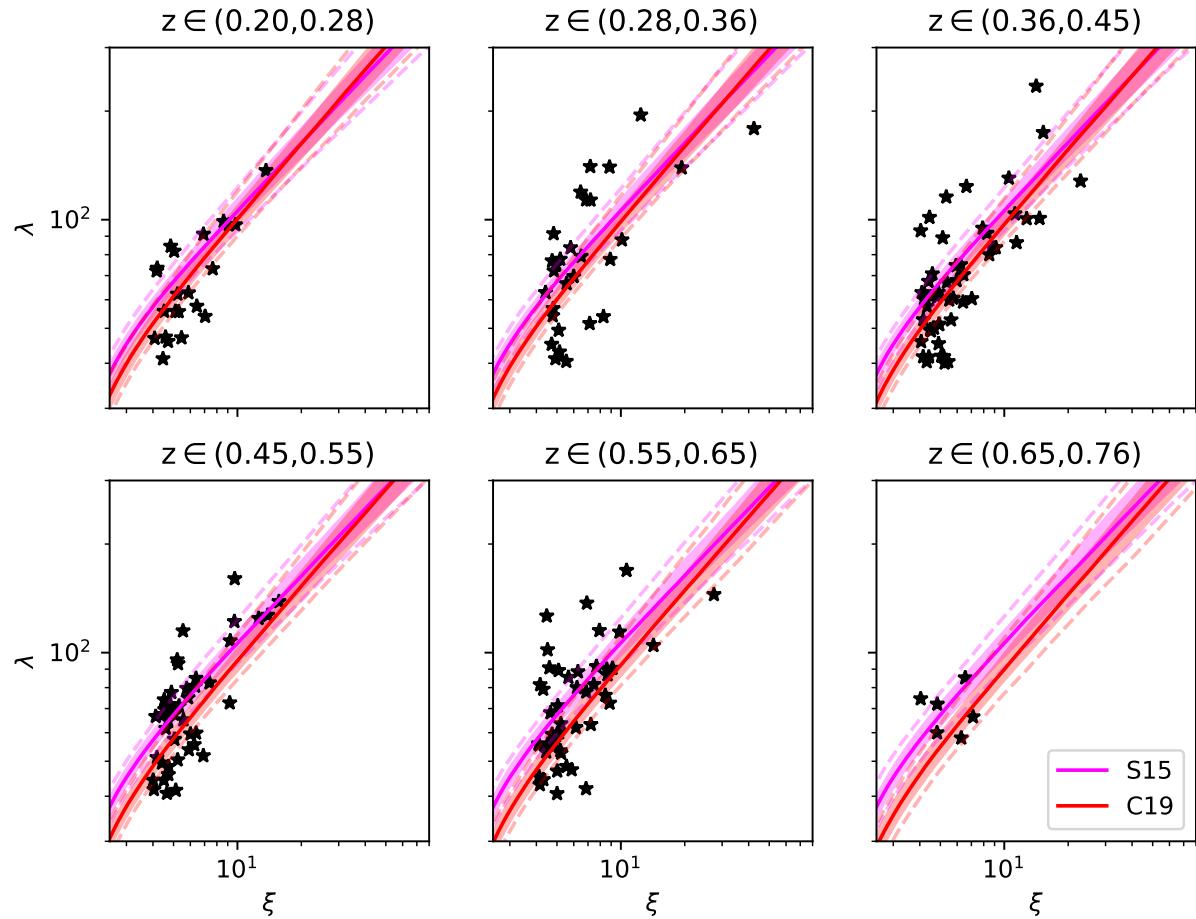


Abbildung 4.5: Scatter plot of the richnesses and SZe signals of the matched sample (black dot with errorbars) in redshift bins. Overplotted the mean richness–SZe signal relation resulting with systematic uncertainties from the fit with the two error models (blue: S15, red: C19). The small deviation in the mean relation reflect the impact of projection effects.

distribution

$$P(n|\hat{N}_{\text{RM}}(j)) \propto e^{-n} n^{N_{\text{RM}}(j)}, \quad (4.22)$$

which, if interpreted as a distribution in  $\hat{N}_{\text{RM}}(j)$  would be the aforementioned Poissonian, but for given  $\hat{N}_{\text{RM}}(j)$  is a distribution in  $N_{\text{RM}}(j)$ . This provides as sample of  $N_{\text{RM}}(j)$ . We then use our posterior on the scaling relation parameters to create a sample of  $N(j)$ . We also use  $\Omega_M = 0.276 \pm 0.047$  and  $\sigma_8 = 0.781 \pm 0.037$  from Bocquet et al. (2019a) to properly account for uncertainties in cosmology when sampling  $N(j)$ . For pairs of entries from the  $N_{\text{RM}}(j)$  sample and the  $N(j)$  sample we estimate  $\pi_{\text{pur,NC}}(j)$ , thus creating a sample thereof which account for systematic uncertainties on the richness–mass relation and on cosmology as well as the statistical uncertainties of the number counts. Note that under these assumptions, the inferred purity can also be larger than 1, indicating excess incompleteness instead.

#### 4.2.6 Priors

We use priors on the SZe signal–mass scaling relation parameters derived from Bocquet et al. (2019a). In that work, constraints on the SZe signal–mass scaling relation parameters are derived by jointly fitting the number counts of SZe selected clusters, and the mass calibration by means of pointed weak lensing follow up measurements. The resulting constraints in the baseline  $\nu$ - $\Lambda$ CDM model read  $A_{\text{SZ}} = 5.25 \pm 1.01$ ,  $B_{\text{SZ}} = 1.53 \pm 0.10$ ,  $C_{\text{SZ}} = 0.47 \pm 0.41$ , and  $\sigma_{\text{SZ}} = 0.16 \pm 0.08$ . These priors encode the systematic mass uncertainty on the SZe derived masses. We like to stress here that these masses do not assume hydrostatic equilibrium, but are instead empirically calibrated using number counts and weak lensing information.

### 4.3 Results

In this section we present our main results, starting with the calibration of the richness–mass relation based on the SPT cross matching. We then proceed to constrain the RM purity and the outlier fractions, subsequently showing that the SPT confirmation fraction of RM cluster matched with our prediction based on the cross calibration. We finally seek to extent the measurement of the RM purity to lower richnesses, by comparing measurement and prediction for the number counts of optically selected clusters and their stacked WL signal.

#### 4.3.1 Cross Calibration

We shall focus first on the measurement of the richness–mass scaling relation parameters, the purity and the outlier fraction from our analysis of the cross matched sample.

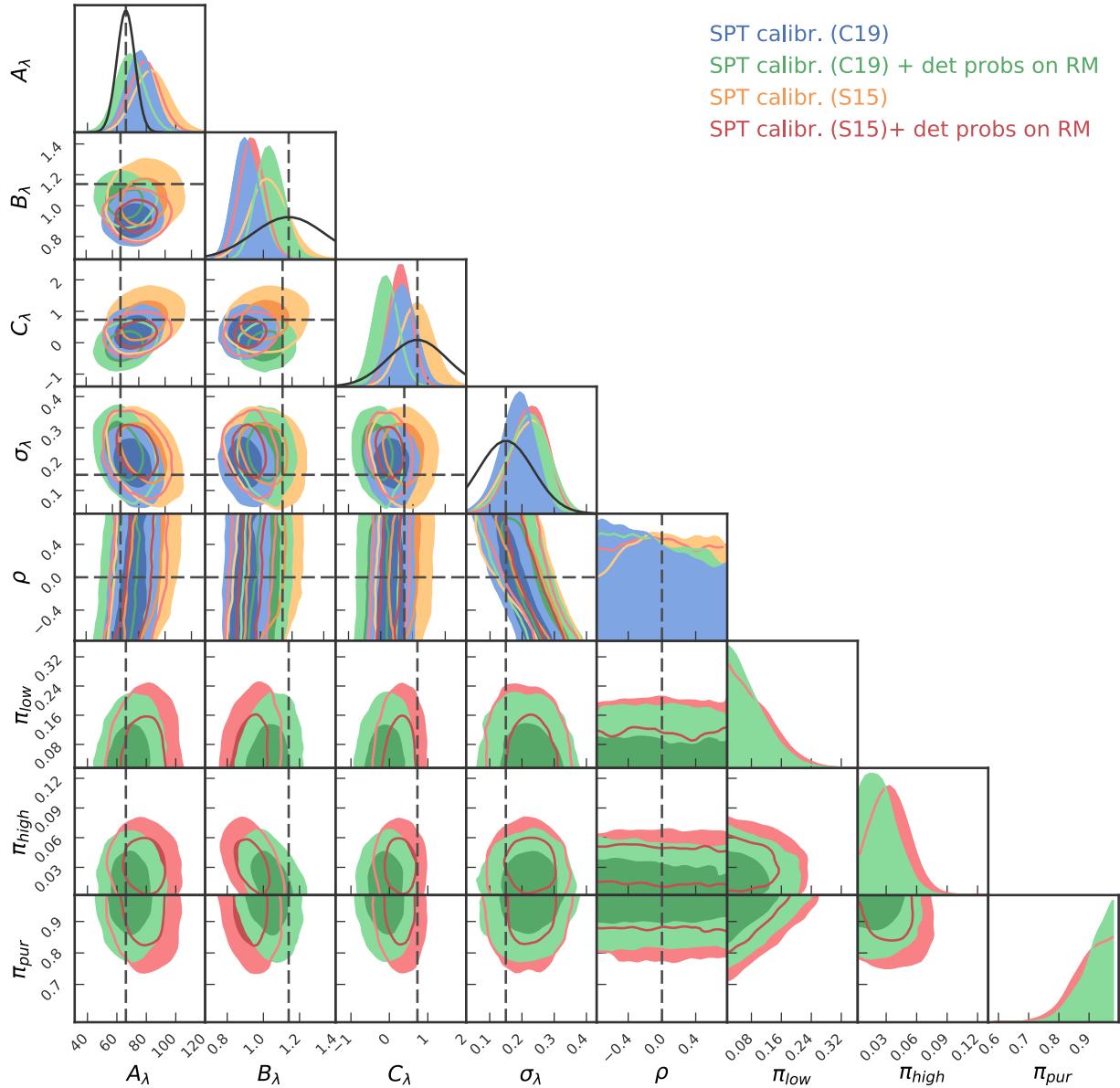


Abbildung 4.6: Posteriors on the parameters of the richness–mass scaling relation, the scatter around that relation and the correlation coefficient between the intrinsic scatter in SZe signal and in richness. Furthermore, where applicable, also the purities and the outlier fractions posterior are shown. We show the results for different combinations of likelihoods and optical error models: blue and yellow the C19 model, and the S15 model, respectively, when sampling just the cross calibration likelihood, while in green and red the C19 model, and the S15 model, respectively, when also considering the detection probability likelihood. In black the richness mass relation obtained by S15 from a SPT-DES matched sample of 25 clusters.

Tabelle 4.1: Mean and standard deviation estimated from the one dimensional marginal posterior plots for the parameters the richness scaling relation, the purity and the outlier fractions.

	$A_\lambda$	$B_\lambda$	$C_\lambda$	$\sigma_\lambda$	$\rho$	$\pi_{\text{pur}}$	$\pi_{\text{low}}$	$\pi_{\text{high}}$
SPT calibr. (C19)	$75.7 \pm 8.8$	$0.91 \pm 0.07$	$0.34 \pm 0.34$	$0.18 \pm 0.05$	–	–	–	–
+ det. prob.	$69.9 \pm 9.2$	$1.04 \pm 0.07$	$-0.09 \pm 0.31$	$0.22 \pm 0.06$	–	$0.80$	$0.20$	$0.022 \pm 0.016$
SPT calibr. (S15)	$83.7 \pm 11.5$	$1.03 \pm 0.10$	$0.71 \pm 0.41$	$0.22 \pm 0.06$	–	–	–	–
+ det. prob.	$78.6 \pm 10.1$	$0.94 \pm 0.07$	$0.28 \pm 0.28$	$0.22 \pm 0.06$	–	$0.76$	$0.22$	$0.033 \pm 0.018$

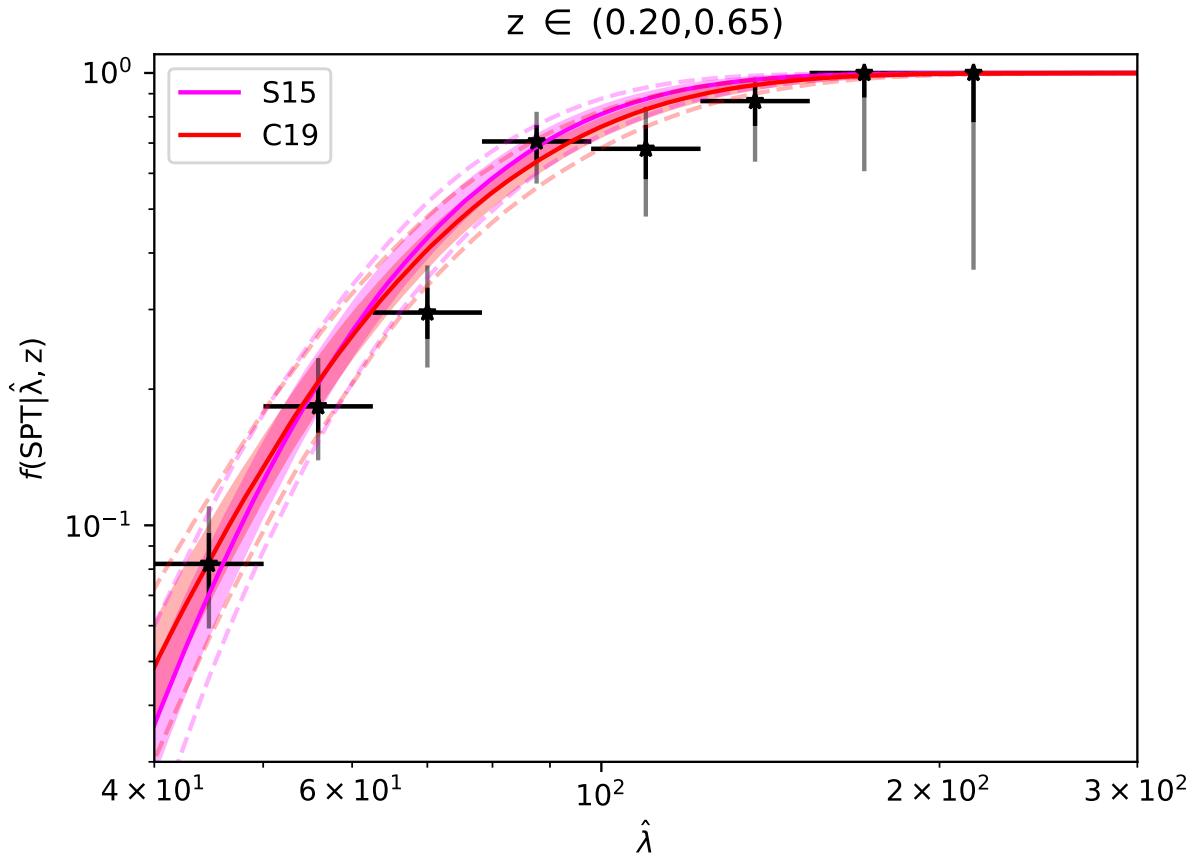


Abbildung 4.7: SPT confirmation fraction as a function of measured richness, predicted from the posteriors on the scaling relation parameters for the S15 error model (blue) and when considering projection effects (red), overlaid with the measurement as black points. This comparison reveals no indication for significant contamination or low purity above measured richness  $\hat{\lambda} > 40$ , consistent with our statistical analysis.

### Scaling relation

As outlined in section 4.2.3 the measured richnesses and measured SZe signals for clusters in the cross matched RM-SPT sample can be used to transfer the calibration of the SZe signal–mass relation given by the priors discussed in section 4.2.6 onto the richness–mass relation. To this end, the sum of the likelihoods given in equation 4.8 are sampled with the aforementioned priors. The scatter plot of the matched sample in redshift bins is shown in figure 4.5 with black points.

The resulting posteriors on the parameters of the richness–mass scaling relation are summarized in terms of their means and standard deviations in Table 4.1 and shown with their marginal distributions in Fig. 4.6. In blue the constraints from adopting the C19 optical error model and sampling the cross calibration likelihood, while in yellow the posteriors when using the S15 optical error model. The constraints are in very good agreement, highlighting that at the level of statistical constraining power of the cross matched sample the two error models are not distinguishable. The minor changes induced by the projection effects can be seen on the mean relation over plotting the data scatter plot in figure 4.5. We also plot in black previous results by S15 on DES-SV.

### Purity and Outlier Fraction

As discussed at length in section 4.2.4 the probability of detecting a RM cluster in SPT is very sensitive on the respective scaling relation parameters of the two selection observables. Consequently, one needs to marginalize over a reasonable range of scaling relation parameters when inferring the purity of the RM sample and the outlier fractions with the likelihood given in equation 4.13. This is accomplished by sampling the likelihood of detection probabilities simultaneously with the cross calibration likelihood (equation 4.8). This ensures proper accounting for the systematic uncertainties on the scaling relation when inferring the purity and outlier fraction. The resulting posteriors, depending on the assumed optical error model, are shown as green and red contours in Fig. 4.6 for the C19 and the S15 model, respectively, and summarized in Table 4.1.

The posteriors on the scaling relation parameters are generally in good agreement with the results without the detection probability likelihood. We find an upper limit  $\pi_{\text{low}} < 0.20$  and  $\pi_{\text{low}} < 0.22$  (at 97.5%) on the low outlier fraction for the C19 model and the S15 model, respectively. We also find a lower limit for the purity of  $\pi_{\text{pur}} > 0.80$  and  $\pi_{\text{pur}} > 0.76$  (at 2.5%) for the C19 model and the S15 model, respectively. For the high outlier fraction, we find  $\pi_{\text{high}} = 0.022 \pm 0.016$  and  $\pi_{\text{high}} = 0.033 \pm 0.018$  for the two error models. Neither of the two measurements is statistically significant. Yet the projection effect model shows consistently larger purity and smaller outlier fraction, hinting at a more adequate description of the scatter.

Interestingly, the detection probability likelihood slightly alters the posteriors on the scaling relation parameters, with the largest impact on the slope and the redshift evolution. This is not surprising when considering that not detecting a RM object in SPT is equivalent to the measurement  $\xi < 4$ , which given priors on the SZe signal–mass relation carries some

mass information, at least in the form of an upper limit. This information is however quite weak, as can be seen by directly comparing the improvement in measurement uncertainties. Furthermore, it is consistent with the information recovered from the matched sample alone, as the shifts in mean values also do not exceed  $1\sigma$ .

The measurement of purity consistent with unity is confirmed when comparing the SPT confirmation fraction of RM objects as a function of richness to the measured fraction, as done in Fig. 4.7. Following equation 4.16, we predict the SPT confirmation fraction of the RM ( $\hat{\lambda} > 40$ ) sample as a function of measured richness. The shaded region reflects the  $1\sigma$  systematic uncertainty as propagated from the posterior samples. Note also here that the difference between the predictions for the two error models is small. As discussed by C19 the values of the scaling relation and scatter around it absorb a majority of the differences between the two error models. We over plot the predicted fraction with the measured fraction in richness bins and associated statistical uncertainties. Within the statistic and systematic uncertainties the prediction and the measurement are in good agreement, confirming the high purity of the RM ( $\hat{\lambda} > 40$ ) sample.

### 4.3.2 Comparison to $\hat{\lambda} > 20$

In the previous section we determined the systematic uncertainty on the richness–mass relation in the regime of measured richness  $\gtrsim 40$ , that is at the high mass end. In this section we extrapolate the relation determined at high masses to lower masses. We first compare our prediction for the mean mass in richness bins to a measurement thereof from stacked weak lensing by McC19. Second we predict the number counts of RM clusters in redshift and richness bins to our prediction. Finally, both these comparison are turned into estimates of purity as a function of redshift and richness, extending to  $\hat{\lambda} > 20$ .

#### Stacked Weak Lensing

The mean mass in redshift and richness bin for our scaling relation parameters constraints are derived following equation 4.18 for the richness bins with edges (20, 30, 45, 60, 300) and redshift bins with edges (0.20, 0.35, 0.50, 0.60) as in McC19. The resulting masses are shown in Fig. 4.8 with the systematic uncertainties as blue and red boxes for the S15 and the C19 error model, respectively. Over plotted are also the measurements by McC19 with statistical uncertainties. The resulting tension  $T(j)$  in the bin  $j$  between the measurements  $\hat{M}_{\text{WL}}(j) \pm \hat{\sigma}$  and the prediction  $\bar{M}(j) \pm \sigma_{\text{sys}}$  is estimated as

$$T = \frac{\hat{M}_{\text{WL}}(j) - \bar{M}(j)}{\sqrt{\hat{\sigma}^2 + \sigma_{\text{sys}}^2}}. \quad (4.23)$$

We report the resulting tension as numbers in the plot. While at high mass the agreement is good, it worsens when moving to lower richness and higher redshift. In the lowest richness, highest redshift bin it clearly exceeds  $2\sigma$  for both optical models.

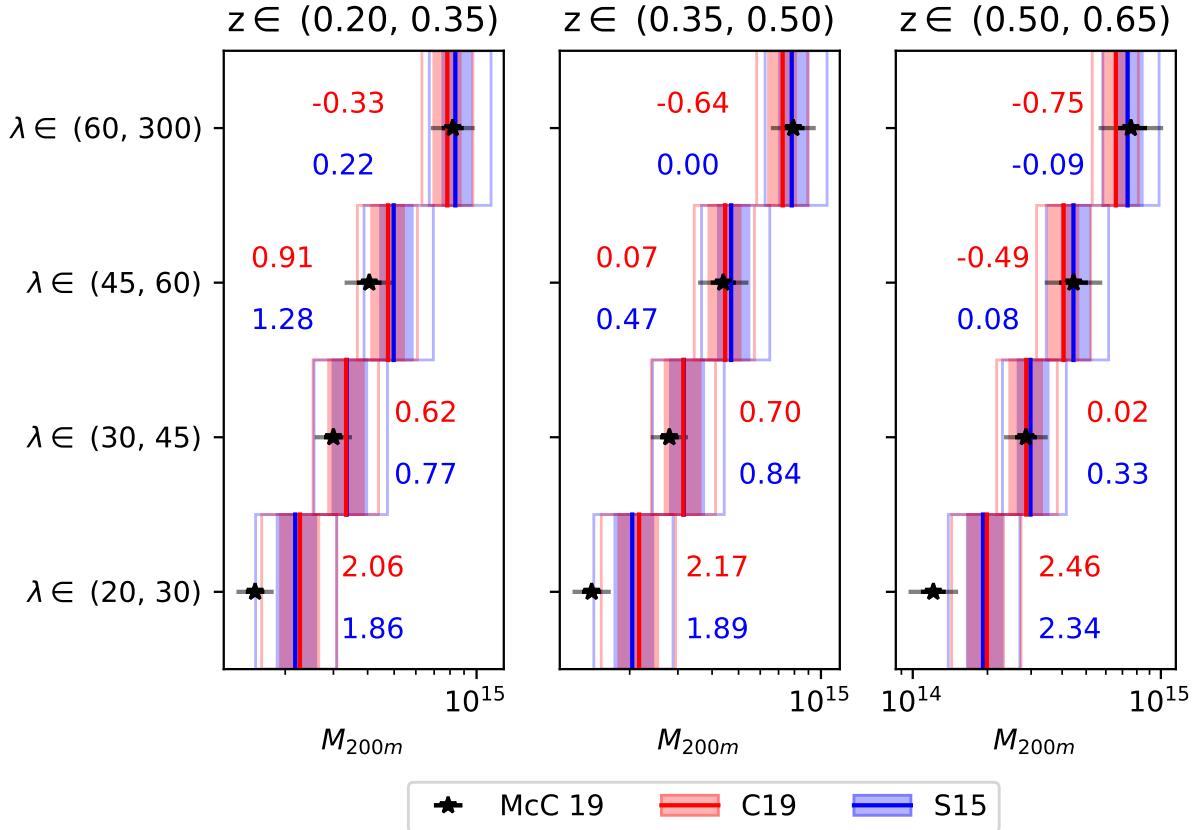


Abbildung 4.8: Mean mass in redshift and richness bins predicted from our posteriors for the S15 error model (blue), and the C19 error model (red), where the filled region represents  $1\sigma$  and the empty one  $2\sigma$ . Over plotted the mean masses reported by McC19 from stacked weak lensing as black points with errorbars. We also report the levels of tension between our prediction and the measurement as numbers next to the respective predictions (red for C19, blue for S15). In the lowest richness bins the tension exceeds  $2\sigma$  for the C19 model, while for the S15 model it does so in the highest redshift, lowest richness bin.

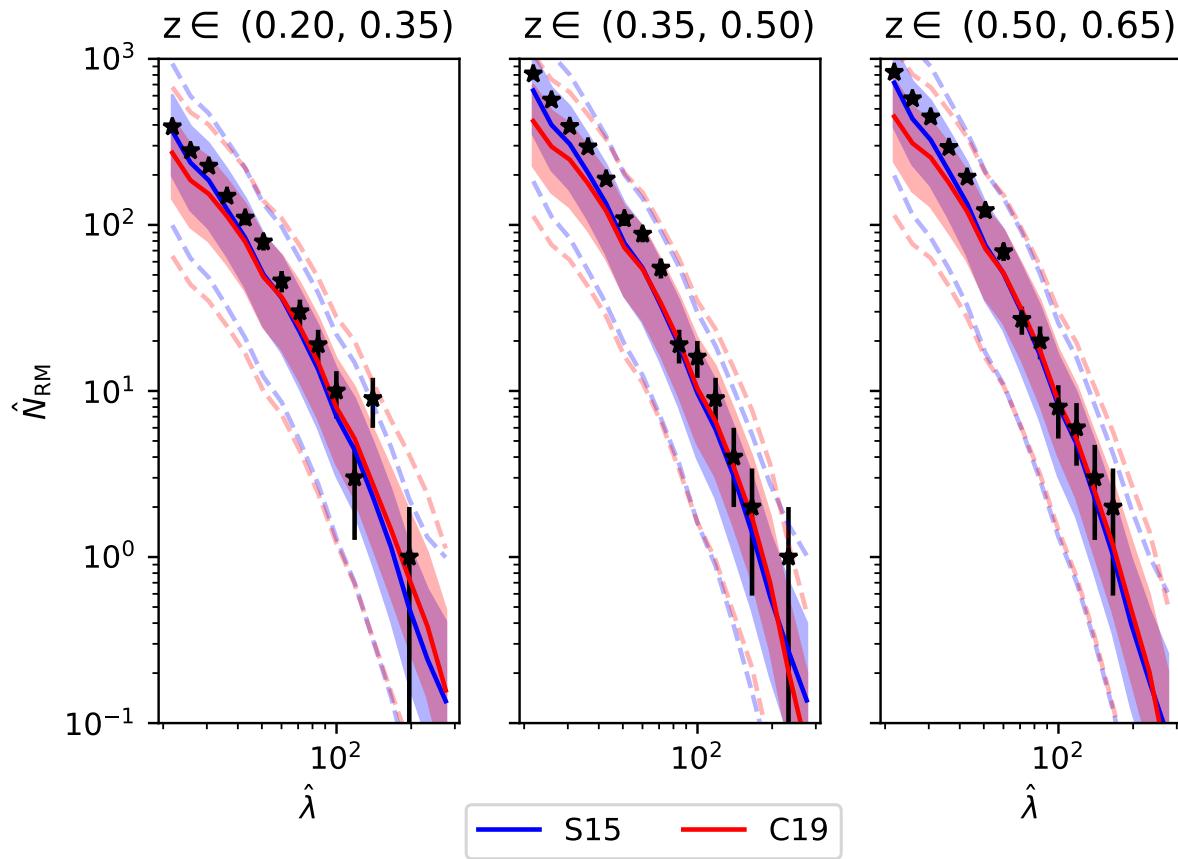


Abbildung 4.9: Number of RM objects in bins of richness and redshift. In color the prediction from our richness–mass relation constraints and SPT cluster number counts for the different optical error models (blue S15, red C19). The shaded region represents the  $1\sigma$  systematic uncertainty. Over plotted in black the number of RM objects with the associated Poissonian error bar. The prediction matched the data within  $1\sigma$  systematic uncertainty.

## Number counts

We also compare the number of RM clusters predicted by our error model to the measured number of cluster in richness and redshift bins. To this end, we predict the number of RM clusters following equation 4.19. As stated above, we do not only propagate the uncertainty on the richness–mass scaling relation parameters, but also on the cosmological parameters. This results in a prediction the number of objects with large systematic uncertainties, as shown in the red (C19) and blue (S15) bands in Fig. 4.9. In the same figure we also plot as black points with error bars the number of RM objects with their statistical Poisson uncertainties, which are considerably smaller than the systematic uncertainties for almost all richnesses. Noticeably, the data is in good agreement with our prediction, independently on the error model. SPT derived cosmology and our richness–mass relation are thus able to predict the number counts of RM objects. This in turn means that if direct mass information was to confirm our calibration of the richness–mass relation, the combination of that calibration and the RM number counts would provide a cosmology that is consistent with SPT constraints.

## 4.4 Discussion

In the following, we discuss the major results presented above. Based on a comparison of literature results, we confirm that our recovered scatter values are consistent with those from other works. Thus, larger scatter does not provide an alternative for the mismatch between our scaling relation and the stacked weak lensing results. We then review the measurements of the mass trend linking it to our measurement of contamination. Thereafter, we discuss other possible sources of the tension between between SPT cross calibration and stacked weak lensing. Finally, we outline upcoming improvements in observation that likely will help to discriminated among the proposed scenarios.

### 4.4.1 Comparison to literature

Our comparison to the literature focusses on the two scaling relation parameters which are most closely linked to the purity, that is the intrinsic scatter and the mass slope. In the following, we will consider also results based on RM on the Sloan Digital Sky Survey (SDSS, see Rykoff et al., 2014, for the discussion of the RM application), as those richnesses are consistent with the richnesses extracted from (DES see, McC19, equations 66-67).

### Constraints on scatter

A larger intrinsic scatter increases the range of masses associated to a given measured richness, thereby reducing the amount of objects which are classified as contamination due to their low mass. This is especially true for optical cluster finding, as every red galaxy lives in a halo of some mass (Cohn et al., 2007; Farahi et al., 2016). The presence of a red galaxy thus indicates the presence of a halo. This situation is different in the SZ, where

noise fluctuation are not necessarily associated to real halos, or in the X-ray, where AGNs, although themselves living in halos, do not display the same luminosity–mass relation as clusters.

Measurements of the scatter are made all the more relevant, as traditionally stacked weak lensing is used in studies of optically selected clusters (e.g. Simet et al., 2017; Murata et al., 2018; McClintock et al., 2019a; Murata et al., 2019). Due to the stacking, such studies loose leverage on the scatter, while Bayesian population modeling retains some constraining power (Grandis et al., 2019). Cross calibration with ICM based mass proxies, like the one done in this work, are thus used to constrain the scatter. Rozo et al. (2015) cross calibrated the SDSS-RM sample (Rykoff et al., 2014) with the Planck 13 SZ selected cluster catalog (Planck Collaboration et al., 2015), investigating the scaling between the richness and the SZ-inferred mass for 191 cluster. The scatter around that relation is  $\sigma_{\ln \lambda | M_{\text{SZ}}} = 0.266 \pm 0.017$ . Investigating the relation between the DES-RM richness and the temperature in 58 archival *Chandra* observations and 110 XMM observation, Farahi et al. (2019) find  $\sigma_{\ln \lambda | M} = 0.20^{+0.10}_{-0.08}$ . Both of these measurements are in good agreement with our results, for instance  $\sigma_\lambda = 0.22 \pm 0.06$ . We furthermore confirm that the impact of the projection model is small compared to the current uncertainties. Also deviation from log-normality as probed by the outlier fraction are not detected. Note, however, that all these results were derived on the high richness end of the population. It remains open to see if they can be extrapolated to lower richnesses. As such, the variance and the deviation from log-normality in the range  $\hat{\lambda} < 40$  are currently, to the authors knowledge, unknown.

### Mass trend of richness

The mis-match between our prediction for the mean mass in richness-redshift bins and its measurement trough stacked weak lensing by McC19 can also be stated as a difference in mass trend. Indeed, McC19 finds a mass trend of  $B_{\lambda, \text{McC19}} = 0.73 \pm 0.03$ . The tension with our results is representative of the varying results reported in recent years in different works (see McClintock et al., 2019a; Capasso et al., 2019c, for summaries). Given the variety of methods employed we propose to separate the works along the following criterion: is the an ICM based selection, that is SZ or X-ray, involved in the sample selections, or is the selection exclusively optical. Note also that we will ignore works that used the number counts at fixed cosmology to determine the mass slope, such as Murata et al. (2018), as the choice of cosmology is directly degenerate with the resulting slope. In the following and in Fig. 4.10 we present a representative selection. In the class of works based on ICM selection we find:

- Rozo et al. (2015) cross calibrating of SDSS-RM with Planck 13 SZ selected clusters finding  $B_\lambda = 0.965 \pm 0.067$ ;
- Mantz et al. (2016) determining  $\lambda \sim M_{\text{gas}}^{0.75 \pm 0.12}$  on a sample ROSAT selected cluster. Combining this with the scaling  $M_{\text{gas}} \sim M^{1.29 \pm 0.09}$  for the gas mass content of SPT selected clusters (Chiu et al., 2018) results in  $B_\lambda = 0.97 \pm 0.17$ . Note here that this inference is based upon the relation between halo mass and gas mass. It is generally

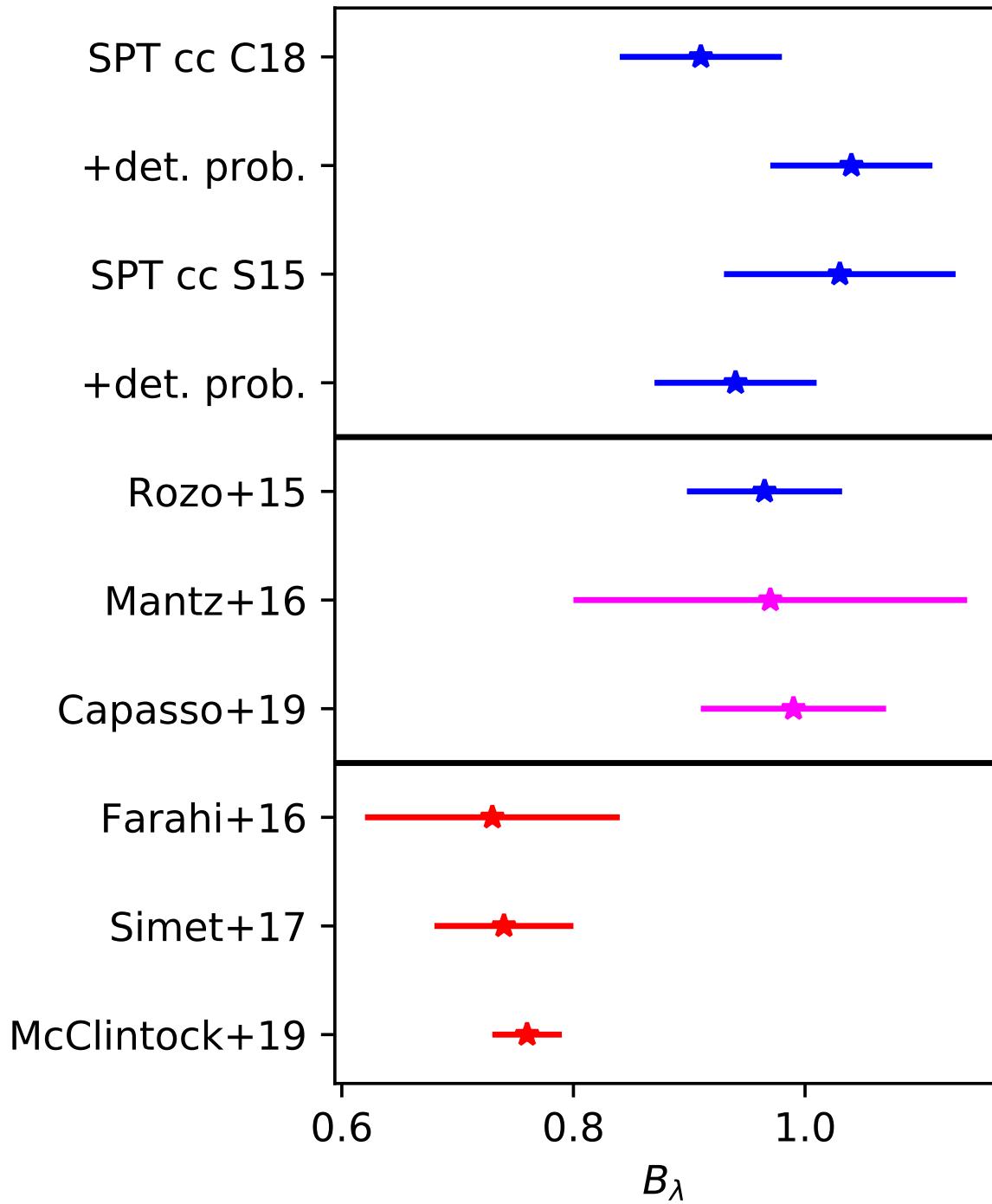


Abbildung 4.10: Comparison of the constraints on the richness-mass slope from different works to our own results. The color coding represents the selection method used: blue for SZe, magenta for X-rays and red for solely optical selection. The discrepancy between the results based on ICM selection (SZe and X-rays) can clearly be seen. This difference might be due to low richness contamination in analyses based on optical selection.

accepted from an empirical and theoretical prospective that the gas mass fraction of low mass systems falls, see for instance the discussion in Bulbul et al. (2019, section 5.2.2.);

- a dynamical analysis of ROSAT selected, SDSS-RM confirmed cluster by Capasso et al. (2019c), reporting  $B_\lambda = 0.99 \pm 0.08$ .

On the other side, constraints obtained on purely optically selected samples are:

- Stacked spectroscopic analysis of SDSS-RM cluster by Farahi et al. (2016), which finds  $B_\lambda = 0.76 \pm 0.11$ .
- Stacked weak lensing of SDSS-RM clusters by Simet et al. (2017), with  $B_\lambda = 0.74 \pm 0.06$ ;
- Stacked weak lensing of DES-RM clusters by McC19 with  $B_\lambda = 0.73 \pm 0.03$ ;
- Stacked analysis of the CMB lensing signal around DES-RM clusters by Baxter et al. (2018) currently do not constrain the mass trend of richness, as they recover the prior imposed on that parameter with no improvement on the uncertainty.

Noticeably, constraints on ICM selected samples are in good agreement with our own results and generally indicate  $B_\lambda \sim 1$ , while results based in stacking methods of optically selected samples suggest  $B_\lambda \sim 0.75$ , as can be clearly seen in Fig. 4.10. As discussed above, this difference could be caused by an increased contamination by low mass systems at low richness which would bias the stacked signals low due to the inclusion of contaminants into the stacks. Caution is however warranted, as in most cases the ICM selected measurements did not reach the richness range  $\hat{\lambda} < 40$ , were conversely most of the objects in the stacked, optically selected analyses lay.

#### 4.4.2 Inferred Purity

Another way of explore the richness range  $\hat{\lambda} < 40$ , is to extrapolate our results. Both the comparison of predicted and measured stack weak lensing mass, as well as the number of RM objects can be turned into an estimated of the purity of the sample, as discussed in section 4.2.5. The results of this prediction are shown in Fig. 4.11 as points for the estimation from number counts and boxes for stacked weak lensing. The color encodes the optical error model used: red for C19 and blue for S15. In both cases the error bars reflect both systematic and statistical uncertainties. In the case of stacked weak lensing the two contributions are of approximately the same magnitude, while the error bar for the number counts estimation are dominated by systematic uncertainties. The two predictions are in mutual agreement. The overestimation of the mean mass compared to the weak lensing is reflected in an estimated purity that deviates at more than  $2\sigma$  from unity.

This provides some support to the idea that the purity of the RM sample might worsen when going to low richness and high redshift. Noticeably, it does so in a manner that is

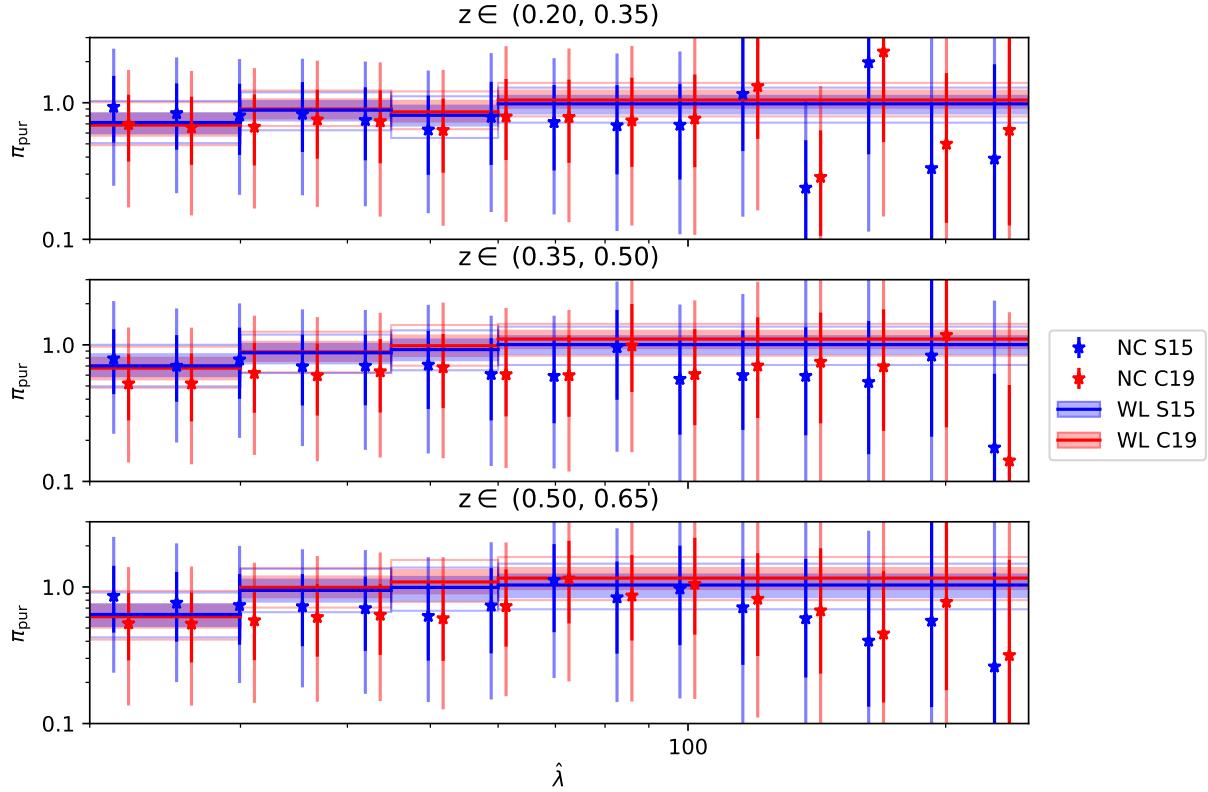


Abbildung 4.11: Estimation of the purity  $\pi_{\text{pur}}$  as a function of richness for different redshift bins. In red the prediction from the C19 error model, while in blue the prediction for the S15 model. Points represent the estimation from number counts, with error bars reflecting the systematic and statistical uncertainties (the latter are however subdominant). As shaded regions the estimates from stacked weak lensing. While the two estimates are in mutual agreement, the purity from stacked weak lensing deviated from 1 at more than  $2\sigma$  for low richness and high redshifts.

consistent between number counts and weak lensing. Assuming up to 30% contamination at low richness would make our richness–mass relation consistent with stacked WL measurements, while being compatible within the systematic uncertainties with the measured number of clusters. It is thus a natural conclusion from extrapolating our richness–mass relation to low richness.

Similar studies by the (Dark Energy Survey Collaboration, in prep.) distinguish themselves by smaller systematic uncertainties, as the richness–mass relation is inferred from the RM number counts itself. In that work, the authors suggest by an analytical argument, that stronger projection effects, being equivalent to contamination by low mass object can not account for all of the observed tension. Empirical confirmation of the project effect model via spectroscopic studies of SDSS-RM objects in the redshift range  $0.08 < z < 0.12$  by (Myles et al., in prep) are currently undertaken. Furthermore, a follow up of the objects in that redshift range with  $20 < \hat{\lambda} < 30$  by (von der Linden et al., in prep.) in the X-ray with *Swift* will also shed light on the mass distribution at low richnesses.

#### 4.4.3 Alternative Explanations

Contamination of the low richness range by even lower mass systems is not the only possible explanation. Alternatively, one could consider the mass range covered by the ICM selected samples, as shown exemplary in Fig. 4.12. Given current sensitivities, it does not extend to such low masses as the optical selection. Indeed, our cross calibration work is limited to  $\hat{\lambda} > 40$ . Considering the different mass ranges, the stacking results might be dominated by the more frequent low mass systems. A mass dependent slope might thus describe the results from both the ICM selected and the purely optically selected samples. The main difficulty of this scenario is to accommodate existing constraints on the stellar mass–halo mass relation, indicating a power law trend down to halo masses of  $M_{\text{halo}} \sim 10^{12} M_{\odot}$  (Behroozi et al., 2019). Nevertheless, comparison to these works is hindered by the amount of stellar mass stripped from galaxies into the intra-cluster light and by the definition of the extraction aperture in the direct RM extraction, which is optimized for selection and not to trace the virial region.

Another possible source of the disagreement between stacked weak lensing results and ICM based cross calibration can be the correlation between the scatter sources of the weak lensing signal and the scatter sources of richness. Both are, for instance affected by miscentering (for the effect of richness see Rykoff et al. (2014), while for weak lensing consider, among others Yang et al. (2006) and Johnston et al. (2007)). Furthermore, at the same halo mass triaxiality impacts both observables similarly (Dietrich et al., 2014). Finally, a part of the lensing signal comes from the two halo term describing the correlation between halos and neighboring structure. This signal is known to display large scatter based on the alignment of the surrounding structure (Osato et al., 2018). It could, furthermore correlate with the percolation procedure in the RM extraction or with the strength of projection effects. Importantly, to explain the current discrepancy on the mass trend, these effects need to be mass dependent, as also pointed out in (Dark Energy Survey Collaboration, in prep.). These effects are currently studied on simulations in preparation for DES-Y3, and

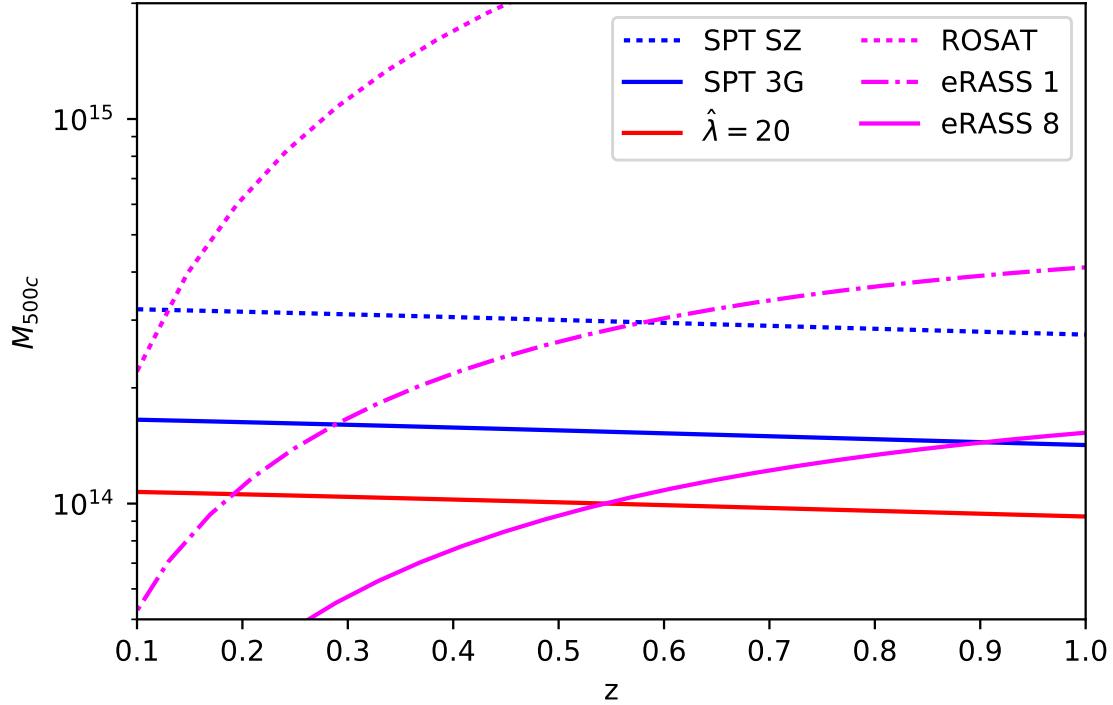


Abbildung 4.12: Current (dotted lines) and future (full and dashed-dotted) mass limits in magenta for X-ray surveys and in blue for SZe surveys. Red the mass corresponding to  $\hat{\lambda} > 20$ . Deeper photometric surveys do not lower the mass limit of optical selection but enable detections to higher redshift. As can be seen in the projected mass limits, eROSITA and SPT-3G will extent the ICM leverage on low mass systems significantly.

LSST. Inference from simulations in these cases is only reliable if the simulations are able to reproduce a realistic distribution of galaxy colors as a function of environmental density (Cohn et al., 2007; Farahi et al., 2016; Costanzi et al., 2019). Even if the prediction was right, simulations still struggle to give a realistic account of the systematic uncertainties induced by the choices in sub grid physics or post-processing. From an empirical perspective, the mass limits of ICM based cluster surveys currently make such studies impossible.

#### 4.4.4 Prospects

As mentioned earlier, direct measurement of contamination in the richness range  $20 < \hat{\lambda} < 40$ , of the mass trend of those richnesses and of the correlation between the scatter in those richnesses and the scatter weak lensing signal currently lays below the sensitivity of X-ray and SZe based cluster surveys. This can be seen clearly by the mass limits of current ICM based cluster surveys, e.g. SPT-SZ (Bleem et al., 2015) and ROSAT (Böhringer et al., 2004) in Fig. 4.12. This situation is however going to change in the next years. In SZ, the third

generation of SPT detectors are currently performing the SPT-3G survey over  $1500 \text{ deg}^2$  (Benson et al., 2014). It will lower the limiting mass of SZ detection by a factor of 2. In X-rays the recent ‘first light’<sup>1</sup> of the eROSITA<sup>2</sup> X-ray telescope (Predehl et al., 2010; Merloni et al., 2012) on board the Russian ‘Spectrum-Roentgen-Gamma’ satellite will drastically improve the sensitivity of X-ray cluster surveys. We present in Fig. 4.12 the mass limits for 15 counts in the first eROSITA All-Sky-Survey (eRASS 1, 0.5 yr of observation) and eighth eROSITA All-Sky-Survey (eRASS 8, 4 yr of observation) following the prediction by Grandis et al. (2019). This would allow us to follow up optically selected objects up to  $z \sim 0.5$  individually. The presence or absence of X-ray emission in eROSITA from optically selected objects will be a powerful tool to extent the kind of analysis presented in this work to lower richesses. It will most likely help to discriminate the different scenarios we identified earlier.

## 4.5 Conclusions

In this work we empirically validate the RM selected DES-Y1 survey by cross calibration with SPT-SZ selected clusters. We first limit ourselves to the high richness regime ( $\hat{\lambda} > 40$ ) to avoid optical incompleteness in the SPT confirmation. We produce a matched sample by positional matching between the RM- $(\hat{\lambda} > 40)$  and the SPT-SZ selected clusters (Bleem et al., 2015). On the matched sample, we model the distribution in SZ-e-signal, richness and redshift with a Bayesian cluster population model. The free parameters of this model, the parameters controlling the scaling between richness and mass, and the scatters around this relation, are constrained from our analysis. We employ prior from previous SPT studies on the parameters on the SZ-e signal–mass relation, effectively transferring the vetted SPT mass calibration onto the richness–mass relation.

In an attempt to explore the effect of projection effects on the richness–mass we employ two different error models: the first by Saro et al. (2015) uses the error bars reported in the RM catalog, while the second by Costanzi et al. (2019) which includes projection effects. Our cross calibration of the richness–mass relation and the scatter around it is not significantly affected by the optical error model. Furthermore, it is consistent with previous cross calibration by (Saro et al., 2015).

We then turn to exploiting the information contained in the fact that some RM- $(\hat{\lambda} > 40)$  in the joint SPT-DES Y1 footprint have been matched by SPT-SZ and some not. By extracting the relative SPT field depth at the RM position, we employ the mass information contained in the richness and the SPT selection function to predict the detection probability by SPT. We expand the probability of SPT detection by accounting for RM purity and outlier fractions with respect to the scatter around the mean relation. Comparing these detection probabilities with the actual occurrence of matches constrains the purity and outlier fractions. We find a purity consistent with 1 and outlier fraction consistent with 0. Accounting for projection effects slightly improves the upper limits on these measurements.

<sup>1</sup><http://www.mpe.mpg.de/7362095/news20191022>

<sup>2</sup><http://www.mpe.mpg.de/eROSITA>

The high purity of the RM- $(\hat{\lambda} > 40)$  sample is confirmed by comparing our prediction of the fraction of SPT detections as a function of measured richness with the measured fraction of matches.

We then extrapolate our richness–mass relation to lower richnesses. Comparison with the mean masses in redshift–richness bins reveals a 2 sigma tension between our relation and stack weak lensing measurements by McClintock et al. (2019a), when considering statistical uncertainties from that work and systematic uncertainties from our cross calibration. Our prediction for the number counts of RM objects based on our richness–mass relation and SPT cosmological constraints is consistent within large systematic error bars with the measured number counts. Interpreting the mis-match in mean mass at low richness as up to 30% contamination at low richness diluting the stack weak lensing signal is also consistent with the comparison between the measured and predicted number counts, due to the large systematic error bars on the latter.

In line with the hypothesis of low richness contamination we discuss how ICM selected measurement of the richness–mass trend are coherently different from studies based exclusively on optical selection. While this effect might indeed be due to the larger contamination the purely optically selected samples, we also discuss that to date, the mass sensitivity of ICM selections does not extend to the mass range spanned by the richnesses between  $20 < \hat{\lambda} < 40$ . As such, we can not exclude alternative explanation for the tension between our relation and the stack weak lensing masses. We then discuss some of these possibilities and highlight how upcoming X-ray surveys like eROSITA, and SZe surveys like SPT-3G will improve the ICM based selection sensitivities to probe the mass regime associated to these lower richnesses, making the different hypothesis we discussed empirically discernable.

## 4.6 Appendix

### 4.6.1 Optical completeness of the SPT sample

As discussed above, we employ an SPT sample that has been confirmed with MCMF by imposing a cut in the probability of random superposition between an optical structure and the SZ candidate,  $f_{\text{cont}} < 0.1$ . As discussed by Klein et al. (2019), Klein et al. (in prep.), and Grandis et al. (in prep.[a]), this cut is equivalent to a redshift dependent cut in the MCMF richness  $\hat{\lambda}_{\text{MCMF}} > \lambda_{\min}(z)$ . Note that the richness  $\hat{\lambda}_{\text{MCMF}}$  extracted by MCMF is not identical to the RM richness  $\hat{\lambda} = \hat{\lambda}_{\text{RM}}$ , as MCMF employs a prior from the SZ candidate on the position and the aperture. The ratio between the two richnesses on the matched sample is shown in Fig. 4.13. When fitting this relation we find

$$\frac{\hat{\lambda}_{\text{MCMF}}}{\hat{\lambda}_{\text{RM}}} = (1.07 \pm 0.02) \left( \frac{E(z)}{E(0.6)} \right)^{-0.24 \pm 0.02}, \quad (4.24)$$

with intrinsic log-normal scatter  $0.23 \pm 0.01$ .

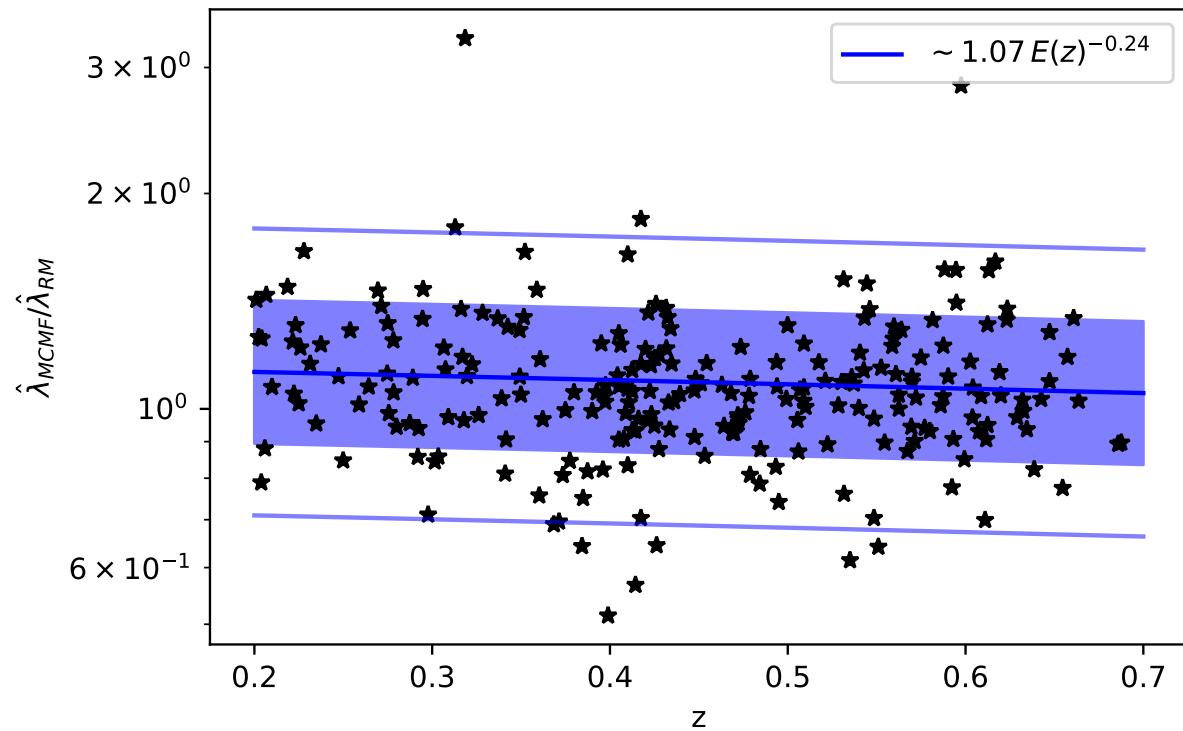


Abbildung 4.13: Ratio between the MCMF richness  $\hat{\lambda}_{MCMF}$  employed in the optical confirmation of SPT-SZ cluster to the RM richness  $\hat{\lambda}_{RM}$  as a function redshift for the matched sample. In blue the mean relation between the two with the intrinsic scatter show by the shaded (1 sigma) and transparent lines (2 sigma). Different centering and apertures lead to intrinsic scatter, a deviation from unity and redshift trend.

Using this relation we estimate that objects with  $\hat{\lambda}_{\text{RM}} > 40$  are always more than 2 sigma above the minimal MCMF richness for all at redshift we consider. This implies at least 97.5% completeness. Every RM- $(\hat{\lambda} > 40)$  object therefore (almost) certainly makes it past the SPT optical confirmation. In the case of the matched sample, where all objects have  $\hat{\lambda} > 40$ , we therefore can safely ignore optical incompleteness in SPT. The same holds for the study of the SPT detection of RM- $(\hat{\lambda} > 40)$ , which we model as solely dependent on the SZe signal. We intentionally omit studying the RM detection probability of SPT objects. The probability of not finding an optically confirmed SPT cluster is given by the probability of  $\hat{\lambda}_{\text{MCMF}} > \lambda_{\min}(z)$  and  $\hat{\lambda} < 40$  at the clusters SZe signal and redshift, and thus depends on the optical incompleteness of the SPT sample.

# Kapitel 5

## Conclusion

After having discussed in detail the individual works that make up this thesis, we now summarize in a more overarching manner the scientific results of these works. Taken together, in this thesis we present several contributions to the field of observational cluster cosmology:

- We demonstrate how to extract cosmological information from an X-ray selected cluster survey while simultaneously calibrating an observable–mass relation with weak lensing measurements in an empirical fashion. The relation is intentionally chosen to be possibly similar to the selection criterion and requiring only a minimal amount of data processing. This method has the promise to reduce the amount of post-processing needed to extract mass information from X-ray observations of clusters, thereby reducing the chance of introducing systematic effects.
- We demonstrate a new method to assess the optimal level of mass calibration for cluster number counts experiments. At this level of mass calibration, all of the statistical power of the number counts contributes towards constraining the cosmology. The scaling relation parameters, which are degenerate with the cosmological information, are thus constrained optimally in this regime.
- We discovered and explored a degeneracy between the scaling relation parameters and the distance–redshift relation in the case of X-ray selected, weak lensing calibrated cluster number count experiments. This degeneracy limits X-ray clusters’ potential to explore the expansion of the Universe. It is likely present also in other selection methods, albeit currently not considered.
- We introduce a method to empirically construct the selection function of X-ray selected surveys. This method has the benefit of being completely empirical: it does not rely on image simulations and the resulting assumptions on cluster morphology. The free parameters of this model, which are calibrated in the data directly, provide an indicator of possible internal tensions. This is a valuable asset to assess the level of internal consistency of cluster cosmological inference.

- We demonstrate how to account for the optical incompleteness introduced by strong optical cleaning of highly contaminated cluster candidate lists. This opens the way to significantly increase the depth of SZe and X-ray cluster surveys, allowing us to construct larger samples extending to higher redshift.
- We further developed and successfully applied a coherent formalism to fit cluster number counts as a function of different combinations of observables. Besides the main X-ray observable, we also fitted number counts in the optical follow-up observable richness alone and in both observables. The resulting consistent mass information validates the selection function modelling. In the presence of external direct mass information, this technique can be used to vet ones cosmological results.
- We invented a statistically sound method to empirically determine the contamination, the incompleteness and the outlier fraction of a mass calibration and selection function modeling by cross-matching two cluster samples. This method computes the probabilities for any cluster in one sample to be detected in the other. Contamination, incompleteness and outlier fractions are then measured by comparing the predicted detection probability to the actual occurrence of detections. This method naturally takes account of the uncertainty in the observable–mass relation, which impacts the uncertainty on the detection probabilities. Furthermore, this method is independent on absolute number of clusters and its redshift evolution, making it an ideal validation test for cluster number count experiments.

As outlined above, in this thesis we touch on all major aspects of the extraction of cosmological information from the number of galaxy clusters and their multi-wavelength signatures: calibration of the observable–mass relation, fitting for cosmology, construction of the surveys selection function, and empirical validation thereof. We developed new techniques and improved established ones, in the quest to better control the systematic effects that impact cluster cosmological studies. The plethora of ongoing and future surveys in all the wavelength regimes that we employed is going to provide ample occasion for the application of these methods in the near future.

# Bibliography

Abbott, T. M. C. et al. (2018). “Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing”. In: *Physical Review D* 98.4, 043526, p. 043526. DOI: 10.1103/PhysRevD.98.043526. arXiv: 1708.01530 [astro-ph.CO].

Albrecht, Andreas et al. (2006). “Report of the Dark Energy Task Force”. In: *arXiv e-prints*, astro-ph/0609591, astro-ph/0609591. arXiv: astro-ph/0609591 [astro-ph].

Allen, S. W., A. E. Evrard, and A. B. Mantz (2011). “Cosmological Parameters from Observations of Galaxy Clusters”. In: *Annual Review of Astronomy and Astrophysics* 49, pp. 409–470. DOI: 10.1146/annurev-astro-081710-102514. arXiv: 1103.4829 [astro-ph.CO].

Anders, E. and N. Grevesse (1989). “Abundances of the elements - Meteoritic and solar”. In: *Geochimica et Cosmochimica Acta* 53, pp. 197–214. DOI: 10.1016/0016-7037(89)90286-X.

Applegate, D. E. et al. (2014). “Weighing the Giants - III. Methods and measurements of accurate galaxy cluster weak-lensing masses”. In: *Monthly Notices of the Royal Astronomical Society* 439, pp. 48–72. DOI: 10.1093/mnras/stt2129. arXiv: 1208.0605.

Barnes, D. J. et al. (2017). “The Cluster-EAGLE project: global properties of simulated clusters with resolved galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 471, pp. 1088–1106. DOI: 10.1093/mnras/stx1647. arXiv: 1703.10907.

Bartelmann, M. (1996). “Arcs from a universal dark-matter halo profile.” In: *Astronomy & Astrophysics* 313, pp. 697–702. eprint: astro-ph/9602053.

Baxter, E. J. et al. (2015). “A Measurement of Gravitational Lensing of the Cosmic Microwave Background by Galaxy Clusters Using Data from the South Pole Telescope”. In: *The Astrophysical Journal* 806, 247, p. 247. DOI: 10.1088/0004-637X/806/2/247. arXiv: 1412.7521.

Baxter, E. J. et al. (2018). “A measurement of CMB cluster lensing with SPT and DES year 1 data”. In: *Monthly Notices of the Royal Astronomical Society* 476, pp. 2674–2688. DOI: 10.1093/mnras/sty305. arXiv: 1708.01360.

Becker, M. R. and A. V. Kravtsov (2011). “On the Accuracy of Weak-lensing Cluster Mass Reconstructions”. In: *The Astrophysical Journal* 740, 25, p. 25. DOI: 10.1088/0004-637X/740/1/25. arXiv: 1011.1681 [astro-ph.CO].

Behroozi, Peter et al. (2019). “UniverseMachine: The Correlation between Galaxy Growth and Dark Matter Halo Assembly from  $z = 0 - 10$ ”. In: *Monthly Notices of the Royal*

*Astronomical Society*, p. 1134. DOI: 10.1093/mnras/stz1182. arXiv: 1806.07893 [astro-ph.GA].

Benson, B. A. et al. (2013). “Cosmological Constraints from Sunyaev-Zel’dovich-selected Clusters with X-Ray Observations in the First 178 deg<sup>2</sup> of the South Pole Telescope Survey”. In: *The Astrophysical Journal* 763, 147, p. 147. DOI: 10.1088/0004-637X/763/2/147. arXiv: 1112.5435 [astro-ph.CO].

Benson, Bradford et al. (2014). “The South Pole Telescope Cluster Survey”. In: *American Astronomical Society Meeting Abstracts #223*. Vol. 223. American Astronomical Society Meeting Abstracts, p. 135.02.

Bharadwaj, V. et al. (2014). “Intracluster medium cooling, AGN feedback, and brightest cluster galaxy properties of galaxy groups. Five properties where groups differ from clusters”. In: *Astronomy & Astrophysics* 572, A46, A46. DOI: 10.1051/0004-6361/201322684. arXiv: 1402.0868.

Bleem, L. E. et al. (2015). “Galaxy Clusters Discovered via the Sunyaev-Zel’dovich Effect in the 2500-Square-Degree SPT-SZ Survey”. In: *The Astrophysical Journal Supplement* 216, 27, p. 27. DOI: 10.1088/0067-0049/216/2/27. arXiv: 1409.0850.

Bocquet, S. et al. (2015). “Mass Calibration and Cosmological Analysis of the SPT-SZ Galaxy Cluster Sample Using Velocity Dispersion  $\sigma_v$  and X-Ray  $Y_X$  Measurements”. In: *The Astrophysical Journal* 799, 214, p. 214. DOI: 10.1088/0004-637X/799/2/214. arXiv: 1407.2942.

Bocquet, S. et al. (2016). “Halo mass function: baryon impact, fitting formulae, and implications for cluster cosmology”. In: *Monthly Notices of the Royal Astronomical Society* 456, pp. 2361–2373. DOI: 10.1093/mnras/stv2657. arXiv: 1502.07357.

Bocquet, S. et al. (2019a). “Cluster Cosmology Constraints from the 2500 deg<sup>2</sup> SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope”. In: *The Astrophysical Journal* 878.1, 55, p. 55. DOI: 10.3847/1538-4357/ab1f10. arXiv: 1812.01679 [astro-ph.CO].

Bocquet, S. et al. (2019b). “Cluster Cosmology Constraints from the 2500 deg<sup>2</sup> SPT-SZ Survey: Inclusion of Weak Gravitational Lensing Data from Magellan and the Hubble Space Telescope”. In: *The Astrophysical Journal* 878.1, 55, p. 55. DOI: 10.3847/1538-4357/ab1f10. arXiv: 1812.01679 [astro-ph.CO].

Böhringer, H. et al. (2001). “The ROSAT-ESO flux limited X-ray (REFLEX) galaxy cluster survey. I. The construction of the cluster sample”. In: *Astronomy & Astrophysics* 369, pp. 826–850. DOI: 10.1051/0004-6361:20010240. eprint: astro-ph/0012266.

Böhringer, H. et al. (2004). “The ROSAT-ESO Flux Limited X-ray (REFLEX) Galaxy cluster survey. V. The cluster catalogue”. In: *Astronomy & Astrophysics* 425, pp. 367–383. DOI: 10.1051/0004-6361:20034484. arXiv: astro-ph/0405546 [astro-ph].

Böhringer, Hans and Norbert Werner (2010). “X-ray spectroscopy of galaxy clusters: studying astrophysical processes in the largest celestial laboratories”. In: *Astronomy and Astrophysics Reviews* 18.1-2, pp. 127–196. DOI: 10.1007/s00159-009-0023-3.

Boller, T. et al. (2016). “Second ROSAT all-sky survey (2RXS) source catalogue”. In: *Astronomy & Astrophysics* 588, A103, A103. DOI: 10.1051/0004-6361/201525648. arXiv: 1609.09244 [astro-ph.HE].

Borgani, S. et al. (2004). “X-ray properties of galaxy clusters and groups from a cosmological hydrodynamical simulation”. In: *Monthly Notices of the Royal Astronomical Society* 348.3, pp. 1078–1096. DOI: 10.1111/j.1365-2966.2004.07431.x. arXiv: astro-ph/0310794 [astro-ph].

Borm, K. et al. (2014). “Constraining galaxy cluster temperatures and redshifts with eROSITA survey data”. In: *Astronomy & Astrophysics* 567, A65, A65. DOI: 10.1051/0004-6361/201322643. arXiv: 1404.5312.

Brunner, H. et al. (2018). “eROSITA ground operations”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 10699. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 106995G. DOI: 10.1117/12.2315138.

Buchner, J. et al. (2014). “X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue”. In: *Astronomy & Astrophysics* 564, A125, A125. DOI: 10.1051/0004-6361/201322971. arXiv: 1402.0004 [astro-ph.HE].

Bulbul, E. et al. (2019). “X-Ray Properties of SPT-selected Galaxy Clusters at  $0.2 < z < 1.5$  Observed with XMM-Newton”. In: *The Astrophysical Journal* 871, 50, p. 50. DOI: 10.3847/1538-4357/aaf230. arXiv: 1807.02556.

Capasso, R. et al. (2019a). “Galaxy kinematics and mass calibration in massive SZE-selected galaxy clusters to  $z = 1.3$ ”. In: *Monthly Notices of the Royal Astronomical Society* 482, pp. 1043–1061. DOI: 10.1093/mnras/sty2645. arXiv: 1711.09903.

— (2019b). “Galaxy kinematics and mass calibration in massive SZE-selected galaxy clusters to  $z = 1.3$ ”. In: *Monthly Notices of the Royal Astronomical Society* 482, pp. 1043–1061. DOI: 10.1093/mnras/sty2645. arXiv: 1711.09903.

Capasso, R. et al. (2019c). “Mass calibration of the CODEX cluster sample using SPIDERS spectroscopy - I. The richness-mass relation”. In: *Monthly Notices of the Royal Astronomical Society* 486.2, pp. 1594–1607. DOI: 10.1093/mnras/stz931. arXiv: 1812.06094 [astro-ph.CO].

Carlstrom, J.~E. et al. (2011). “The 10 Meter South Pole Telescope”. In: *Publications of the Astronomical Society of the Pacific* 123, pp. 568–581. DOI: 10.1086/659879. arXiv: 0907.4445.

Castorina, E. et al. (2014). “Cosmology with massive neutrinos II: on the universality of the halo mass function and bias”. In: *Journal of Cosmology and Astroparticle Physics* 2, 049, p. 049. DOI: 10.1088/1475-7516/2014/02/049. arXiv: 1311.1212 [astro-ph.CO].

Cavaliere, A. and R. Fusco-Femiano (1976). “X-rays from hot plasma in clusters of galaxies”. In: *Astronomy & Astrophysics* 49, pp. 137–144.

Chiu, I et al. (2016). “Baryon content of massive galaxy clusters at  $0.57 \leq z \leq 1.33$ ”. In: *Monthly Notices of the Royal Astronomical Society* 455, pp. 258–275. DOI: 10.1093/mnras/stv2303. arXiv: 1412.7823.

Chiu, I. et al. (2018). “Baryon content in a sample of 91 galaxy clusters selected by the South Pole Telescope at  $0.2 < z < 1.25$ ”. In: *Monthly Notices of the Royal Astronomical Society* 478, pp. 3072–3099. DOI: 10.1093/mnras/sty1284. arXiv: 1711.00917.

Clerc, N. et al. (2014). “The XMM-LSS survey: the Class 1 cluster sample over the extended  $11 \text{ deg}^2$  and its spatial distribution”. In: *Monthly Notices of the Royal Astronomical Society* 444, pp. 2723–2753. DOI: 10.1093/mnras/stu1625. arXiv: 1408.6325.

Clerc, N. et al. (2018). “Synthetic simulations of the extragalactic sky seen by eROSITA. I. Pre-launch selection functions from Monte-Carlo simulations”. In: *Astronomy & Astrophysics* 617, A92, A92. DOI: 10.1051/0004-6361/201732119. arXiv: 1806.08652.

Cohn, J. D. et al. (2007). “Red-sequence cluster finding in the Millennium Simulation”. In: *Monthly Notices of the Royal Astronomical Society* 382.4, pp. 1738–1750. DOI: 10.1111/j.1365-2966.2007.12479.x. arXiv: 0706.0211 [astro-ph].

Cooke, R. J. et al. (2014). “Precision Measures of the Primordial Abundance of Deuterium”. In: *The Astrophysical Journal* 781, 31, p. 31. DOI: 10.1088/0004-637X/781/1/31. arXiv: 1308.3240.

Costanzi, M. et al. (2013). “Cosmology with massive neutrinos III: the halo mass function and an application to galaxy clusters”. In: *Journal of Cosmology and Astroparticle Physics* 12, 012, p. 012. DOI: 10.1088/1475-7516/2013/12/012. arXiv: 1311.1514 [astro-ph.CO].

Costanzi, M. et al. (2019). “Modelling projection effects in optically selected cluster catalogues”. In: *Monthly Notices of the Royal Astronomical Society* 482, pp. 490–505. DOI: 10.1093/mnras/sty2665. arXiv: 1807.07072.

Dark Energy Survey Collaboration (in prep.).

de Haan, T. et al. (2016). “Cosmological Constraints from Galaxy Clusters in the 2500 Square-degree SPT-SZ Survey”. In: *The Astrophysical Journal* 832, 95, p. 95. DOI: 10.3847/0004-637X/832/1/95. arXiv: 1603.06522.

Dietrich, J. P. et al. (2019). “Sunyaev-Zel’dovich effect and X-ray scaling relations from weak lensing mass calibration of 32 South Pole Telescope selected galaxy clusters”. In: *Monthly Notices of the Royal Astronomical Society* 483, pp. 2871–2906. DOI: 10.1093/mnras/sty3088. arXiv: 1711.05344.

Dietrich, Jörg P. et al. (2014). “Orientation bias of optically selected galaxy clusters and its impact on stacked weak-lensing analyses”. In: *Monthly Notices of the Royal Astronomical Society* 443.2, pp. 1713–1722. DOI: 10.1093/mnras/stu1282. arXiv: 1405.2923 [astro-ph.CO].

Dressler, A. (1984). “The Evolution of Galaxies in Clusters”. In: *Annual Review of Astronomy and Astrophysics* 22, pp. 185–222. DOI: 10.1146/annurev.astro.22.1.185.

Drlica-Wagner, A. et al. (2018). “Dark Energy Survey Year 1 Results: The Photometric Data Set for Cosmology”. In: *The Astrophysical Journal Supplement* 235.2, 33, p. 33. DOI: 10.3847/1538-4365/aab4f5. arXiv: 1708.01531 [astro-ph.CO].

Duffy, A. R. et al. (2008). “Dark matter halo concentrations in the Wilkinson Microwave Anisotropy Probe year 5 cosmology”. In: *Monthly Notices of the Royal Astronomical Society* 390, pp. L64–L68. DOI: 10.1111/j.1745-3933.2008.00537.x. arXiv: 0804.2486.

Eckmiller, H. J., D. S. Hudson, and T. H. Reiprich (2011). “Testing the low-mass end of X-ray scaling relations with a sample of Chandra galaxy groups”. In: *Astronomy &*

*Astrophysics* 535, A105, A105. DOI: 10.1051/0004-6361/201116734. arXiv: 1109.6498.

Efstathiou, G. and J. R. Bond (1999). “Cosmic confusion: degeneracies among cosmological parameters derived from measurements of microwave background anisotropies”. In: *Monthly Notices of the Royal Astronomical Society* 304, pp. 75–97. DOI: 10.1046/j.1365-8711.1999.02274.x. eprint: astro-ph/9807103.

Evrard, A. E. et al. (2002). “Galaxy Clusters in Hubble Volume Simulations: Cosmological Constraints from Sky Survey Populations”. In: *The Astrophysical Journal* 573.1, pp. 7–36. DOI: 10.1086/340551. arXiv: astro-ph/0110246 [astro-ph].

Farahi, A. et al. (2019). “Mass Variance from Archival X-ray Properties of Dark Energy Survey Year-1 Galaxy Clusters”. In: *arXiv e-prints*, arXiv:1903.08042, arXiv:1903.08042. arXiv: 1903.08042 [astro-ph.CO].

Farahi, Arya et al. (2016). “Galaxy cluster mass estimation from stacked spectroscopic analysis”. In: *Monthly Notices of the Royal Astronomical Society* 460.4, pp. 3900–3912. DOI: 10.1093/mnras/stw1143. arXiv: 1601.05773 [astro-ph.CO].

Feroz, F., M. P. Hobson, and M. Bridges (2009). “MULTINEST: an efficient and robust Bayesian inference tool for cosmology and particle physics”. In: *Monthly Notices of the Royal Astronomical Society* 398, pp. 1601–1614. DOI: 10.1111/j.1365-2966.2009.14548.x. arXiv: 0809.3437.

Font-Ribera, A. et al. (2014). “DESI and other Dark Energy experiments in the era of neutrino mass measurements”. In: *Journal of Cosmology and Astroparticle Physics* 5, 023, p. 023. DOI: 10.1088/1475-7516/2014/05/023. arXiv: 1308.4164.

Giannantonio, T. et al. (2012). “Constraining primordial non-Gaussianity with future galaxy surveys”. In: *Monthly Notices of the Royal Astronomical Society* 422, pp. 2854–2877. DOI: 10.1111/j.1365-2966.2012.20604.x. arXiv: 1109.0958 [astro-ph.CO].

Grandis, S. et al. (2019). “Impact of Weak Lensing Mass Calibration on eROSITA Galaxy Cluster Cosmological Studies - a Forecast”. In: *Monthly Notices of the Royal Astronomical Society*, p. 1738. DOI: 10.1093/mnras/stz1778. arXiv: 1810.10553 [astro-ph.CO].

Grandis, S. et al. (in prep.[a]).

— (in prep.[b]).

Gruen, D. et al. (2019). “Dark Energy Survey Year 1 results: the effect of intraccluster light on photometric redshifts for weak gravitational lensing”. In: *Monthly Notices of the Royal Astronomical Society* 488.3, pp. 4389–4399. DOI: 10.1093/mnras/stz2036. arXiv: 1809.04599 [astro-ph.CO].

Gupta, N. et al. (2017). “High-frequency cluster radio galaxies: luminosity functions and implications for SZE-selected cluster samples”. In: *Monthly Notices of the Royal Astronomical Society* 467, pp. 3737–3750. DOI: 10.1093/mnras/stx095. arXiv: 1605.05329.

Haiman, Z., J. J. Mohr, and G. P. Holder (2001). “Constraints on Cosmological Parameters from Future Galaxy Cluster Surveys”. In: *The Astrophysical Journal* 553, pp. 545–561. DOI: 10.1086/320939. eprint: astro-ph/0002336.

Hennig, C. et al. (2017). “Galaxy populations in massive galaxy clusters to  $z = 1.1$ : colour distribution, concentration, halo occupation number and red sequence fraction”. In:

*Monthly Notices of the Royal Astronomical Society* 467, pp. 4015–4035. DOI: 10.1093/mnras/stx175. arXiv: 1604.00988.

Hoekstra, H. (2003). “How well can we determine cluster mass profiles from weak lensing?” In: *Monthly Notices of the Royal Astronomical Society* 339, pp. 1155–1162. DOI: 10.1046/j.1365-8711.2003.06264.x. eprint: astro-ph/0208351.

Hoekstra, H. et al. (2013). “Masses of Galaxy Clusters from Gravitational Lensing”. In: *Space Science Reviews* 177, pp. 75–118. DOI: 10.1007/s11214-013-9978-5. arXiv: 1303.3274.

Hofmann, F. et al. (2017). “eROSITA cluster cosmology forecasts: Cluster temperature substructure bias”. In: *Astronomy & Astrophysics* 606, A118, A118. DOI: 10.1051/0004-6361/201730742. arXiv: 1708.05205.

Holder, G., Z. Haiman, and J. J. Mohr (2001). “Constraints on  $\Omega_m$ ,  $\Omega_?$ , and  $\sigma_8$  from Galaxy Cluster Redshift Distributions”. In: *The Astrophysical Journal* 560, pp. L111–L114. DOI: 10.1086/324309. eprint: astro-ph/0105396.

Howlett, C. et al. (2012). “CMB power spectrum parameter degeneracies in the era of precision cosmology”. In: *Journal of Cosmology and Astroparticle Physics* 4, 027, p. 027. DOI: 10.1088/1475-7516/2012/04/027. arXiv: 1201.3654 [astro-ph.CO].

Hu, Wayne (2003). “Self-consistency and calibration of cluster number count surveys for dark energy”. In: *Physical Review D* 67.8, 081304, p. 081304. DOI: 10.1103/PhysRevD.67.081304. arXiv: astro-ph/0301416 [astro-ph].

Jenkins, A. et al. (2001). “The mass function of dark matter haloes”. In: *Monthly Notices of the Royal Astronomical Society* 321.2, pp. 372–384. DOI: 10.1046/j.1365-8711.2001.04029.x. arXiv: astro-ph/0005260 [astro-ph].

Johnston, D. E. et al. (2007). “Cross-correlation Weak Lensing of SDSS galaxy Clusters II: Cluster Density Profiles and the Mass–Richness Relation”. In: *arXiv e-prints*. arXiv: 0709.1159.

Kaiser, N. (1986). “Evolution and clustering of rich clusters.” In: *Monthly Notices of the Royal Astronomical Society* 222, pp. 323–345. DOI: 10.1093/mnras/222.2.323.

Kalberla, P. M. W. et al. (2005). “The Leiden/Argentine/Bonn (LAB) Survey of Galactic HI. Final data release of the combined LDS and IAR surveys with improved stray-radiation corrections”. In: *Astronomy & Astrophysics* 440, pp. 775–782. DOI: 10.1051/0004-6361:20041864. eprint: astro-ph/0504140.

Klein, M. et al. (2018). “A multicomponent matched filter cluster confirmation tool for eROSITA: initial application to the RASS and DES-SV data sets”. In: *Monthly Notices of the Royal Astronomical Society* 474, pp. 3324–3343. DOI: 10.1093/mnras/stx2929.

Klein, M. et al. (2019). “A new RASS galaxy cluster catalogue with low contamination extending to  $z \approx 1$  in the DES overlap region”. In: *Monthly Notices of the Royal Astronomical Society* 488.1, pp. 739–769. DOI: 10.1093/mnras/stz1463. arXiv: 1812.09956 [astro-ph.CO].

Klein, M. et al. (in prep.).

Koester, B. P. et al. (2007). “A MaxBCG Catalog of 13,823 Galaxy Clusters from the Sloan Digital Sky Survey”. In: *The Astrophysical Journal* 660, pp. 239–255. DOI: 10.1086/509599. eprint: astro-ph/0701265.

Kravtsov, Andrey V. and Stefano Borgani (2012). “Formation of Galaxy Clusters”. In: *Annual Review of Astronomy and Astrophysics* 50, pp. 353–409. DOI: 10.1146/annurev-astro-081811-125502. arXiv: 1205.5556 [astro-ph.CO].

Kravtsov, Andrey V., Daisuke Nagai, and Alexey A. Vikhlinin (2005). “Effects of Cooling and Star Formation on the Baryon Fractions in Clusters”. In: *The Astrophysical Journal* 625.2, pp. 588–598. DOI: 10.1086/429796. arXiv: astro-ph/0501227 [astro-ph].

Laureijs, R. et al. (2011). “Euclid Definition Study Report”. In: *ArXiv e-prints*. arXiv: 1110.3193 [astro-ph.CO].

Lee, B. E. et al. (2018). “The relative impact of baryons and cluster shape on weak lensing mass estimates of galaxy clusters”. In: *Monthly Notices of the Royal Astronomical Society* 479, pp. 890–899. DOI: 10.1093/mnras/sty1377. arXiv: 1805.12186.

Levi, M. et al. (2013). “The DESI Experiment, a whitepaper for Snowmass 2013”. In: *ArXiv e-prints*. arXiv: 1308.0847 [astro-ph.CO].

Lin, Y.-T. et al. (2012). “Baryon Content of Massive Galaxy Clusters at  $z = 0\text{--}0.6$ ”. In: *The Astrophysical Journal* 745, L3, p. L3. DOI: 10.1088/2041-8205/745/1/L3. arXiv: 1112.1705.

Liu, J. et al. (2015). “Analysis of Sunyaev-Zel’dovich effect mass-observable relations using South Pole Telescope observations of an X-ray selected sample of low-mass galaxy clusters and groups”. In: *Monthly Notices of the Royal Astronomical Society* 448.3, pp. 2085–2099. DOI: 10.1093/mnras/stv080. arXiv: 1407.7520 [astro-ph.CO].

Lovisari, L., T. H. Reiprich, and G. Schellenberger (2015). “Scaling properties of a complete X-ray selected galaxy group sample”. In: *Astronomy & Astrophysics* 573, A118, A118. DOI: 10.1051/0004-6361/201423954. arXiv: 1409.3845.

LSST DESC (2012). “Large Synoptic Survey Telescope: Dark Energy Science Collaboration”. In: *ArXiv e-prints*. arXiv: 1211.0310 [astro-ph.CO].

LSST DESC et al. (2018). “The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document”. In: *arXiv e-prints*. arXiv: 1809.01669.

Lynden-Bell, D. (1967). “Statistical mechanics of violent relaxation in stellar systems”. In: *Monthly Notices of the Royal Astronomical Society* 136, p. 101. DOI: 10.1093/mnras/136.1.101.

Majumdar, S. and J. J. Mohr (2003). “Importance of Cluster Structural Evolution in Using X-Ray and Sunyaev-Zeldovich Effect Galaxy Cluster Surveys to Study Dark Energy”. In: *The Astrophysical Journal* 585, pp. 603–610. DOI: 10.1086/346179. eprint: astro-ph/0208002.

— (2004). “Self-Calibration in Cluster Studies of Dark Energy: Combining the Cluster Redshift Distribution, the Power Spectrum, and Mass Measurements”. In: *The Astrophysical Journal* 613, pp. 41–50. DOI: 10.1086/422829. eprint: astro-ph/0305341.

Mantz, A. et al. (2010). “The observed growth of massive galaxy clusters - I. Statistical methods and cosmological constraints”. In: *Monthly Notices of the Royal Astronomical Society* 406, pp. 1759–1772. DOI: 10.1111/j.1365-2966.2010.16992.x. arXiv: 0909.3098.

Mantz, A. B. et al. (2014). “Cosmology and astrophysics from relaxed galaxy clusters - II. Cosmological constraints”. In: *Monthly Notices of the Royal Astronomical Society* 440, pp. 2077–2098. DOI: 10.1093/mnras/stu368. arXiv: 1402.6212.

Mantz, A. B. et al. (2015). “Weighing the giants - IV. Cosmology and neutrino mass”. In: *Monthly Notices of the Royal Astronomical Society* 446, pp. 2205–2225. DOI: 10.1093/mnras/stu2096. arXiv: 1407.4516.

Mantz, Adam B. et al. (2016). “Weighing the giants- V. Galaxy cluster scaling relations”. In: *Monthly Notices of the Royal Astronomical Society* 463.4, pp. 3582–3603. DOI: 10.1093/mnras/stw2250. arXiv: 1606.03407 [astro-ph.CO].

Marulli, F. et al. (2018). “The XXL Survey. XVI. The clustering of X-ray selected galaxy clusters at  $z \sim 0.3$ ”. In: *Astronomy & Astrophysics* 620, A1, A1. DOI: 10.1051/0004-6361/201833238. arXiv: 1807.04760.

Mauch, T. et al. (2003). “SUMSS: a wide-field radio imaging survey of the southern sky - II. The source catalogue”. In: *Monthly Notices of the Royal Astronomical Society* 342, pp. 1117–1130. DOI: 10.1046/j.1365-8711.2003.06605.x. eprint: astro-ph/0303188.

McClintock, T. et al. (2019a). “Dark Energy Survey Year 1 results: weak lensing mass calibration of redMaPPer galaxy clusters”. In: *Monthly Notices of the Royal Astronomical Society* 482, pp. 1352–1378. DOI: 10.1093/mnras/sty2711. arXiv: 1805.00039.

McClintock, Thomas et al. (2019b). “The Aemulus Project. II. Emulating the Halo Mass Function”. In: *The Astrophysical Journal* 872.1, 53, p. 53. DOI: 10.3847/1538-4357/aaf568. arXiv: 1804.05866 [astro-ph.CO].

McDonald, M. et al. (2014). “The Redshift Evolution of the Mean Temperature, Pressure, and Entropy Profiles in 80 SPT-Selected Galaxy Clusters”. In: *The Astrophysical Journal* 794.1, 67, p. 67. DOI: 10.1088/0004-637X/794/1/67. arXiv: 1404.6250 [astro-ph.HE].

McDonald, M. et al. (2016). “The Evolution of the Intracluster Medium Metallicity in Sunyaev Zel’dovich-selected Galaxy Clusters at  $0 < z < 1.5$ ”. In: *The Astrophysical Journal* 826, 124, p. 124. DOI: 10.3847/0004-637X/826/2/124. arXiv: 1603.03035.

Merloni, A. et al. (2012). “eROSITA Science Book: Mapping the Structure of the Energetic Universe”. In: *ArXiv e-prints*. arXiv: 1209.3114 [astro-ph.HE].

Mohr, J. J., B. Mathiesen, and A. E. Evrard (1999). “Properties of the Intracluster Medium in an Ensemble of Nearby Galaxy Clusters”. In: *The Astrophysical Journal* 517, pp. 627–649. DOI: 10.1086/307227. eprint: astro-ph/9901281.

Mulroy, S. L. et al. (2019). “LoCuSS: scaling relations between galaxy cluster mass, gas, and stellar content”. In: *Monthly Notices of the Royal Astronomical Society* 484, pp. 60–80. DOI: 10.1093/mnras/sty3484. arXiv: 1901.11276.

Murata, R. et al. (2018). “Constraints on the Mass-Richness Relation from the Abundance and Weak Lensing of SDSS Clusters”. In: *The Astrophysical Journal* 854, 120, p. 120. DOI: 10.3847/1538-4357/aaaab8. arXiv: 1707.01907.

Murata, Ryoma et al. (2019). “The mass-richness relation of optically-selected clusters from weak gravitational lensing and abundance with Subaru HSC first-year data”. In: *arXiv e-prints*, arXiv:1904.07524, arXiv:1904.07524. arXiv: 1904.07524 [astro-ph.CO].

Myles, J. et al. (in prep).

Nagai, D., A. V. Kravtsov, and A. Vikhlinin (2007). “Effects of Galaxy Formation on Thermodynamics of the Intracluster Medium”. In: *The Astrophysical Journal* 668, pp. 1–14. DOI: 10.1086/521328. eprint: arXiv:astro-ph/0703661.

Navarro, J. F., C. S. Frenk, and S. D. M. White (1996). “The Structure of Cold Dark Matter Halos”. In: *The Astrophysical Journal* 462, p. 563. DOI: 10.1086/177173. eprint: astro-ph/9508025.

Osato, Ken et al. (2018). “Strong orientation dependence of surface mass density profiles of dark haloes at large scales”. In: *Monthly Notices of the Royal Astronomical Society* 477.2, pp. 2141–2153. DOI: 10.1093/mnras/sty762. arXiv: 1712.00094 [astro-ph.CO].

Pacaud, F. et al. (2006). “The XMM Large-Scale Structure survey: the X-ray pipeline and survey selection function”. In: *Monthly Notices of the Royal Astronomical Society* 372, pp. 578–590. DOI: 10.1111/j.1365-2966.2006.10881.x. eprint: astro-ph/0607177.

Piffaretti, R. et al. (2011). “The MCXC: a meta-catalogue of x-ray detected clusters of galaxies”. In: *Astronomy & Astrophysics* 534, A109, A109. DOI: 10.1051/0004-6361/201015377. arXiv: 1007.1916.

Pillepich, A., C. Porciani, and T. H. Reiprich (2012). “The X-ray cluster survey with eRosita: forecasts for cosmology, cluster physics and primordial non-Gaussianity”. In: *Monthly Notices of the Royal Astronomical Society* 422, pp. 44–69. DOI: 10.1111/j.1365-2966.2012.20443.x. arXiv: 1111.6587.

Pillepich, A. et al. (2018). “Forecasts on dark energy from the X-ray cluster survey with eROSITA: constraints from counts and clustering”. In: *Monthly Notices of the Royal Astronomical Society* 481, pp. 613–626. DOI: 10.1093/mnras/sty2240. arXiv: 1807.06021.

Planck Collaboration et al. (2014). “Planck 2013 results. XVI. Cosmological parameters”. In: *Astronomy & Astrophysics* 571, A16, A16. DOI: 10.1051/0004-6361/201321591. arXiv: 1303.5076.

Planck Collaboration et al. (2015). “Planck 2013 results. XXXII. The updated Planck catalogue of Sunyaev-Zeldovich sources”. In: *Astronomy & Astrophysics* 581, A14, A14. DOI: 10.1051/0004-6361/201525787. arXiv: 1502.00543 [astro-ph.CO].

Planck Collaboration et al. (2016a). “Planck 2015 results. XIII. Cosmological parameters”. In: *Astronomy & Astrophysics* 594, A13, A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589.

— (2016b). “Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts”. In: *Astronomy & Astrophysics* 594, A24, A24. DOI: 10.1051/0004-6361/201525833. arXiv: 1502.01597.

Pratt, G. W. et al. (2019). “The Galaxy Cluster Mass Scale and Its Impact on Cosmological Constraints from the Cluster Population”. In: *Space Science Reviews* 215, 25, p. 25. DOI: 10.1007/s11214-019-0591-0. arXiv: 1902.10837.

Predehl, P. et al. (2010). “eROSITA on SRG”. In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 7732. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series.

tion Engineers (SPIE) Conference Series. DOI: 10.1117/12.856577. arXiv: 1001.2502 [astro-ph.CO].

Press, W. H. and P. Schechter (1974). "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation". In: *The Astrophysical Journal* 187, pp. 425–438. DOI: 10.1086/152650.

Riess, A. G. et al. (2016). "A 2.4% Determination of the Local Value of the Hubble Constant". In: *The Astrophysical Journal* 826, 56, p. 56. DOI: 10.3847/0004-637X/826/1/56. arXiv: 1604.01424.

Rigault, M. et al. (2013). "Evidence of environmental dependencies of Type Ia supernovae from the Nearby Supernova Factory indicated by local H $\alpha$ ". In: *Astronomy & Astrophysics* 560, A66, A66. DOI: 10.1051/0004-6361/201322104. arXiv: 1309.1182 [astro-ph.CO].

Rigault, M. et al. (2015). "Confirmation of a Star Formation Bias in Type Ia Supernova Distances and its Effect on the Measurement of the Hubble Constant". In: *The Astrophysical Journal* 802, 20, p. 20. DOI: 10.1088/0004-637X/802/1/20. arXiv: 1412.6501.

Rigault, M. et al. (2018). "Strong Dependence of Type Ia Supernova Standardization on the Local Specific Star Formation Rate". In: *ArXiv e-prints*. arXiv: 1806.03849.

Romer, A. K. et al. (2001). "A Serendipitous Galaxy Cluster Survey with XMM: Expected Catalog Properties and Scientific Applications". In: *The Astrophysical Journal* 547, pp. 594–608. DOI: 10.1086/318382.

Rozo, E. et al. (2010). "Cosmological Constraints from the Sloan Digital Sky Survey maxBCG Cluster Catalog". In: *The Astrophysical Journal* 708, pp. 645–660. DOI: 10.1088/0004-637X/708/1/645. arXiv: 0902.3702 [astro-ph.CO].

Rozo, E. et al. (2015). "redMaPPer - III. A detailed comparison of the Planck 2013 and SDSS DR8 redMaPPer cluster catalogues". In: *Monthly Notices of the Royal Astronomical Society* 450.1, pp. 592–605. DOI: 10.1093/mnras/stv605. arXiv: 1401.7716 [astro-ph.CO].

Rozo, Eduardo et al. (2009). "Improvement of the Richness Estimates of maxBCG Clusters". In: *The Astrophysical Journal* 703.1, pp. 601–613. DOI: 10.1088/0004-637X/703/1/601.

Rykoff, E. S. et al. (2014). "redMaPPer. I. Algorithm and SDSS DR8 Catalog". In: *The Astrophysical Journal* 785.2, 104, p. 104. DOI: 10.1088/0004-637X/785/2/104. arXiv: 1303.3562 [astro-ph.CO].

Rykoff, E. S. et al. (2016). "The RedMaPPer Galaxy Cluster Catalog From DES Science Verification Data". In: *The Astrophysical Journal Supplement* 224, 1, p. 1. DOI: 10.3847/0067-0049/224/1/1. arXiv: 1601.00621.

Sánchez, C. et al. (2014). "Photometric redshift analysis in the Dark Energy Survey Science Verification data". In: *Monthly Notices of the Royal Astronomical Society* 445, pp. 1482–1506. DOI: 10.1093/mnras/stu1836. arXiv: 1406.4407 [astro-ph.IM].

Sarazin, Craig L. (1988). *X-ray emission from clusters of galaxies*.

Saro, A. et al. (2015). "Constraints on the richness-mass relation and the optical-SZE positional offset distribution for SZE-selected clusters". In: *Monthly Notices of the Royal*

*Astronomical Society* 454, pp. 2305–2319. DOI: 10.1093/mnras/stv2141. arXiv: 1506.07814.

Sartoris, B. et al. (2016). “Next generation cosmology: constraints from the Euclid galaxy cluster survey”. In: *Monthly Notices of the Royal Astronomical Society* 459, pp. 1764–1780. DOI: 10.1093/mnras/stw630. arXiv: 1505.02165.

Schellenberger, G. and T. H. Reiprich (2017). “HICOSMO: cosmology with a complete sample of galaxy clusters - II. Cosmological results”. In: *Monthly Notices of the Royal Astronomical Society* 471, pp. 1370–1389. DOI: 10.1093/mnras/stx1583. arXiv: 1705.05843.

Schrabback, T. et al. (2018a). “Cluster mass calibration at high redshift: HST weak lensing analysis of 13 distant galaxy clusters from the South Pole Telescope Sunyaev-Zel'dovich Survey”. In: *Monthly Notices of the Royal Astronomical Society* 474, pp. 2635–2678. DOI: 10.1093/mnras/stx2666. arXiv: 1611.03866.

Schrabback, T. et al. (2018b). “Precise weak lensing constraints from deep high-resolution  $K_s$  images: VLT/HAWK-I analysis of the super-massive galaxy cluster RCS2 J 232727.7-020437 at  $z = 0.70$ ”. In: *Astronomy & Astrophysics* 610, A85, A85. DOI: 10.1051/0004-6361/201731730. arXiv: 1711.00475.

Seitz, C. and P. Schneider (1997). “Steps towards nonlinear cluster inversion through gravitational distortions. III. Including a redshift distribution of the sources.” In: *Astronomy & Astrophysics* 318, pp. 687–699. eprint: astro-ph/9601079.

Simet, Melanie et al. (2017). “Weak lensing measurement of the mass-richness relation of SDSS redMaPPer clusters”. In: *Monthly Notices of the Royal Astronomical Society* 466.3, pp. 3103–3118. DOI: 10.1093/mnras/stw3250. arXiv: 1603.06953 [astro-ph.CO].

Skilling, John (2006). “Nested sampling for general Bayesian computation”. In: *Bayesian Anal.* 1.4, pp. 833–859. DOI: 10.1214/06-BA127. URL: <https://doi.org/10.1214/06-BA127>.

Smail, I., R. S. Ellis, and M. J. Fitchett (1994). “Gravitational Lensing of Distant Field Galaxies by Rich Clusters - Part One - Faint Galaxy Redshift Distributions”. In: *Monthly Notices of the Royal Astronomical Society* 270, p. 245. DOI: 10.1093/mnras/270.2.245. eprint: astro-ph/9402048.

Smith, R. K. et al. (2001). “Collisional Plasma Models with APEC/APED: Emission-Line Diagnostics of Hydrogen-like and Helium-like Ions”. In: *The Astrophysical Journal* 556, pp. L91–L95. DOI: 10.1086/322992. eprint: astro-ph/0106478.

Stern, C. et al. (2019). “Weak-lensing analysis of SPT-selected galaxy clusters using Dark Energy Survey Science Verification data”. In: *Monthly Notices of the Royal Astronomical Society* 485, pp. 69–87. DOI: 10.1093/mnras/stz234. arXiv: 1802.04533.

Sunyaev, R. A. and Y. B. Zeldovich (1972). “The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies”. In: *Comments on Astrophysics and Space Physics* 4, p. 173.

Tanabashi, M. et al. (2018). “Review of Particle Physics”. In: *Phys. Rev. D* 98 (3), p. 030001. DOI: 10.1103/PhysRevD.98.030001. URL: <https://link.aps.org/doi/10.1103/PhysRevD.98.030001>.

Tinker, J. et al. (2008). “Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality”. In: *The Astrophysical Journal* 688, 709–728, pp. 709–728. DOI: 10.1086/591439. arXiv: 0803.2706.

Tozzi, Paolo and Colin Norman (2001). “The Evolution of X-Ray Clusters and the Entropy of the Intracluster Medium”. In: *The Astrophysical Journal* 546.1, pp. 63–84. DOI: 10.1086/318237. arXiv: astro-ph/0003289 [astro-ph].

Vanderlinde, K. et al. (2010). “Galaxy Clusters Selected with the Sunyaev-Zel’dovich Effect from 2008 South Pole Telescope Observations”. In: *The Astrophysical Journal* 722, pp. 1180–1196. DOI: 10.1088/0004-637X/722/2/1180. arXiv: 1003.0003.

Vikhlinin, A. et al. (1998). “A Catalog of 200 Galaxy Clusters Serendipitously Detected in the ROSAT PSPC Pointed Observations”. In: *The Astrophysical Journal* 502, pp. 558–581. DOI: 10.1086/305951. eprint: astro-ph/9803099.

Vikhlinin, A. et al. (2009a). “Chandra Cluster Cosmology Project. II. Samples and X-Ray Data Reduction”. In: *The Astrophysical Journal* 692, pp. 1033–1059. DOI: 10.1088/0004-637X/692/2/1033. arXiv: 0805.2207.

Vikhlinin, A. et al. (2009b). “Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints”. In: *The Astrophysical Journal* 692, pp. 1060–1074. DOI: 10.1088/0004-637X/692/2/1060. arXiv: 0812.2720.

von der Linden, A. et al. (in prep.).

Werner, N. et al. (2007). “Complex X-ray morphology of Abell 3128: a distant cluster behind a disturbed cluster”. In: *Astronomy & Astrophysics* 474, pp. 707–716. DOI: 10.1051/0004-6361:20077990. arXiv: 0708.3253.

Yang, X. et al. (2006). “Weak lensing by galaxies in groups and clusters - I. Theoretical expectations”. In: *Monthly Notices of the Royal Astronomical Society* 373, pp. 1159–1172. DOI: 10.1111/j.1365-2966.2006.11091.x. eprint: astro-ph/0607552.

Zuntz, J. et al. (2018). “Dark Energy Survey Year 1 results: weak lensing shape catalogues”. In: *Monthly Notices of the Royal Astronomical Society* 481, pp. 1149–1182. DOI: 10.1093/mnras/sty2219. arXiv: 1708.01533.

# Lebenslauf: Sebastian Grandis

## Nationalities

German, Italian, Swiss

## Education

PhD in Astronomy and Astrophysics, Department of Physics, Ludwig-Maximilian University (LMU) Munich, since Oct 2015-present Cosmological Studies with Galaxy Clusters at X-ray, optical and millimeter wavelengths Supervisor: Prof. Dr. J. J. Mohr, co-supervisors: Dr. S. Bocquet, Dr. A. Saro, Dr. J. P. Dietrich

Master in Physics, Swiss Federal Institute of Technology Zurich (ETH Zurich), 2014-2015 Passed with distinction, marks: 5.84 out of 6 Thesis: Information Gains from Cosmological Probes, supervisor: Prof. Dr. A. Refregier

Bachelor in Physics, Swiss Federal Institute of Technology Zurich (ETH Zurich), 2011-2014 Passed with distinction, marks: 5.87 out of 6

## Teaching Experience

Teaching assistant 'Cosmology and Large Scale Structure', Faculty of Physics, LMU, 2019

Teaching assistant in Physics, Department of Physics, ETH Zurich, 2014-2015 (4 semesters)

Teaching assistant in Numerical Methods, Department of Applied Mathematics, ETH Zurich, 2013 (2 semesters)

## Research Skills

Analysis of cosmological and astrophysical observation, treatment of systematic effects, Bayesian statistics and applications thereof to astrophysics and cosmology

## Presentations

'Cosmology & ICM Physics with eROSITA', presented at 'ICM Physics and Modelling', Garching, Bavaria, Germany, Oct 2018

'Forecasting Impact of DES Weak Lensing Mass Calibration on eROSITA Cluster Cosmology Constraints', presented at 'Annual Meeting of Astronomische Gesellschaft' (German Astronomical Society), Göttingen, Lower Saxony, Germany, Sep 2017

'Quantifying Tensions in Cosmology', Laboratoire de l'Accélérateur Linéaire

(LAL), Centre Scientifique d'Orsay, Orsay Cedex, France, Jan 2017

'Quantifying Tension between Independent Datasets', presented at 'Challenges in Statistical Inference', International Graduate School of Science and Engineering (IGSSE), Garching, Bavaria, Germany, Nov 2016

**Language Skills**

German (native), Italian (native), English (fluent, IELTS, mark 8 out of 9), French (fluent), Python